UNIFORM $C^{1,\alpha}$-REGULARITY FOR ALMOST-MINIMIZERS OF SOME NONLOCAL PERTURBATIONS OF THE PERIMETER

M. GOLDMAN, B. MERLET, AND M. PEGON

ABSTRACT. In this paper, we establish a $C^{1,\alpha}$-regularity theorem for almost-minimizers of the functional $\mathcal{F}_{\varepsilon,\gamma} = P - \gamma P_\varepsilon$, where $\gamma \in (0,1)$ and P_ε is a nonlocal energy converging to the perimeter as ε vanishes. Our theorem provides a criterion for $C^{1,\alpha}$-regularity at a point of the boundary which is *uniform* as the parameter ε goes to 0. As a consequence we obtain that volume-constrained minimizers of $\mathcal{F}_{\varepsilon,\gamma}$ are balls for any ε small enough. For small ε, this minimization problem corresponds to the large mass regime for a Gamow-type problem where the nonlocal repulsive term is given by an integrable kernel G with sufficiently fast decay at infinity.

CONTENTS

1. Introduction 1
2. Preliminary 7
2.1. Nonlocal perimeter and first variation 7
2.2. Perimeter quasi-minimizing properties of minimizers 12
2.3. Basic properties of the excess 14
2.4. The height bound 14
3. Lipschitz approximation theorem 14
3.1. Lipschitz approximation and harmonic comparison 15
3.2. Proof of Theorem 3.2 18
4. Caccioppoli inequality 23
4.1. A refined quasi-minimality condition 24
4.2. A Caccioppoli-type inequality 30
5. Uniform regularity 33
5.1. Excess decay for $r \lesssim \varepsilon$ 33
5.2. Excess decay for $r \gg \varepsilon$ 34
5.3. $C^{1,\alpha}$-regularity 36
References 37

1. INTRODUCTION

The aim of this paper is to complete the program started in [29, 27] regarding the behavior of minimizers of a variant of Gamow's liquid drop model (see [7]) in the regime of large mass. We are interested in the minimization problem

$$\min\left\{\mathcal{F}_{\gamma,\varepsilon}(E) \coloneqq P(E) - \gamma P_\varepsilon(E) \ : \ E \subseteq \mathbb{R}^n \text{ with } |E| = |B_1|\right\}, \qquad (\mathcal{P})$$

where $n \geq 2$, $\gamma \in (0,1)$, P is the Euclidean perimeter and P_ε is a nonlocal perimeter functional such that $P_\varepsilon \to P$ (both pointwise and in the sense of Γ-convergence) as $\varepsilon \to 0$. More precisely, given a radial function $G : \mathbb{R}^n \mapsto (0,\infty)$ with finite first moment, we define the rescaled kernels G_ε by $G(z) \coloneqq \varepsilon^{-(n+1)}G(\varepsilon^{-1}z)$ for all $z \in \mathbb{R}^n$ and the nonlocal perimeter P_ε by

$$P_\varepsilon(E) \coloneqq \int_{\mathbb{R}^n\times\mathbb{R}^n} |\mathbf{1}_E(x) - \mathbf{1}_E(y)| G_\varepsilon(x-y)\,\mathrm{d}x\,\mathrm{d}y = 2\int_{E\times E^c} G_\varepsilon(x-y)\,\mathrm{d}x\,\mathrm{d}y.$$

2020 *Mathematics Subject Classification.* 28A75, 49Q05, 49Q10, 49Q20.
Key words and phrases. Geometric variational problems, nonlocal isoperimetric problems, nonlocal perimeters, regularity, liquid drop model.

It is well-known that when G is integrable, $(\mathcal{P})$ is indeed equivalent to Gamow's model after appropriate rescaling (see [29] for instance). In particular the regime of small ε in $(\mathcal{P})$ corresponds to large mass in Gamow's model.

The main contribution of this paper is a $C^{1,\alpha}$-regularity theorem for almost-minimizers of $\mathcal{F}_{\varepsilon,\gamma}$ which is *uniform* as ε goes to 0, under suitable assumptions on G. This result is stated further on as Theorem B. To the best of our knowledge, this latter is the first uniform regularity result (in dimension higher than 2) for a problem involving the competition of two local/nonlocal perimeters, where neither of the terms is negligible in front of the other.

In combination with the Fuglede type computations done in [27], we obtain the following characterization of minimizers of $(\mathcal{P})$ for small ε. This extends to any arbitrary space dimension the two-dimensional result [27, Theorem A] which is not based on regularity theory.

Theorem A (Minimality of the unit ball). *Assume that $n \geq 2$, $\gamma \in (0,1)$ and that G satisfies the assumptions (H1) to (H5) described below. Then, there exists $\varepsilon_{\text{ball}} = \varepsilon_{\text{ball}}(n, G, \gamma) > 0$, such that, for every $\varepsilon \leq \varepsilon_{\text{ball}}$, the unit ball is the unique minimizer of $(\mathcal{P})$, up to translations.*

Proof. By [29, Theorem B] (see also [27, proof of Theorem 2.7] about the hypothesis $G \in L^1$), if ε is small enough we have existence of minimizers E_ε for $(\mathcal{P})$. Moreover, still by [29, Theorem B], they converge up to translation to B_1 as $\varepsilon \to 0$. The convergence is meant here both in L^1 for the sets and in the Hausdorff distance for the boundaries. In addition, we have convergence of the perimeters. This yields continuity of the excess (see e.g. [26]). Since B_1 is smooth, this implies that Theorem B may be applied at every point of the boundary of E_ε at a scale R which is uniform in ε (and the point). By a standard covering argument, see e.g. [8] this upgrades the Hausdorff convergence of the boundaries to $C^{1,\alpha}$. In particular, for ε small enough, ∂E_ε are small $C^{1,\alpha}$ graphs over ∂B_1. Then, Theorem A is an immediate consequence of [27, Theorem 3]. □

For our $C^{1,\alpha}$-regularity theorem, we work with a classical notion of almost-minimality for $\mathcal{F}_{\varepsilon,\gamma}$.

Definition 1.1 (Almost-minimizers). Let $\gamma \in (0,1)$ and $\varepsilon > 0$. For any positive constants Λ and r_0, we say that E is a (Λ, r_0)-minimizer of $\mathcal{F}_{\varepsilon,\gamma}$ if for every set of finite perimeter $F \subseteq \mathbb{R}^n$ such that $E \triangle F \subset\subset B_r(x)$ with $0 < r \leq r_0$ and $x \in \mathbb{R}^n$, we have

$$\mathcal{F}_{\varepsilon,\gamma}(E) \leq \mathcal{F}_{\varepsilon,\gamma}(F) + \Lambda |E \triangle F|.$$

We will show in Proposition 2.6 that (volume constrained) minimizers of $(\mathcal{P})$ are indeed (unconstrained) (Λ, r_0)-minimizers of $\mathcal{F}_{\varepsilon,\gamma}$ for any r_0 and some constant Λ, not depending on ε. This type of relaxation of the volume constraint is standard for this kind of problems (see e.g. [30, 17, 18, 14]).

Remark 1.2. We could generalize the above definition to an open subset $\Omega \subseteq \mathbb{R}^n$, imposing that competitors F differ from E only in balls $B_r(x) \subseteq \Omega$. Our arguments work just the same and yield uniform regularity of ∂E in compact subsets of Ω. This applies for instance to sets E which are prescribed outside Ω and minimize $\mathcal{F}_{\varepsilon,\gamma}$ locally in Ω.

For $k \in \mathbb{N}$ and a general kernel K, it is convenient to introduce the k-th moment of K, which is defined by

$$I_K^k := \int_{\mathbb{R}^n} |z|^k |K(z)| \,\mathrm{d}z. \tag{1.1}$$

In this work, G always satisfies the following hypotheses:

(H1) G is a measurable, nonnegative, radial function, that is, there exists $g : (0,\infty) \to [0,\infty)$ such that $G(z) = g(|z|)$ for every $z \in \mathbb{R}^n \setminus \{0\}$;

(H2) $z \mapsto |z|G(z) \in L^1(\mathbb{R}^n)$ and the first moment is normalized by

$$I_G^1 = \frac{1}{\mathbb{K}_{1,n}}, \tag{1.2}$$

where $\mathbb{K}_{1,n} := \fint_{\mathbb{S}^{n-1}} |x_n| \,\mathrm{d}\mathcal{H}^{n-1}$.

We will also use the following additional assumptions on G:

(H3) $G \in W^{1,1}_{\text{loc}}(\mathbb{R}^n \setminus \{0\})$, $I^2_{|\nabla G|} < \infty$, and $|g'(r)| \leq \frac{C}{r^{n+1}}$ for $r \geq 1$;

(H4) $\displaystyle\int_{B_1\setminus B_r} G(z)\,\mathrm{d}z \le \frac{C}{r^{s_0}}$ for every $r \in (0,1)$, for some constants $C > 0$ and $s_0 \in (0,1)$;

(H5) Denoting $Q : \mathbb{R}_+ \to \mathbb{R}_+$ the rate function defined by

$$Q(r) := \int_{\mathbb{R}^n\setminus B_r} |z|G(z)\,\mathrm{d}z, \qquad \forall r \in [0,\infty), \tag{1.3}$$

there holds $Q(r) \le \dfrac{C}{r^{n-1+p_0}}$ for every $r > 0$, for some constants $C > 0$ and $p_0 > 0$.

Let us briefly comment on these assumptions.

(i) (H1) and (H2) are needed to ensure that P_ε is well-defined on sets of finite perimeter and that it converges to the standard perimeter. In particular, it is used in [29] to obtain existence of minimizers for small ε and convergence to the ball as ε vanishes.

(ii) (H3) (in the form of $I^2_{|\nabla G|} < \infty$) is used to compute the first variation of $\mathcal{F}_{\varepsilon,\gamma}$ (see Lemma 2.5). It is also needed in its full version in order to apply the stability result [27, Theorem 3] for nearly spherical sets.

(iii) (H4) states that for small scales the perimeter is dominant over P_ε, leading to regularity at small scales by classical regularity theory for almost minimizers of the perimeter (see Proposition 2.10 and Proposition 5.2). Notice that this assumption, which is weaker than $G \in L^1$, roughly states that close to 0, G behaves at most like the kernel of the s_0-fractional perimeter (see for instance [4, 13, 14]).

(iv) Finally, (H5) is used in the proof of Theorem 5.5 to bridge between the excess decay at large scales obtained in Proposition 5.4 and the excess decay at small scales from Proposition 5.2 (see also the discussion below).

We now also comment on the restrictions these conditions impose on the kernel and give a few examples where these assumptions are satisfied.

(i) With (H2), one can check that assumption (H5) is equivalent to $I_G^{n+q_0} < \infty$ for some positive q_0 (possibly different from p_0).

(ii) If G is a power law function near the origin, that is, $G \propto |z|^{-\alpha}$ for some $\alpha > 0$ in a neighborhood of 0, then (H4) states that $\alpha \le n + s_0$. Notice that in that particular example, $|\cdot|^2\nabla G$ is integrable near the origin, which is a part of (H3).

(iii) If G is a power law function at infinity, that is $G \propto |z|^{-\beta}$ at infinity, (H5) states that $\beta \ge 2n+p_0$. In that particular example, $|\cdot|^2\nabla G$ is integrable at infinity, and $|g'(r)| \le \frac{C}{r^{n+1}}$ when $r \to \infty$, which is the other part of (H3).

(iv) From the two previous points, we see that the kernel G defined by

$$G(z) \propto \min\left(\frac{1}{|z|^{n+s_0}}, \frac{1}{|z|^{2n+p_0}}\right)$$

with $s_0 \in (0,1)$, $p_0 > 0$, satisfies assumptions (H1) to (H5).
Other admissible kernels are multiples of the Bessel kernels $\mathcal{B}_{\alpha,\kappa}$, defined for any $\alpha > 0$ and $\kappa > 0$ as the fundamental solution of the operator $(\mathrm{Id} - \kappa\Delta)^{\frac{\alpha}{2}}$. Indeed, Bessel kernels are smooth away from zero, decay exponentially at infinity and, near the origin

$$\mathcal{B}_{\alpha,\kappa} \propto \begin{cases} \frac{1}{|z|^{n-\alpha}} & \text{for } \alpha \in (0,n), \\ -\log(|z|) & \text{for } \alpha = n, \\ 1 & \text{for } \alpha > n. \end{cases}$$

Let us point out that they correspond to screened Coulomb kernels in the case $\alpha = 2$, see [23]. Eventually, our paper covers the case of integrable and compactly supported kernels (with the extra assumption (H3)), which was first studied in [30].

To state our $C^{1,\alpha}$-regularity theorem and sketch its proof, we need to introduce the notion of (spherical) excess, which measures the variation of the normal vector to the boundary of a set near a point.

For a set of finite perimeter E, we will always implicitly assume that E denotes a well-chosen representative such that its topological boundary ∂E satisfies (see e.g. [26, Proposition 12.19])

$$\partial E = \operatorname{spt}|D\mathbf{1}_E| = \Big\{x \in \mathbb{R}^n \ : \ 0 < |E \cap B_r(x)| < |B_r(x)| \text{ for all } r > 0\Big\}.$$

We denote by $\partial^* E$ the reduced boundary of E, and by $\nu_E(x)$ the outer unit normal to $\partial^* E$ at x.

Definition 1.3 (Spherical excess)**.** For any set of finite perimeter $E \subseteq \mathbb{R}^n$ we define the spherical excess (or simply excess) of E in $x \in \partial E$ at scale $r > 0$ by

$$\mathbf{e}(E, x, r) \coloneqq \inf_{\nu \in \mathbb{S}^{n-1}} \frac{1}{r^{n-1}} \int_{\partial^* E \cap B_r(x)} \frac{|\nu - \nu_E(y)|^2}{2} \, \mathrm{d}\mathcal{H}^{n-1}_y,$$

where we used the short-hand notation $\mathrm{d}\mathcal{H}^{n-1}_y$ for $\mathrm{d}\mathcal{H}^{n-1}(y)$.

When $x = 0$ we simply denote $\mathbf{e}(E, r) = \mathbf{e}(E, 0, r)$ (which we will usually assume by translation invariance of the statements). We can now state our main "epsilon-regularity" theorem.

Theorem B. *Assume that G satisfies (H1) to (H5), and let $\gamma \in (0,1)$ and $\Lambda > 0$. Then there exist positive constants τ_{reg}, $\varepsilon_{\mathrm{reg}}$, $\beta \in (0,1)$, and $\alpha \in (0,1)$ depending only on n, G and γ such that the following holds. Let E be a (Λ, r_0)-minimizer of $\mathcal{F}_{\varepsilon,\gamma}$ with $\varepsilon \in (0, \varepsilon_{\mathrm{reg}})$ and $0 \in \partial E$. Assume that for some $R \in [\varepsilon^{1-\beta}, r_0]$,*

$$\mathbf{e}(E, R) + \Lambda R \leq \tau_{\mathrm{reg}},$$

then, up to a rotation, ∂E coincides in $B_{R/2}$ with the graph of a $C^{1,\alpha}$ function $u : \mathbb{R}^{n-1} \mapsto \mathbb{R}$. Moreover,

$$[\nabla u]^2_{\alpha, \frac{R}{2}} \leq C \left(\frac{1}{R^{2\alpha}} (\mathbf{e}(E, R) + \Lambda R) + 1 \right)$$

for some $C = C(n, G, \gamma) > 0$. Here $[\cdot]_{\alpha,R}$ denotes the Hölder semi-norm in the ball of radius R in $\mathbb{R}^{n-1}$.

Proof. The proof is a standard consequence of the excess decay proven in Theorem 5.5 and Campanato's criterion for Hölder-continuous functions. We refer to [26] for more details. □

Let us stress the fact that by (H4), as explained above, we essentially already know that such minimizers are $C^{1,\alpha}$ regular by the classical regularity theory of almost-minimizers of the perimeter, see Proposition 5.2. The main point of Theorem B is that the estimate holds at a "large" scale R uniformly as ε goes to 0.

We now give the main steps of the proof of Theorem B or more precisely of Theorem 5.5. The overall strategy follows the classical regularity theory for minimizers of the perimeter as pioneered by De Giorgi, Federer, Almgren and Allard to name a few. We follow here the presentation from [26]. The first step is to obtain density upper and lower bounds, both for the volume and the perimeter. This is a direct consequence of a weak quasi-minimality property for (Λ, r_0)-minimizers E of $\mathcal{F}_{\varepsilon,\gamma}$ (see [16, Theorem 5.6] or [9]). Indeed, we prove in Proposition 2.8 that E satisfies

$$P(E; B_r(x)) \leq C P(F; B_r(x))$$

for every F such that $E \triangle F \subset\subset B_r(x)$ with $x \in \mathbb{R}^n$, $0 < \Lambda r \leq 1 - \gamma$ and C depends only on n, G and γ.

The next step is to prove the excess decay itself. To this aim we argue differently for scales $r \geq r_+ \gg \varepsilon$ and $r \leq r_- \ll \varepsilon$. As already explained above, for the latter the nonlocal term is of higher order with respect to the perimeter so that we are able to rely on the classical regularity theory for almost-minimizers of the perimeter, see Proposition 5.2. Before focusing on the scales $r \gg \varepsilon$, let us point out that one difficulty is to bridge the gap between r_+ and r_-. We solve this issue using a relatively naive estimate coming from the scaling properties of the excess, see Proposition 2.12. To compensate the loss introduced at this step we need the excess to have already decayed enough when reaching the scale r_+. This explains both hypothesis (H5) and $R \geq \varepsilon^{1-\beta}$ in Theorem B.

We are thus left with the excess decay for $r \gg \varepsilon$. This is done in Proposition 5.4 and represents most of the work. The proof goes through a Campanato iteration scheme which relies on the improvement of the excess by tilting proven in Lemma 5.3. In turn this Lemma states that if the excess is small at some scale $r \gg \varepsilon$ then up to tilting and an error of the order of $Q(r/\varepsilon)$ (with Q defined by (1.3)), the excess is much smaller at a scale λr for some $\lambda \ll 1$. Notice that as opposed to the usual applications of this idea, here the error term gets larger as r decreases.

Very roughly speaking, the idea of the proof of the tilt Lemma is that for $r \gg \varepsilon$, P_ε is equal at leading order to the perimeter so that we can write $\mathcal{F}_{\varepsilon,\gamma} = (1-\gamma)P + \gamma(P - P_\varepsilon)$ and we can hope to reproduce the classical strategy for the excess decay treating $P - P_\varepsilon$ has a higher order term. This is actually quite delicate since formal computations show that on smooth sets the energy $P - P_\varepsilon$ penalizes curvature rather than volume (or even perimeter). Let us sketch the four main steps of this strategy, where Steps 2 & 4 differ substantially from the case of the perimeter functional. We first need to introduce the cylindrical excess and some more notation. Let

$$\mathbf{C}(x,r,\nu) = x + \Big\{ y + t\nu \;:\; y \in \nu^\perp \text{ such that } |y| < r \text{ and } t \in (-r,r) \Big\} \tag{1.4}$$

denote the (truncated) cylinder centered at $x \in \mathbb{R}^n$ with direction $\nu \in \mathbb{S}^{n-1}$, basis radius r and height $2r$.

Definition 1.4 (Cylindrical excess)**.** For any set of finite perimeter $E \subseteq \mathbb{R}^n$ and any cylinder $\mathbf{C}(x,r,\nu)$ centered at $x \in \partial E$ we define the cylindrical excess of E in $\mathbf{C}(x,r,\nu)$ by

$$\mathbf{e}(E,x,r,\nu) \coloneqq \frac{1}{r^{n-1}} \int_{\partial^* E \cap \mathbf{C}(x,r,\nu)} \frac{|\nu - \nu_E(y)|^2}{2} \,\mathrm{d}\mathcal{H}^{n-1}_y. \tag{1.5}$$

As above, if $x = 0$ we simply denote $\mathbf{C}(r,\nu) = \mathbf{C}(0,r,\nu)$ and $\mathbf{e}(E,r,\nu) = \mathbf{e}(E,0,r,\nu)$. We can now proceed with the sketch of the proof for large scales.

Step 1. We show in Theorem 3.1 that if the excess of a (Λ, r_0)-almost minimizer E of $\mathcal{F}_{\varepsilon,\gamma}$ is small in a cylinder $\mathbf{C}(4r,\nu)$, then $\partial E \cap \mathbf{C}(2r,\nu)$ is almost flat and almost entirely covered by the graph of a Lipschitz function u. As observed in [12], this is based on the so-called *height bound* (see Proposition 2.14) which relies only on the density estimates so that we can directly appeal to [26].

Step 2. In Theorem 3.2, we show that the function u "almost" satisfies an equation of the form $(\Delta - \gamma\Delta_{G_\varepsilon})u = 0$ in $\mathbf{C}(r,\nu)$, where Δ_{G_ε} is a nonlocal operator converging to the Laplacian as $\varepsilon \to 0$. For this part, we proceed as follows:

1. We write the Euler–Lagrange equation (see Lemma 2.5) associated with deformations of E in the direction of ν.
2. Carefully discarding the negligible long-range interaction terms, we "localize" in Lemma 3.4 the equation to the cylinder $\mathbf{C}(2r,\nu)$.
3. In Lemma 3.5, we pass the equation on ∂E to the graph of u using their proximity.
4. We linearize the equation.
5. Eventually, since r is much larger than ε, formally $(\Delta - \gamma\Delta_{G_\varepsilon}) \simeq (1-\gamma)\Delta$ in $\mathbf{C}(r,\nu)$, so that u is close to a harmonic function, see Proposition 3.3.

Step 3. Since u is close to a harmonic function, we show that the *flatness* of E (see Definition 4.1) at some smaller scale λr is much smaller than the excess at scale $4r$, up to tilting the direction (see (5.11)). This part is relatively standard.

Step 4. By analogy with functions, one should think of the excess of E as the Dirichlet energy of u, and think of the flatness of E as the L^2 norm of u. To transfer the smallness of the flatness at scale λr to the excess, we prove in Proposition 4.6 a Caccioppoli-type inequality (or Reverse Poincaré), stating roughly

$$\mathbf{e}(E,\lambda r/2,\nu) \lesssim \mathbf{f}(E,\lambda r,\nu) + \left(\frac{\varepsilon}{\lambda r}\right)^\theta \mathbf{e}(E,\lambda r,\nu) + \text{"smaller terms"}$$

whenever λr is still much larger than ε. Our proof of the Caccioppoli inequality relies on an improved quasi-minimality condition when the set E is already known to be sufficiently flat (see Proposition 4.2). To obtain this improved quasi-minimality, we heavily use the 1D slicing techniques already introduced in [27] and end up having to prove that the unit $(n-1)$-ball B' minimizes a quantity which can be interpreted as the average over all lighting directions of the shadow of the boundary of a set obstructing the cylinder $B' \times \mathbb{R}$ (see (4.18)).

Motivation and related results. As already alluded to, under the additional hypothesis that $G \in L^1$ and after rescaling, $\mathcal{F}_{\varepsilon,\gamma}$ is equivalent to the generalized Gamow functionals (see [5]),

$$\min \Big\{ P(E) + \gamma \int_{E\times E} G(x-y)\,\mathrm{d}x\,\mathrm{d}y \;:\; |E| = m \Big\}. \tag{1.6}$$

Besides the case of compactly supported kernels studied in [30] and for which existence of minimizers holds for any m, the main example studied in the literature is the case of Riesz interaction energies, $G(z) = |z|^{-(n-\alpha)}$ with $\alpha \in (0, n)$. The case $n = 3$ and $\alpha = 2$ corresponding to Gamow's liquid drop model for the atomic nucleus, see [7] for a short overview on this problem. In this problem it has been shown in [22, 14] that for small m minimizers are balls while for large m there is non-existence of minimizers under the assumption that $\alpha \in [n-2, n)$, see [22, 25, 15]. The question of the non-existence of minimizers in the case $\alpha \in (0, n-2)$ is still open. The proof of the rigidity of balls for small m in [22, 14] (see also [5] for the case of quite general kernels G) is of the same spirit as for Theorem A and goes through a Selection Principle along the lines of [8]. Notice however that in that case one can directly rely on the classical regularity theory for quasi-minimizers of the perimeter, see (2.4). Let us point out that in a related direction, it has been shown in [1] that if we replace the Euclidean perimeter in (1.6) by an isotropic but weighted perimeter with a weight growing fast enough at infinity then just as in Theorem A balls are the unique minimizers for large m.

Recently, there has been a growing interest for related nonlocal isoperimetric problems which do not fall within the standard regularity theory for perimeter almost-minimizers. A first example comes from a variational model for charged liquid drops where the kernel is still given by the Riesz interaction kernel but now the charge is not assumed to be uniformly distributed on E. This leads in general to much more singular interactions, see [19]. However, introducing either an additional penalization of the charge as in [12] or restricting to $\alpha \leq 1$ as in [21], it is possible to obtain an ε-regularity theorem in the same spirit as the one for minimal surfaces. A major difference between our setting and [12, 21] is that in the charged liquid drop model, it is possible to show that for smooth sets the nonlocal term is actually a volume term (while for us it penalizes curvature). Another example studied in [28] and which is strongly related to $(\mathcal{P})$, is formally (1.6) with the Riesz kernel but for $n = 2$ and $\alpha = -1$. This is motivated by dipolar repulsion. In order to make the model meaningful, a small-scale cut-off has to be introduced (otherwise the energy is always infinite). This cut-off plays a similar role in that model to our parameter ε. In particular, just like in our case, in the limit of vanishing cut-off length and after proper renormalization, the nonlocal term converges to the perimeter. Among many other things, it is shown in [28] that as in Theorem A, in the sub-critical regime (in our language $\gamma < 1$) minimizers are disks for small but finite cut-off lengths. Just like in our problem, the main issue in [28] is to obtain regularity estimates which hold at a macroscopic scale. However, our strategy to obtain these estimates is very different from [28]. Indeed, while we propagate regularity from the macroscopic scale down to the microscopic scale (in the form of excess decay), [28] relies on the Euler–Lagrange equation to bootstrap the regularity from the microscopic scale up to the macroscopic scale. Let us point out that on the one hand, the proof in [28] does not seem to be easily adapted to dimensions higher than two and that on the other hand the logarithmic scaling in [28] allows to directly pass (in our notation) from a scale $r \ll \varepsilon$ to a scale $r \gg \varepsilon$. We refer to [6, 24, 11] for other related models where however rigidity of the ball has not been investigated.

In conclusion, besides [28] which concerns a two dimensional problem, Theorem B is the first uniform regularity theorem for quasi-minimizers of a functional built with two competing local/nonlocal perimeters which remain asymptotically of the same order.
One may compare our regularity result with the one of [4]. Therein, the authors establish a uniform $C^{1,\alpha}$-regularity result for local minimizers of the s-perimeter which is uniform in s as $s \to 1^-$. However, due to the lack of a competing term, the problem and its analysis are rather different from the ones of the present work.

Outline of the paper. The structure of the paper is as follows. In Section 2, we recall and prove a few facts about nonlocal perimeters as well as some useful results from [29, 27] on minimizers of $(\mathcal{P})$. We then establish uniform density estimates for (Λ, r_0)-minimizers of $\mathcal{F}_{\varepsilon,\gamma}$ and show that minimizers of $(\mathcal{P})$ are almost-minimizers of $\mathcal{F}_{\varepsilon,\gamma}$. Eventually, we recall some basic properties of the excess and argue that almost-minimizers satisfy the height bound property. In Section 3 we prove the Lipschitz approximation theorem at scales which are much larger than ε (Theorems 3.1 and 3.2). In Section 4, we establish the Caccioppoli inequality for (Λ, r_0)-minimizers of $\mathcal{F}_{\varepsilon,\gamma}$. Finally, building upon Sections 3 and 4, Section 5 is devoted to establishing power decay of the excess from large scales down to arbitrarily small scales.

Notation.

We write any point $x \in \mathbb{R}^n$ as $x = (x', x_n)$. We denote by $B_r(x) \subseteq \mathbb{R}^n$ the open ball of radius r in $\mathbb{R}^n$ centered at x. When $x = 0$ we simply write B_r for $B_r(0)$. For open balls in $\mathbb{R}^{n-1}$, we write $D_r(x')$ and simply D_r when $x' = 0$. For any $m \in \mathbb{N}$, ω_m denotes the m-volume of the unit ball in $\mathbb{R}^m$, that is, its Lebesgue measure in $\mathbb{R}^m$.
For any set $E \subseteq \mathbb{R}^n$, we denote by $E^{\mathrm{c}} := \mathbb{R}^n \setminus E$ its complement, and write $|E|$ for its volume whenever E is measurable. For any $m \in \mathbb{N}$ we denote by $\mathcal{H}^m$ the m-dimensional Hausdorff measure in $\mathbb{R}^n$. When integrating with respect to the measure $\mathcal{H}^m$ in a variable x, we use the compact notation $\mathrm{d}\mathcal{H}^m_x$ instead of the standard $\mathrm{d}\mathcal{H}^m(x)$. If A is of dimension m and f is $\mathcal{H}^m$-measurable, we may simply write

$$\int_A f := \int_A f(x)\,\mathrm{d}\mathcal{H}^m_x.$$

Similarly, we sometimes use the notation $f_x := f(x)$.
When $\nu = e_n$ (the n-th vector of the canonical basis of $\mathbb{R}^n$) we write $\mathbf{C}_r(x)$ for the cylinder $\mathbf{C}(x, r, \nu)$ (recall (1.4)) and simply $\mathbf{C}_r$ if in addition $x = 0$. We also write $\mathbf{e}_n(E, x, r)$ for $\mathbf{e}(E, x, r, e_n)$ (recall (1.5)) and for $x = 0$, $\mathbf{e}_n(E, r) = \mathbf{e}_n(E, 0, r)$.

2. Preliminary

2.1. **Nonlocal perimeter and first variation.** In this section we recall a few basic properties of the nonlocal perimeter depending on our assumptions on G. The following proposition is a consequence of [10] and our choice of I^1_G. It ensures that P_ε is well-defined on sets of finite perimeter and is bounded from above by the standard perimeter. We also state it for a general K, not necessarily normalized, since we will often use it with other kernels than G.

Proposition 2.1 (Upper bound). *Assume that $K : \mathbb{R}^n \to [0, \infty)$ satisfies (H1) and $x \mapsto |x|K(x) \in L^1(\mathbb{R}^n)$. Then, for every set of finite perimeter E in $\mathbb{R}^n$, we have*

$$P_K(E) \le \mathbb{K}_{1,n} I^1_K P(E). \tag{2.1}$$

In particular, for the kernels G_ε, we have

$$P_\varepsilon(E) \le P(E), \qquad \forall \varepsilon > 0. \tag{2.2}$$

Let us recall that P_ε is continuous with respect to the L^1 topology along sequences with bounded perimeter.

Lemma 2.2 (Continuity). *Assume that G satisfies (H1) and (H2). Let E_k be a sequence of sets of finite perimeter in $\mathbb{R}^n$ and $E \subseteq \mathbb{R}^n$ such that*

$$\sup_k (P(E_k) + |E_k|) < \infty \qquad \text{and} \qquad E_k \xrightarrow{L^1} E.$$

Then, for every $\varepsilon > 0$, we have

$$\lim_k P_\varepsilon(E_k) = P_\varepsilon(E).$$

Proof. Let $C := \sup_k \big(P(E_k) + |E_k|\big) < \infty$. Setting

$$u_k(z) := \int_{\mathbb{R}^n} |\mathbf{1}_{E_k}(x+z) - \mathbf{1}_{E_k}(x)|\,\mathrm{d}x \qquad \text{and} \qquad u(z) := \int_{\mathbb{R}^n} |\mathbf{1}_{E}(x+z) - \mathbf{1}_{E}(x)|\,\mathrm{d}x,$$

by the L^1 convergence of E_k to E, for every $z \in \mathbb{R}^n$, $u_k(z)$ converges to $u(z)$. In addition, we have

$$P_\varepsilon(E_k) = \int_{\mathbb{R}^n} u_k(z) G_\varepsilon(z)\,\mathrm{d}z$$

and

$$u_k(z)G_\varepsilon(z) \le P(E_k)|z|G_\varepsilon(z) \le C|z|G_\varepsilon(z) \in L^1(\mathbb{R}^n).$$

Hence by dominated convergence, $\lim_k P_\varepsilon(E_k) = P_\varepsilon(E)$. ∎

Depending on the integrability assumptions on G, we may estimate the difference $P_\varepsilon(E)-P_\varepsilon(F)$ from above by a perimeter term, a volume term, or an interpolation of the two. This type of estimates is relatively standard in the context of nonlocal perimeters (see for instance [13, Lemma 5.3] for a similar statement in the case of s-perimeters). The last interpolation estimate will allow us to show a useful quasi-minimality property at small scales for (Λ, r_0)-minimizers of $\mathcal{F}_{\varepsilon,\gamma}$ (see Proposition 2.10). We will not use (2.4) below but include it for completeness.

Lemma 2.3. *Let $E, F \subseteq \mathbb{R}^n$ be two measurable sets, and let $\varepsilon > 0$. We have:*

(i) if G satisfies (H1) and (H2), then

$$P_\varepsilon(E) - P_\varepsilon(F) \le P(E\triangle F); \tag{2.3}$$

(ii) if G satisfies (H1) and $G \in L^1(\mathbb{R}^n)$, then

$$P_\varepsilon(E) - P_\varepsilon(F) \le \frac{2I_G^0}{\varepsilon}|E\triangle F|; \tag{2.4}$$

(iii) if G satisfies (H1), (H2) and (H4), then there exists $C = C(n,G) > 0$ such that

$$P_\varepsilon(E) - P_\varepsilon(F) \le C\left(\frac{|E\triangle F|}{\varepsilon}\right)^{1-s_0} P(E\triangle F)^{s_0}. \tag{2.5}$$

Proof. We decompose the proof in two steps.

Step 1. We establish $P_\varepsilon(E) - P_\varepsilon(F) \le P_\varepsilon(E\triangle F)$. To this aim we note for $A, B \subseteq \mathbb{R}^n$,

$$\Phi_\varepsilon(A,B) := \int_{A\times B} G_\varepsilon(x-y)\,\mathrm{d}x\,\mathrm{d}y$$

so that $P_\varepsilon(E) = 2\Phi_\varepsilon(E,E^c)$. It is readily checked that

$$\begin{aligned}\Phi_\varepsilon(E,E^c) - \Phi_\varepsilon(F,F^c) &= \Phi_\varepsilon(E\cap F, F\backslash E) + \Phi_\varepsilon(E\backslash F, E^c\cap F^c) - \Phi_\varepsilon(E\cap F, E\backslash F)\\ &\qquad - \Phi_\varepsilon(F\backslash E, F^c\cap E^c)\\ &= \Phi_\varepsilon(E\triangle F, (E\triangle F)^c) - 2\left[\Phi_\varepsilon(E\cap F, E\backslash F) + \Phi_\varepsilon(F\backslash E, F^c\cap E^c)\right]\\ &\le \Phi_\varepsilon(E\triangle F, (E\triangle F)^c).\end{aligned}$$

This concludes the first step.

Step 2. We deduce the different cases. Case (i) is direct consequence of Step 1 and (2.2). If $G \in L^1(\mathbb{R}^n)$ then

$$P_\varepsilon(E) \le 2\|G_\varepsilon\|_{L^1(\mathbb{R}^n)}|E|$$

which gives (ii). For (iii), let us write, for any $R > 0$ and any $E \subseteq \mathbb{R}^n$,

$$P_\varepsilon(E) = \int_{\mathbb{R}^n\backslash B_R} G_\varepsilon(z)\int_{\mathbb{R}^n} |\chi_E(x+z) - \chi_E(x)|\,\mathrm{d}x\,\mathrm{d}z + \int_{B_R} G_\varepsilon(z)\int_{\mathbb{R}^n} |\chi_E(x+z)-\chi_E(x)|\,\mathrm{d}x\,\mathrm{d}z.$$

Using

$$\int_{\mathbb{R}^n} |\chi_E(x+z) - \chi_E(x)|\,\mathrm{d}x \le 2|E|$$

and

$$\int_{\mathbb{R}^n} |\chi_E(x+z) - \chi_E(x)|\,\mathrm{d}x \le |z|P(E),$$

we deduce

$$\begin{aligned}P_\varepsilon(E) &\le 2|E|\int_{\mathbb{R}^n\backslash B_R} G_\varepsilon(z)\,\mathrm{d}z + P(E)\int_{B_R} |z|G_\varepsilon(z)\,\mathrm{d}z\\ &= \frac{2|E|}{\varepsilon}\int_{\mathbb{R}^n\backslash B_{R/\varepsilon}} G(z)\,\mathrm{d}z + P(E)\int_{B_{R/\varepsilon}} |z|G(z)\,\mathrm{d}z.\end{aligned} \tag{2.6}$$

Next, we claim that (H4) implies

$$\int_{\mathbb{R}^n\backslash B_r} G(x)\,\mathrm{d}x \le \frac{C}{r^{s_0}}, \qquad \forall r > 0 \tag{2.7}$$

and

$$\int_{B_r} |x|G(x)\,\mathrm{d}x \le Cr^{1-s_0}, \qquad \forall r > 0, \tag{2.8}$$

for some $C = C(n, G) > 0$. It is of course enough to check these statements for either small or large r. We start with (2.7). Thanks to (H4), it holds for small r. If instead $r \geq 1$,

$$\int_{\mathbb{R}^n\setminus B_r} G(x)\,\mathrm{d}x \leq \frac{1}{r}\int_{\mathbb{R}^n\setminus B_r} |x|G(x)\,\mathrm{d}x \leq \frac{I_G^1}{r} \leq \frac{C}{r^{s_0}}.$$

We now turn to (2.8). By (H2) it is enough to prove it for $r \in (0,1)$. In this case, we have

$$\begin{aligned}\int_{B_r} |x|G(x)\,\mathrm{d}x &= \sum_{k=0}^{\infty}\int_{B_{2^{-k}r}\setminus B_{2^{-(k+1)}r}} |x|G(x)\,\mathrm{d}x \leq \sum_{k=0}^{\infty}\frac{r}{2^k}\int_{B_1\setminus B_{2^{-(k+1)}r}} G(x)\,\mathrm{d}x\\ &\overset{(2.7)}{\leq} C\sum_{k=0}^{\infty}\frac{r}{2^k}\left(\frac{2^k}{r}\right)^{s_0} = Cr^{1-s_0}\sum_{k=0}^{\infty}\frac{1}{2^{k(1-s_0)}} \leq Cr^{1-s_0},\end{aligned}$$

proving (2.8).
Plugging (2.7) and (2.8) into (2.6) yields

$$P_\varepsilon(E) \leq C\left(\frac{|E|}{\varepsilon}\left(\frac{\varepsilon}{R}\right)^{s_0} + P(E)\left(\frac{R}{\varepsilon}\right)^{1-s_0}\right).$$

Finally choosing $R = \frac{|E|}{P(E)}$, we get

$$P_\varepsilon(E) \leq C\left(\frac{|E|}{\varepsilon}\right)^{1-s_0} P(E)^{s_0}.$$

This concludes the proof of (iii). □

We will use the following computation from [27, Lemma 2.3] to estimate the derivative of the nonlocal perimeter under rescaling.

Lemma 2.4. *Assume that G satisfies (H1), (H2) and (H3). Then, for any set of finite perimeter $E \subseteq \mathbb{R}^n$, the function $t \mapsto P_\varepsilon(tE)$ is locally Lipschitz continuous in $(0, +\infty)$, and for almost every t, we have*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[P_\varepsilon(tE)\right] = \frac{n}{t}P_\varepsilon(tE) - \frac{1}{t}\widetilde{P}_\varepsilon(tE),$$

where $\widetilde{P}_\varepsilon(E)$ is defined by

$$\widetilde{P}_\varepsilon(E) := 2\int_E\int_{\partial^* E} G_\varepsilon(x-y)\,(y-x)\cdot\nu_E(y)\,\mathrm{d}\mathcal{H}^{n-1}_y\,\mathrm{d}x. \tag{2.9}$$

We now compute the first variation of the energy.

Lemma 2.5. *Assume that G satisfies (H1), (H2) and (H3). Let $T \in C^1_c(\mathbb{R}^n;\mathbb{R}^n)$ a compactly supported vector field, and let us define $f_t := \mathrm{Id}_{\mathbb{R}^n} + tT$. Then for any set of finite perimeter $E \subseteq \mathbb{R}^n$, $\varepsilon > 0$, $\gamma \in (0,1)$ and $\Lambda \geq 0$, the function $t \mapsto \mathcal{F}_{\varepsilon,\gamma}(f_t(E))$ is differentiable at $t = 0$ with $\delta\mathcal{F}_{\varepsilon,\gamma}(E)[T] := \left[\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}_{\varepsilon,\gamma}(f_t(E))\right]_{|t=0}$ given by*

$$\begin{aligned}\delta\mathcal{F}_{\varepsilon,\gamma}(E)[T] &= \int_{\partial^* E}\mathrm{div}_E\, T\,\mathrm{d}\mathcal{H}^{n-1}\\ &\quad - 2\gamma\left(\int_{E\times E^c}\mathrm{div}\, T(x)G_\varepsilon(x-y)\,\mathrm{d}x\,\mathrm{d}y + \int_{\partial^* E}\int_E G_\varepsilon(x-y)\,(T(x)-T(y))\cdot\nu_E(y)\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y\right)\end{aligned}$$

where $\mathrm{div}_E\, T$ is the boundary divergence of T on E, defined by

$$\mathrm{div}_E\, T(x) := \mathrm{div}\, T(x) - \nu_E(x)\cdot\nabla T(x)\nu_E(x), \quad \forall x \in \partial^* E.$$

Proof. Since the computation of the first variation of the perimeter is standard, see e.g. [26, Theorem 17.5], it is enough to compute the first variation of P_ε. We will show that (recall the notation $T_x = T(x)$)

$$\begin{aligned}\left[\frac{\mathrm{d}}{\mathrm{d}t}P_\varepsilon(f_t(E))\right]_{|t=0} &= 2\int_{E\times E^c}\mathrm{div}\, T(x)G_\varepsilon(x-y)\,\mathrm{d}x\,\mathrm{d}y\\ &\quad + 2\int_{\partial^* E}\int_E G_\varepsilon(x-y)\,(T_x - T_y)\cdot\nu_E(y)\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y.\end{aligned}$$

Notice that using (2.2) and the fact that T is Lipschitz continuous, (H1) and (H2) imply that both terms on the right-hand side are well-defined. Since ε does not play any role we may assume without loss of generality that $\varepsilon = 1$. We set $F_G(t) := \frac{1}{2}P_1(f_t(E))$. Note that choosing $t_0 \le 1/\|\nabla T\|_{L^\infty}$, f_t is a diffeomorphism of $\mathbb{R}^n$ for every t such that $|t| \le t_0$. In particular $f_t(E)$ is a set of finite perimeter (see e.g. [26, Proposition 17.1]). Thus, $F_G(t)$ is well-defined for every $t \in (-t_0, t_0)$. We then set (for the moment this is just a notation)

$$F'_G(0) := \int_{E\times E^c} \operatorname{div} T(x) G(x-y)\,\mathrm{d}x\,\mathrm{d}y + \int_{\partial^* E}\int_E G(x-y)\,(T_x - T_y)\cdot \nu_E(y)\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y.$$

We claim that as $t \to 0$,

$$F_G(t) - F_G(0) - tF'_G(0) = o(t). \tag{2.10}$$

This would show that F_G is differentiable in 0 with derivative $F'_G(0)$, concluding the proof. Changing variables, for any t small enough we have

$$F_G(t) = \int_{E\times E^c} G(f_t(x) - f_t(y)) \det Df_t(x)\, \det Df_t(y)\,\mathrm{d}x\,\mathrm{d}y.$$

Since $\det Df_t(x) = 1 + t\operatorname{div} T(x) + O(t^2)$, we find

$$F_G(t) = \int_{E\times E^c} G(f_t(x) - f_t(y))(1 + t\operatorname{div} T(x) + t\operatorname{div} T(y) + O(t^2))\,\mathrm{d}x\,\mathrm{d}y.$$

Notice that by the reverse change of variables and (2.2),

$$\int_{E\times E^c} G(f_t(x) - f_t(y))\,\mathrm{d}x\,\mathrm{d}y \le C\int_{f_t(E)\times f_t(E)^c} G(x-y)\,\mathrm{d}x\,\mathrm{d}y \le CP(f_t(E)) \le CP(E).$$

Therefore

$$F_G(t) = \int_{E\times E^c} G(f_t(x) - f_t(y))(1 + t\operatorname{div} T(x) + t\operatorname{div} T(y))\,\mathrm{d}x\,\mathrm{d}y + O(t^2).$$

Now, using that

$$G(f_t(x) - f_t(y)) - G(x-y) = t\int_0^1 \nabla G(f_{st}(x) - f_{st}(y))\cdot (T_x - T_y)\,\mathrm{d}s$$

and the Lipschitz continuity of T, we have

$$\begin{aligned}
&\left|\int_{E\times E^c} G(f_t(x) - f_t(y))\operatorname{div} T(x)\,\mathrm{d}x\,\mathrm{d}y - \int_{E\times E^c} G(x-y)\operatorname{div} T(x)\,\mathrm{d}x\,\mathrm{d}y\right| \\
&\quad\le C|t|\int_0^1\int_{E\times E^c} |\nabla G(f_{st}(x) - f_{st}(y))||x-y|\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s \\
&\quad\le C|t|\int_0^1\int_{E\times E^c} |\nabla G(f_{st}(x) - f_{st}(y))||f_{st}(x) - f_{st}(y)|\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s \\
&\quad\le C|t|\int_0^1\int_{f_{st}(E)\times(f_{st}(E))^c} |\nabla G(x-y)||x-y|\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s \\
&\quad\le CI^2_{|\nabla G|}|t|\int_0^1 P(f_{st}(E))\,\mathrm{d}s \le CI^2_{|\nabla G|}|t|,
\end{aligned}$$

where we used again (2.1) but for the kernel $K = |\cdot||\nabla G|$. Since the same holds with $\operatorname{div} T(x)$ replaced by $\operatorname{div} T(y)$, in order to prove (2.10) it is thus enough to show

$$\begin{aligned}
&\int_{E\times E^c} G(f_t(x) - f_t(y))\,\mathrm{d}x\,\mathrm{d}y - \int_{E\times E^c} G(x-y)\,\mathrm{d}x\,\mathrm{d}y \\
&\qquad + t\left(\int_{E\times E^c} G(x-y)\operatorname{div} T(y)\,\mathrm{d}x\,\mathrm{d}y - \int_{\partial^* E}\int_E G(x-y)\,(T_x - T_y)\cdot\nu_E(y)\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y\right) = o(t).
\end{aligned}$$

Writing as above that

$$\int_{E\times E^c} G(f_t(x)-f_t(y))\,\mathrm{d}x\,\mathrm{d}y - \int_{E\times E^c} G(x-y)\,\mathrm{d}x\,\mathrm{d}y = t\int_0^1\int_{E\times E^c}\nabla G(f_{st}(x)-f_{st}(y))\cdot(T_x-T_y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s$$

we reduce it further to the proof of

$$\lim_{t\to 0}\int_0^1\int_{E\times E^c}\nabla G(f_{st}(x)-f_{st}(y))\cdot(T_x-T_y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s = \int_{E\times E^c}\nabla G(x-y)\cdot(T_x-T_y)\,\mathrm{d}x\,\mathrm{d}y \tag{2.11}$$

together with the integration by parts formula

$$\int_{E\times E^c}\nabla G(x-y)\cdot(T_x-T_y)+G(x-y)\operatorname{div}T(y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\partial^* E}\int_E G(x-y)\,(T_x-T_y)\cdot\nu_E(y)\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y. \tag{2.12}$$

Notice that this would be easy to prove if G was a smooth kernel with compact support. However, since our assumptions on G seem too weak to prove these directly we will argue by approximation. Let G_k be a sequence of smooth compactly supported radial kernels with

$$\lim_{k\to\infty}\int_{\mathbb{R}^n}|z||[G-G_k](z)|\,\mathrm{d}z = 0 \qquad \text{and} \qquad \lim_{k\to\infty}\int_{\mathbb{R}^n}|z|^2|\nabla[G-G_k](z)|\,\mathrm{d}z = 0.$$

Since we assumed that $I^1_G + I^2_{|\nabla G|} < \infty$ it is not difficult to construct such a sequence. We start with (2.11). For every fixed $s\in[0,1]$, we have

$$\begin{aligned}
&\left|\int_{E\times E^c}\nabla G(f_{st}(x)-f_{st}(y))\cdot(T_x-T_y)\,\mathrm{d}x\,\mathrm{d}y - \int_{E\times E^c}\nabla G_k(f_{st}(x)-f_{st}(y))\cdot(T_x-T_y)\,\mathrm{d}x\,\mathrm{d}y\right|\\
&\quad\le C\int_{E\times E^c}|[\nabla G-\nabla G_k](f_{st}(x)-f_{st}(y))||x-y|\,\mathrm{d}x\,\mathrm{d}y\\
&\quad\le C\int_{f_{st}(E)\times f_{st}(E)^c}|\nabla[G-G_k](x-y)||x-y|\,\mathrm{d}x\,\mathrm{d}y\\
&\quad\le C\left(\int_{\mathbb{R}^n}|z|^2|\nabla[G-G_k](z)|\,\mathrm{d}z\right)P(f_{st}(E))\\
&\quad\le C\int_{\mathbb{R}^n}|z|^2|\nabla[G-G_k](z)|\,\mathrm{d}z,
\end{aligned}$$

where we used (2.1) with $K=|\cdot||\nabla[G-G_k]|$ (which is radially symmetric). Integrating in s and using a simple diagonal argument, this proves (2.11). We now turn to (2.12). Since

$$-\operatorname{div}_y(G(x-y)(T_x-T_y)) = \nabla G(x-y)\cdot(T_x-T_y)+G(x-y)\operatorname{div}T(y),$$

the integration by parts formula (2.12) holds with G replaced by G_k. By the previous computations it is therefore enough to observe that on the one hand

$$\left|\int_{E\times E^c}G(x-y)\operatorname{div}T(y)\,\mathrm{d}x\,\mathrm{d}y - \int_{E\times E^c}G_k(x-y)\operatorname{div}T(y)\,\mathrm{d}x\,\mathrm{d}y\right| \le CP(E)\int_{\mathbb{R}^n}|z||[G-G_k](z)|\,\mathrm{d}z$$

and on the other hand,

$$\begin{aligned}
&\left|\int_{\partial^* E}\int_E G(x-y)\,(T_x-T_y)\cdot\nu_E(y)\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y - \int_{\partial^* E}\int_E G_k(x-y)\,(T_x-T_y)\cdot\nu_E(y)\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y\right|\\
&\qquad\le C\int_{\partial^* E}\int_E|x-y||[G-G_k](x-y)|\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y \le CP(E)\int_{\mathbb{R}^n}|z||[G-G_k](z)|\,\mathrm{d}z.
\end{aligned}$$

□

2.2. **Perimeter quasi-minimizing properties of minimizers.** We recall from [29, (4.2)] that using (2.2) it readily follows that if E satisfies $\mathcal{F}_{\varepsilon,\gamma}(E) \le \mathcal{F}_{\varepsilon,\gamma}(B_1)$, then

$$P(E) \le P(B_1) + \frac{\gamma}{1-\gamma}\left(P(B_1) - P_\varepsilon(B_1)\right) \le \frac{1}{1-\gamma} P(B_1). \tag{2.13}$$

We now use the scaling properties given in Lemma 2.4 to prove the equivalence between $(\mathcal{P})$ and the unconstrained minimization problem

$$\min \left\{ \mathcal{F}_{\varepsilon,\gamma}(E) + \Lambda\big||E| - |B_1|\big| \ : \ E \subseteq \mathbb{R}^n \text{ measurable} \right\} \tag{$\mathcal{P}'$}$$

if Λ is large enough, not depending on ε. As a consequence, minimizers of $(\mathcal{P})$ are (Λ, r_0)-minimizers of $\mathcal{F}_{\varepsilon,\gamma}$.

Proposition 2.6. *Assume that G satisfies (H1) and (H2) and let $\gamma \in (0,1)$. There exists $C = C(n) > 0$ such that for every $\gamma \in (0,1)$, $\varepsilon > 0$ and $\Lambda \ge C/(1-\gamma)$, problems $(\mathcal{P})$ and $(\mathcal{P}')$ are equivalent, in the sense that $(\mathcal{P}')$ admits a minimizer if and only if $(\mathcal{P})$ does, and their minimizers coincide. In particular, any minimizer of $(\mathcal{P})$ is a (Λ, r_0)-minimizer of $\mathcal{F}_{\varepsilon,\gamma}$ for any $\Lambda \ge C/(1-\gamma)$ and any $r_0 > 0$.*

Proof. Let us set

$$\Lambda_0 \coloneqq \frac{1}{1-\gamma}\left(1 + \left(n + \frac{2}{\mathbb{K}_{1,n}}\right)\right)\frac{P(B_1)}{|B_1|}.$$

Since

$$\inf_{|E|=|B_1|} \mathcal{F}_{\varepsilon,\gamma}(E) \ge \inf_E \mathcal{F}_{\varepsilon,\gamma,\Lambda}(E),$$

it is enough to prove that for $\Lambda \ge \Lambda_0$, the converse inequality holds and that any set minimizing $\mathcal{F}_{\varepsilon,\gamma,\Lambda}$ must have measure equal to ω_n. This in turn is equivalent to the claim that if E is such that

$$|E| \ne \omega_n \quad \text{ and } \quad \mathcal{F}_{\varepsilon,\gamma,\Lambda}(E) \le \inf_{|E|=|B_1|} \mathcal{F}_{\varepsilon,\gamma}(E)$$

then $\Lambda < \Lambda_0$. Let E be such a set. Recall that E satisfies (2.13). Let λ be such that $|\lambda E| = |B_1|$. We then have

$$P(E) - \gamma P_\varepsilon(E) + \Lambda||E| - |B_1|| = \mathcal{F}_{\varepsilon,\gamma,\Lambda}(E) \le \mathcal{F}_{\varepsilon,\gamma}(\lambda E) = \lambda^{n-1}P(E) - \gamma P_\varepsilon(\lambda E).$$

Reorganizing terms we find

$$\Lambda\omega_n|\lambda^n - 1| \le (\lambda^{n-1} - 1)P(E) + \gamma|P_\varepsilon(E) - P_\varepsilon(\lambda E)|. \tag{2.14}$$

By Lemma 2.4 and (2.2), for any $t > 0$ we have

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\left[P_\varepsilon(tE)\right]\right| \le \frac{1}{t}\left(nP_\varepsilon(tE) + |\widetilde{P}_{G_\varepsilon}(tE)|\right) \le \frac{1}{t}\left(nP(tE) + 2P(tE)I_G^1\right) \le t^{n-2}\left(n + \frac{2}{\mathbb{K}_{1,n}}\right),$$

thus

$$|P_\varepsilon(E) - P_\varepsilon(\lambda E)| \le \left|\int_\lambda^1 \frac{\mathrm{d}}{\mathrm{d}t}\left[P_\varepsilon(tE)\right]\mathrm{d}t\right| \le C_1|\lambda^{n-1} - 1|P(E),$$

where $C_1 \coloneqq \left(n + \frac{2}{\mathbb{K}_{1,n}}\right)$. Inserting this into (2.14) and using (2.13), this leads to

$$\Lambda\omega_n|\lambda^n - 1| \le \frac{1}{1-\gamma}\left(1 + C_1\gamma\right)P(B_1)|\lambda^{n-1} - 1|.$$

Since $|\lambda^{n-1} - 1| < |\lambda^n - 1|$ we conclude that $\Lambda < \Lambda_0$. □

We recall the following elementary scaling properties of the energy which we will heavily use in the paper.

Proposition 2.7. *For any set of finite perimeter E, any $\varepsilon > 0$ and any $r > 0$ we have*

$$\mathcal{F}_{\varepsilon,\gamma}(E) = r^{n-1}\mathcal{F}_{\varepsilon/r,\gamma}(E/r).$$

In particular E is a (Λ, r_0)-minimizer of $\mathcal{F}_{\varepsilon,\gamma}$ if and only if E/r is a $(\Lambda r, \frac{r_0}{r})$-minimizer of $\mathcal{F}_{\varepsilon/r,\gamma}$.

We now prove that (Λ, r_0)-minimizers of $\mathcal{F}_{\varepsilon,\gamma}$ are quasi-minimizers of the perimeter and thus have density bounds which are uniform in ε.

Proposition 2.8 (Weak quasi-minimality)**.** *Assume that G satisfies (H1) and (H2) and let $\gamma \in (0,1)$, $\varepsilon > 0$, $\Lambda > 0$ and $r_0 > 0$ with $\Lambda r_0 \le 1-\gamma$. Then, for any (Λ, r_0)-minimizer E of $\mathcal{F}_{\varepsilon,\gamma}$ and every set F with $E \triangle F \subset\subset B_r(x)$ with $0 < r \le r_0$ we have*

$$P(E; B_r(x)) \le \frac{4}{1-\gamma} P(F; B_r(x)). \tag{2.15}$$

As a consequence, there exists $C = C(n) > 0$ such that for every $x \in \partial E$ and every $0 < r \le r_0$,

$$\left(\frac{1-\gamma}{4}\right)^n \le \frac{|E \cap B_r(x)|}{r^n} \le 1 - \left(\frac{1-\gamma}{4}\right)^n \quad \text{and} \quad \frac{(1-\gamma)^{n-1}}{C} \le \frac{P(E; B_r(x))}{r^{n-1}} \le \frac{C}{1-\gamma}. \tag{2.16}$$

In particular, we have

$$\mathcal{H}^{n-1}(\partial E \setminus \partial^* E) = 0. \tag{2.17}$$

Proof. We only prove (2.15) since it is standard that weak quasi-minimality implies density upper and lower bounds (see [16, Theorem 5.6]), which then imply (2.17). To obtain the correct scaling in γ, one can repeat the proof in [26, Theorem 21.11]. By scaling and translation, we may assume that $r = 1$ and $x = 0$. Testing the (Λ, r_0)-minimality of E against F, we have

$$\begin{aligned} P(E; B_1) &\le P(F; B_1) + \gamma \left(P_\varepsilon(E) - P_\varepsilon(F)\right) + \Lambda |E \triangle F| \\ &\overset{(2.3)\&(2.2)}{\le} P(F; B_1) + \gamma P(E \triangle F) + \Lambda |E \triangle F|. \end{aligned}$$

We now argue as in [26, Remark 21.7] and use the isoperimetric inequality to infer

$$|E \triangle F| = |E \triangle F|^{\frac{1}{n}} |E \triangle F|^{1-\frac{1}{n}} \le \frac{1}{n} P(E \triangle F).$$

We thus find

$$P(E; B_1) \le P(F; B_1) + \left(\gamma + \frac{\Lambda}{n}\right) P(E \triangle F) \le P(F; B_1) + \left(\gamma + \frac{\Lambda}{n}\right) (P(E; B_1) + P(F; B_1)).$$

Rearranging terms and using that $\Lambda/n \le (1-\gamma)/2$ yields (2.15). □

Remark 2.9. Thanks to (2.17) if E is a (Λ, r_0)-minimizer of $\mathcal{F}_{\varepsilon,\gamma}$ with $\Lambda r_0 \le 1-\gamma$, we will not distinguish anymore between ∂E and $\partial^* E$ when integrating.

Under hypothesis (H4) we prove that (Λ, r_0)-minimizers of $\mathcal{F}_{\varepsilon,\gamma}$ are also almost-minimizers at scales which are small compared to ε.

Proposition 2.10. *Assume that G satisfies (H1), (H2) and (H4). Then there exists $C = C(n, G, \gamma) > 0$ such that for every $\gamma \in (0,1)$, $\varepsilon > 0$, $\Lambda > 0$ and $r_0 > 0$ with $\Lambda r_0 \le 1-\gamma$, every (Λ, r_0)-minimizer E of $\mathcal{F}_{\varepsilon,\gamma}$ and every set F with $E \triangle F \subset\subset B_r(x)$ and $r \le r_0$, we have*

$$P(E; B_r(x)) \le P(F; B_r(x)) + \left(\frac{C}{\varepsilon^{1-s_0}}\right) r^{n-s_0} + \Lambda |E \triangle F|. \tag{2.18}$$

Proof. We may assume that $P(F; B_r(x)) \le P(E; B_r(x))$ otherwise there is nothing to prove. Arguing as above using the (Λ, r_0)-minimality of E we have

$$\begin{aligned} P(E; B_r(x)) &\le P(F; B_r(x)) + \gamma\big(P_\varepsilon(E) - P_\varepsilon(F)\big) + \Lambda |E \triangle F| \\ &\overset{(2.5)}{\le} P(F; B_r(x)) + C\gamma \left(\frac{|E \triangle F|}{\varepsilon}\right)^{1-s_0} P(E \triangle F)^{s_0} + \Lambda |E \triangle F| \\ &\overset{(2.16)}{\le} P(F; B_r(x)) + \left(\frac{C}{\varepsilon^{1-s_0}}\right) r^{n-s_0} + \Lambda |E \triangle F|. \end{aligned}$$

□

Remark 2.11. Proposition 2.10 indeed yields classical almost-minimality for the perimeter at scales smaller than ε since letting $r = \varepsilon \hat{r}$ and $E = x + \varepsilon \hat{E}$, we find for every $\hat{F} \triangle \hat{E} \subset\subset B_{\hat{r}}$,

$$P(\hat{E}; B_{\hat{r}}) \le P(\hat{F}; B_{\hat{r}}) + C \hat{r}^{n-s_0} + \Lambda \varepsilon |\hat{E} \triangle \hat{F}|.$$

2.3. **Basic properties of the excess.** We recall two basic properties of the excess that we use extensively in the rest of the paper. The cylindrical excess and spherical excess are respectively defined in Definition 1.3 and Definition 1.4. We refer to [26, Chapter 22.1] for more details on the excess.

Proposition 2.12 (Scaling properties). *Let $E \subseteq \mathbb{R}^n$ be a set of finite perimeter, $x \in \partial E$, $\nu \in \mathbb{S}^{n-1}$ and $0 < r < R$. Then we have*

$$\mathbf{e}(E, x, r, \nu) \leq \left(\frac{R}{r}\right)^{n-1} \mathbf{e}(E, x, R, \nu) \qquad \textit{and}$$

In addition, setting $E_{x,r} := \frac{(E-x)}{r}$ we have

$$\mathbf{e}(E_{x,r}, 0, 1, \nu) = \mathbf{e}(E, x, R, \nu).$$

Note that this property holds for the spherical excess as well.

Proposition 2.13 (Changes of direction). *Let $\gamma \in (0,1)$ and $\varepsilon > 0$. There exists $C = C(n,\gamma) > 0$ such that for every (Λ, r_0)-minimizer E of $\mathcal{F}_{\varepsilon,\gamma}$ with $\Lambda r_0 \leq 1-\gamma$, every $\nu, \nu_0 \in \mathbb{S}^{n-1}$, $x \in \partial E$ and $r > 0$ such that $\sqrt{2}r \leq r_0$, we have*

$$\mathbf{e}(E, x, r, \nu) \leq C\left(\mathbf{e}(E, x, \sqrt{2}r, \nu_0) + |\nu - \nu_0|^2\right).$$

The proof is identical to the one in [26, Proposition 22.5] and relies only on the density estimates for minimizers. Since it is very short, we write it for the reader's convenience.

Proof. Using the inequality $|\nu - \nu_E(y)|^2 \leq 2|\nu - \nu_0|^2 + 2|\nu_0 - \nu_E(y)|^2$, and the facts that $\mathbf{C}(x, r, \nu) \subseteq \mathbf{C}(x, \sqrt{2}r, \nu_0)$ (recall the definition (1.4)) and $\mathbf{C}(x, r, \nu) \subseteq B_{\sqrt{2}r}(x)$, we have

$$\mathbf{e}(E, x, r, \nu) \leq \frac{2}{r^{n-1}} \int_{\partial E \cap C(x, \sqrt{2}r, \nu_0)} \frac{|\nu_0 - \nu_E(y)|^2}{2} \, \mathrm{d}\mathcal{H}^{n-1}_y + \frac{P(E; B_{\sqrt{2}r}(x))}{r^{n-1}} |\nu - \nu_0|^2.$$

The results follows from (2.16). □

2.4. **The height bound.** Thanks to the density estimates of Proposition 2.8, (Λ, r_0)-minimizers of $\mathcal{F}_{\varepsilon,\gamma}$ satisfy the same "height bound" property as quasi-minimizers of the perimeter (see [26, Theorem 22.8]). This property is a crucial tool for the Lipschitz approximation theorem and the Caccioppoli inequality.

Proposition 2.14 (The height bound). *Let $\varepsilon > 0$, $\gamma \in (0,1)$, $\Lambda > 0$ and $r_0 > 0$ with $\Lambda r_0 \leq 1 - \gamma$. There exist positive constants $\tau_{\text{height}} = \tau_{\text{height}}(n, \gamma)$ and $C = C(n, \gamma)$ such that the following holds. For every (Λ, r_0)-minimizer E of $\mathcal{F}_{\varepsilon,\gamma}$, every $x \in \partial E$, $\nu \in \mathbb{S}^{n-1}$ and $r > 0$ with $2r \leq r_0$, if*

$$\mathbf{e}_n(x, 2r) < \tau_{\text{height}},$$

then

$$\sup\left\{|x_n - y_n| \;:\; (y', y_n) \in \partial E \cap \mathbf{C}_r(x)\right\} \leq C r \mathbf{e}_n(x, 2r)^{\frac{1}{2(n-1)}}.$$

Proof. As recalled by F. Maggi, the only step where the almost-minimality with respect to the perimeter is used in the proof of [26, Theorem 22.8] is to obtain the "small-excess position" of Lemma 22.10 therein. In fact, this lemma holds as long as we have density estimates on the perimeter for E, as shown in [12, Lemma 7.2]. Hence, thanks to (2.16), the same height bound holds for (Λ, r_0)-minimizers of $\mathcal{F}_{\varepsilon,\gamma}$, whenever $\Lambda r_0 \leq 1 - \gamma$ and $2r \leq r_0$. □

3. Lipschitz approximation theorem

This section is devoted to the proof of the Lipschitz approximation theorem for (Λ, r_0)-minimizers of $\mathcal{F}_{\varepsilon,\gamma}$, which can be divided into two parts. A first part states that a small excess of such an almost-minimizer E in a cylinder implies that the boundary of E in that cylinder is almost entirely covered by the graph of a Lipschitz function u. A second step states that the aforementioned function u is close to a harmonic function as long as the scale is much larger than ε.

3.1. **Lipschitz approximation and harmonic comparison.** Since the first part of the Lipschitz approximation theorem relies only on standard properties on the excess, the density estimates and the height bound, by Propositions 2.8 and 2.14, the proof can be reproduced almost verbatim from Steps 1 to 4 of the proof of [26, Theorem 23.7].

Theorem 3.1 (Lipschitz approximation I). *Assume that G satisfies (H1), (H2) and (H3). Let $\varepsilon > 0$, $\gamma \in (0,1)$, $\Lambda > 0$ and $r_0 > 0$ with $\Lambda r_0 \le 1-\gamma$. There exist positive constants $\tau_{\rm lip} = \tau_{\rm lip}(n,\gamma)$, $\delta_0 = \delta_0(n,\gamma)$ and $C = C(n,\gamma)$ such that the following holds. If E is a (Λ, r_0)-minimizer of $\mathcal{F}_{\varepsilon,\gamma}$ with $0 \in \partial E$ and, for some r such that $4\Lambda r \le r_0$,*

$$\mathbf{e}_n(4r) \le \tau_{\rm lip},$$

then, setting

$$M := \partial E \cap \mathbf{C}_{2r},$$

there exists a $\frac{1}{2}$-Lipschitz function $u : \mathbb{R}^{n-1} \to \mathbb{R}$ such that:

(i) $\|u\|_{L^\infty} \le C r \mathbf{e}_n(4r)^{\frac{1}{2(n-1)}} < \frac{r}{4}$;

(ii) $\mathcal{H}^{n-1}(M \triangle \Gamma_u) \le C\mathbf{e}_n(4r) r^{n-1}$;

(iii) $\displaystyle\frac{1}{r^{n-1}} \int_{D_{2r}} |\nabla u|^2 \le C\mathbf{e}_n(4r)$.

We show that the function u in the conclusion of Theorem 3.1 is "almost" a solution to a nonlocal linear equation of the form $(\Delta - \gamma\Delta_{G_\varepsilon})u = 0$ in D_r.

Theorem 3.2 (Lipschitz approximation II). *There exists $C = C(n,\gamma, I^2_{\nabla G}) > 0$ such that under the same assumptions as Theorem 3.1, the function u satisfies for every $\varphi \in C^1_c(D_r)$.*

$$\frac{1}{r^{n-1}}\left(\int_{D_r} \nabla u\cdot\nabla\varphi - \gamma\int_{D_{2r}\times D_{2r}} (u(x')-u(y'))(\varphi(x')-\varphi(y'))G_\varepsilon(x'-y',0)\,\mathrm{d}x'\,\mathrm{d}y'\right)$$
$$\le C\|\nabla\varphi\|_{L^\infty}\left(\mathbf{e}_n(4r) + Q\left(\frac{r}{4\varepsilon}\right) + \Lambda r\right).$$

By scaling it is enough to prove Theorem 3.2 for $r = 1$. Since the proof is quite long, we postpone it to the next section and show first how it leads to a harmonic approximation result.

Proposition 3.3 (Harmonic approximation). *Let $\gamma \in (0,1)$ and assume that G satisfies (H1) and (H2). There exists $\varepsilon_{\rm harm} \in (0,1)$ such that for every $\tau > 0$, there exists $\sigma = \sigma(n, G, \gamma, \tau) > 0$ with the following property. If for some $\varepsilon \in (0, \varepsilon_{\rm harm})$, $u \in H^1(D_2)$ satisfies*

$$\int_{D_2} |\nabla u|^2 \le 1$$

and, for all $\varphi \in C^1_c(D_1)$,

$$\left|\int_{D_1} \nabla u\cdot\nabla\varphi - 2\gamma\int_{D_2\times D_2}(u(x')-u(y'))(\varphi(x')-\varphi(y'))G_\varepsilon(x'-y',0)\,\mathrm{d}x'\,\mathrm{d}y'\right| \le \|\nabla\varphi\|_{L^\infty}\sigma,$$

then there exists a harmonic function v on D_1 such that

$$\int_{D_1} |\nabla v|^2 \le 1 \quad \text{and} \quad \int_{D_1} |u-v|^2 \le \tau.$$

Proof. As there is no risk for confusion, to simplify the notation we use x, y instead of x', y' for points in $\mathbb{R}^{n-1}$, and write $G_\varepsilon(x)$ instead of $G_\varepsilon(x', 0)$. Arguing by contradiction, let us assume that there exist vanishing sequences $(\varepsilon_k) \subseteq (0,1)$ and $(\sigma_k) \subseteq (0,1)$, a positive constant $\tau > 0$ and a sequence $(u_k) \subseteq H^1(D_2)$ such that the following holds:

(i) $\displaystyle\int_{D_2} |\nabla u_k|^2 \le 1$ for all $k \in \mathbb{N}$;

(ii) for every $\varphi \in C^\infty_c(D_1)$ we have

$$\left|\int_{D_1} \nabla u_k\cdot\nabla\varphi - 2\gamma\int_{D_2\times D_2}(u_k(x)-u_k(y))(\varphi(x)-\varphi(y))G_{\varepsilon_k}(x-y)\,\mathrm{d}x\,\mathrm{d}y\right| \le \sigma_k\|\nabla\varphi\|_{L^\infty};$$

(iii) there is no harmonic function u on D_1 such that

$$\int_{D_1} |\nabla u|^2 \le 1 \quad \text{and} \quad \int_{D_1} |u_k - u|^2 \le \tau.$$

Without loss of generality, up to adding a constant to each u_k, one may assume that $\int_{D_2} u_k = 0$, so that by Poincaré–Wirtinger inequality, we have

$$\int_{D_2} |u_k|^2 \le C \int_{D_2} |\nabla u_k|^2 \le C, \qquad \forall k \in \mathbb{N}. \tag{3.1}$$

In particular, (u_k) is bounded in $H^1(D_2)$. Thus, up to extraction of a subsequence (not relabeled), there exists $u \in H^1(D_2)$ such that u_k converges strongly to u in $L^2(D_2)$ and ∇u_k converges weakly to ∇u in $L^2(D_2; \mathbb{R}^{n-1})$. We claim that for every $\varphi \in C_c^\infty(D_1)$,

$$\lim_k \int_{D_2\times D_2} (u_k(x) - u_k(y))(\varphi(x) - \varphi(y)) G_{\varepsilon_k}(x-y)\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{2}\int_{D_1} \nabla u \cdot \nabla\varphi, \tag{3.2}$$

which we prove further below. By the weak convergence of ∇u_k to ∇u, the fact that $\gamma \neq 1$ and (ii), this implies

$$\int_{D_1} \nabla u \cdot \nabla \varphi = 0, \qquad \forall \varphi \in C_c^\infty(D_1);$$

in other words, u is harmonic. By (i) and lower semicontinuity with respect to the weak H^1 convergence, we have

$$\int_{D_1} |\nabla u|^2 \le 1, \tag{3.3}$$

and since u_k converges to u in $L^2(D_1)$, for any k large enough, we have

$$\int_{D_1} |u_k - u|^2 \le \tau.$$

With (3.3), this contradicts (iii).
We now prove (3.2). Using the change of variable $z = x - y$, we have

$$\begin{aligned}
&\int_{D_2\times D_2} (u_k(x) - u_k(y))(\varphi(x) - \varphi(y)) G_{\varepsilon_k}(x-y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_0^1\int_0^1 \int_{D_2\times D_2} \big(\nabla u_k(x + t(y-x)) \cdot (x-y)\big)\big(\nabla\varphi(x + s(y-x))\cdot(x-y)\big) G_{\varepsilon_k}(x-y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\,\mathrm{d}t \\
&= \int_0^1\int_0^1 \int_{\mathbb{R}^{n-1}}\int_{D_2} \mathbf{1}_{D_2}(x+z)\,(\nabla u_k(x+tz)\cdot z)\,(\nabla\varphi(x+sz)\cdot z)\,G_{\varepsilon_k}(z)\,\mathrm{d}x\,\mathrm{d}z\,\mathrm{d}s\,\mathrm{d}t.
\end{aligned} \tag{3.4}$$

Let us set $g_\varepsilon(r) := \varepsilon^{-(n+1)} g(\varepsilon^{-1} r)$ for every $r > 0$ (recall $G(x) = g(|x|)$ for every $x \in \mathbb{R}^n \setminus \{0\}$). Then for each $s, t \in (0,1)$ and each $x \in D_2$, using polar coordinates, we have

$$\begin{aligned}
&\int_{\mathbb{R}^{n-1}} \mathbf{1}_{D_2}(x+z)\,(\nabla u_k(x+tz)\cdot z)\,(\nabla\varphi(x+sz)\cdot z)\,G_{\varepsilon_k}(z)\,\mathrm{d}z \\
&\quad = \int_0^4 r^n g_{\varepsilon_k}(r) \int_{\mathbb{S}^{n-2}} \mathbf{1}_{D_2}(x + r\sigma)\big(\nabla u_k(x + tr\sigma)\cdot\sigma\big)\big(\nabla\varphi(x+sr\sigma)\cdot\sigma\big)\,\mathrm{d}\mathcal{H}^{n-2}_\sigma\,\mathrm{d}r.
\end{aligned} \tag{3.5}$$

Using the fact that for every $s, t \in (0,1)$ we have $|\nabla\varphi(x + sr\sigma) - \nabla\varphi(x+tr\sigma)| \le r\|D^2\varphi\|_{L^\infty}$ and Cauchy–Schwarz inequality, we deduce that for every $s, t \in (0,1)$ and every $\sigma \in \mathbb{S}^{n-2}$,

$$\begin{aligned}
&\bigg|\int_0^4 r^n g_{\varepsilon_k}(r) \int_{D_2} \mathbf{1}_{D_2}(x+r\sigma)\big(\nabla u_k(x+tr\sigma)\cdot\sigma\big)\big(\nabla\varphi(x+sr\sigma)\cdot\sigma\big)\,\mathrm{d}x\,\mathrm{d}r \\
&\qquad - \int_0^4 r^n g_{\varepsilon_k}(r) \int_{D_2} \mathbf{1}_{D_2}(x+r\sigma)\big(\nabla u_k(x+tr\sigma)\cdot\sigma\big)\big(\nabla\varphi(x+tr\sigma)\cdot\sigma\big)\,\mathrm{d}x\,\mathrm{d}r\bigg| \\
&\qquad\qquad \le C\|D^2\varphi\|_{L^\infty}\left(\int_0^4 r^{n+1} g_{\varepsilon_k}(r)\,\mathrm{d}r\right)\left(\int_{D_2}|\nabla u|^2\right)^{\frac{1}{2}}.
\end{aligned}$$

Notice that

$$\lim_{k\to\infty} \int_0^4 r^{n+1} g_{\varepsilon_k}(r)\,\mathrm{d}r = 0$$

since $r \mapsto r^n g(r) \in L^1(\mathbb{R})$ and

$$\begin{aligned}\int_0^4 r^{n+1} g_{\varepsilon_k}(r)\,\mathrm{d}r &= \int_0^{\frac{4}{\varepsilon_k}} (\varepsilon_k r) r^n g(r)\,\mathrm{d}r \\ &= \int_0^{\frac{4}{\sqrt{\varepsilon_k}}} (\varepsilon_k r) r^n g(r)\,\mathrm{d}r + \int_{\frac{4}{\sqrt{\varepsilon_k}}}^{\frac{4}{\varepsilon_k}} (\varepsilon_k r) r^n g(r)\,\mathrm{d}r \\ &\leq 4\sqrt{\varepsilon_k} \int_0^\infty r^n g(r)\,\mathrm{d}r + 4 \int_{\frac{4}{\sqrt{\varepsilon_k}}}^\infty r^n g(r)\,\mathrm{d}r.\end{aligned}$$

Therefore, in view of (3.4) and (3.5), in order to prove (3.2), we only need to compute the limit of

$$\begin{aligned}&\int_0^1 \int_0^4 r^n g_{\varepsilon_k}(r) \int_{\mathbb{S}^{n-2}} \int_{D_2} \mathbf{1}_{D_2}(x + r\sigma)\big(\nabla u_k(x + tr\sigma)\cdot\sigma\big)\big(\nabla\varphi(x + tr\sigma)\cdot\sigma\big)\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-2}_\sigma\,\mathrm{d}r\,\mathrm{d}t \\ &= \int_0^1 \int_0^4 r^n g_{\varepsilon_k}(r) \int_{\mathbb{S}^{n-2}} \int_{\mathbb{R}^{n-1}} \mathbf{1}_{D_2}(y - tr\sigma)\mathbf{1}_{D_2}(y + (1-t)r\sigma)\big(\nabla u_k(y)\cdot\sigma\big)\big(\nabla\varphi(y)\cdot\sigma\big) \\ &\qquad\qquad \mathrm{d}y\,\mathrm{d}\mathcal{H}^{n-2}_\sigma\,\mathrm{d}r\,\mathrm{d}t \\ &= \int_0^1 \int_0^{\frac{4}{\varepsilon_k}} r^n g(r) \int_{\mathbb{S}^{n-2}} \int_{D_1} \mathbf{1}_{D_2}(y - t\varepsilon_k r\sigma)\mathbf{1}_{D_2}(y + (1-t)\varepsilon_k r\sigma)\big(\nabla u_k(y)\cdot\sigma\big)\big(\nabla\varphi(y)\cdot\sigma\big) \\ &\qquad\qquad \mathrm{d}y\,\mathrm{d}\mathcal{H}^{n-2}_\sigma\,\mathrm{d}r\,\mathrm{d}t,\end{aligned} \tag{3.6}$$

where we used a change of variables and the fact that $\varphi \in C^\infty_c(D_1)$. By the weak convergence of ∇u_k to ∇u, for any $r > 0$, $t \in (0,1)$ and $\sigma \in \mathbb{S}^{n-2}$, we have

$$\lim_k \int_{D_1} \big(\nabla u_k\cdot\sigma\big)\big(\nabla\varphi\cdot\sigma\big) = \int_{D_1} \big(\nabla u\cdot\sigma\big)\big(\nabla\varphi\cdot\sigma\big)$$

and

$$\begin{aligned}&\left|\int_{D_1} \mathbf{1}_{D_2}(y - t\varepsilon_k r\sigma)\mathbf{1}_{D_2}(y + (1-t)\varepsilon_k r\sigma)\big(\nabla u_k(y)\cdot\sigma\big)\big(\nabla\varphi(y)\cdot\sigma\big)\,\mathrm{d}y - \int_{D_1}\big(\nabla u_k\cdot\sigma\big)\big(\nabla\varphi\cdot\sigma\big)\right| \\ &\quad\leq \int_{D_1\setminus\big(D_2(t\varepsilon_k r)\cup D_2((1-t)\varepsilon_k r)\big)} |\nabla u_k||\nabla\varphi| \\ &\quad\leq \|\nabla u_k\|_{L^2(D_1)} \left(\int_{D_1\setminus\big(D_2(t\varepsilon_k r)\cup D_2((1-t)\varepsilon_k r)\big)} |\nabla\varphi|^2\right)^{\frac12} \xrightarrow{k\to\infty} 0,\end{aligned}$$

where we used the inequality $\|\nabla u_k\|_{L^2(D_1)} \leq 1$ to pass to the limit. Thus, for any $r > 0$, $t \in (0,1)$ and $\sigma \in \mathbb{S}^{n-2}$, we have

$$\lim_k \int_{D_1} \mathbf{1}_{D_2}(y - t\varepsilon_k r\sigma)\mathbf{1}_{D_2}(y + (1-t)\varepsilon_k r\sigma)\ \big(\nabla u_k(y)\cdot\sigma\big)\big(\nabla\varphi(y)\cdot\sigma\big)\,\mathrm{d}y = \int_{D_1}\big(\nabla u\cdot\sigma\big)\big(\nabla\varphi\cdot\sigma\big).$$

Hence, using once more $\|\nabla u_k\|_{L^2(D_1)} \leq 1$ and Cauchy–Schwarz inequality, applying the dominated convergence theorem yields

$$\begin{aligned}&\lim_k \int_0^1 \int_0^{\frac{4}{\varepsilon_k}} r^n g(r) \int_{\mathbb{S}^{n-2}} \int_{D_1} \mathbf{1}_{D_2}(y-t\varepsilon_k r\sigma)\mathbf{1}_{D_2}(y+(1-t)\varepsilon_k r\sigma)\big(\nabla u_k(y)\cdot\sigma\big)\big(\nabla\varphi(y)\cdot\sigma\big)\,\mathrm{d}y\,\mathrm{d}\mathcal{H}^{n-2}_\sigma\,\mathrm{d}r\,\mathrm{d}t \\ &\qquad = \int_0^\infty r^n g(r) \int_{\mathbb{S}^{n-2}} \int_{D_1} \big(\nabla u(y)\cdot\sigma\big)\big(\nabla\varphi(y)\cdot\sigma\big)\,\mathrm{d}y\,\mathrm{d}\mathcal{H}^{n-2}_\sigma\,\mathrm{d}r.\end{aligned} \tag{3.7}$$

This concludes the proof of (3.2) in view of the normalization (1.2) and the fact that for every $x, y \in \mathbb{R}^{n-1}$,

$$\begin{aligned}\int_{\mathbb{S}^{n-2}} (x\cdot\sigma)(y\cdot\sigma)\,\mathrm{d}\mathcal{H}^{n-2}_\sigma &= x\cdot\left(\int_{\mathbb{S}^{n-2}} \sigma\otimes\sigma\,\mathrm{d}\mathcal{H}^{n-2}_\sigma\right) y = \left(\int_{\mathbb{S}^{n-2}} |\sigma_1|^2\,\mathrm{d}\mathcal{H}^{n-2}_\sigma\right) x\cdot y \\ &= \frac12\left(\int_{\mathbb{S}^{n-1}} |\sigma_1|\,\mathrm{d}\mathcal{H}^{n-1}_\sigma\right) x\cdot y,\end{aligned} \tag{3.8}$$

where the last equality comes from a direct computation (see [29, Lemma 3.13]). □

3.2. **Proof of Theorem 3.2.** We start by "localizing" the Euler–Lagrange equation implied by the (Λ, r_0)-minimality condition and the first variation of $\mathcal{F}_{\varepsilon,\gamma}$ given by Lemma 2.5.

Lemma 3.4. *Under the assumptions of Theorem 3.1, there exists $C = C(n, \gamma) > 0$ such that for every $\varphi \in C^1_c(D_1)$ we have (with a slight abuse of notation we identify φ with a function of $\mathbb{R}^n$)*

$$\left| \int_{\partial E \cap \mathbf{C}_2} (\nabla\varphi \cdot \nu_E)(\nu_E \cdot e_n) + 2\gamma \int_{\partial E \cap \mathbf{C}_2} \int_{E \cap \mathbf{C}_2} G_\varepsilon(x - y)(\varphi(x) - \varphi(y))(\nu_E(y) \cdot e_n) \, \mathrm{d}x \, \mathrm{d}\mathcal{H}^{n-1}_y \right|$$
$$\leq C \left(Q\left(\frac{1}{4\varepsilon}\right) + \Lambda \right) \|\nabla\varphi\|_{L^\infty}. \quad (3.9)$$

Proof. By Proposition 2.14 we may choose $\tau_{\text{lip}} = \tau_{\text{lip}}(n, \gamma)$ small enough so that

$$\left\{ x_n < -\frac{1}{4} \right\} \cap \mathbf{C}_2 \subseteq E \cap \mathbf{C}_2 \subseteq \left\{ x_n < \frac{1}{4} \right\} \cap \mathbf{C}_2. \quad (3.10)$$

To simplify notation we write ν for ν_E and recall the convention T_x for $T(x)$. We may assume without loss of generality that $\|\nabla\varphi\|_{L^\infty} = 1$. We start with the following simple observation. For every measure μ and every sets A, B such that $A \times B \subseteq \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \, : \, |x - y| > 1/4\}$,

$$\int_{A \times B} G_\varepsilon(x - y) \mathrm{d}\mu(x) \, \mathrm{d}y \leq 4\mu(A) \int_{\mathbb{R}^n \setminus B_{\frac{1}{4}}} |z| G_\varepsilon(z) \, \mathrm{d}z = 4\mu(A) Q\left(\frac{1}{4\varepsilon}\right). \quad (3.11)$$

Let now $\alpha \in C^1_c((-1, 1); [0, 1])$ be such that $\alpha \equiv 1$ in $(-\frac{1}{2}, \frac{1}{2})$ and $\|\alpha'\|_{L^\infty} \leq 4$. We then consider the vector field $T \in C^1_c(\mathbf{C}_1)$ defined by $T(x) = \varphi(x')\alpha(x_n)e_n$ for all $x \in \mathbb{R}^n$. We first claim that

$$\left| \int_{\partial E} \nu \cdot (\nabla T \nu) + 2\gamma \int_{\partial E \cap \mathbf{C}_2} \int_{E \cap \mathbf{C}_2} G_\varepsilon(x - y)(T_x - T_y) \cdot \nu_y \, \mathrm{d}x \, \mathrm{d}\mathcal{H}^{n-1}_y \right| \leq C \left(Q\left(\frac{1}{4\varepsilon}\right) + \Lambda \right). \quad (3.12)$$

This would conclude the proof of (3.9) since $T(x) = \varphi(x')e_n$ in $D_2 \times (-\frac{1}{2}, \frac{1}{2})$ and

$$\left| \int_{\partial E \cap \mathbf{C}_2} \int_{E \cap \left(\mathbf{C}_2 \setminus \left(D_2 \times (-\frac{1}{2}, \frac{1}{2})\right)\right)} G_\varepsilon(x - y)(T_x - \varphi_{x'} e_n) \cdot \nu_y \, \mathrm{d}x \, \mathrm{d}\mathcal{H}^{n-1}_y \right|$$
$$\leq C \int_{\partial E \cap \mathbf{C}_2} \int_{E \cap \left(\mathbf{C}_2 \setminus \left(D_2 \times (-\frac{1}{2}, \frac{1}{2})\right)\right)} G_\varepsilon(x - y) \, \mathrm{d}x \, \mathrm{d}\mathcal{H}^{n-1}_y \overset{(3.11)}{\leq} CQ\left(\frac{1}{4\varepsilon}\right),$$

where we used (3.11) with $\mu = \mathcal{H}^{n-1} \llcorner \partial E$ and the fact that $P(E; \mathbf{C}_2) \leq C$ by (2.16).
We thus prove (3.12). Notice that $\operatorname{div} T(x) = \varphi(x')\alpha'(x_n)$. In particular, by (3.10), $\operatorname{div} T$ vanishes in

$$(\partial E \cap \mathbf{C}_1) \cup \left(E \cap \left\{ x_n \leq -1 \text{ or } x_n \geq -\frac{1}{2} \right\} \right).$$

By (Λ, r_0)-minimality of E, setting $f_t(x) = x + tT(x)$ we have

$$\mathcal{F}_{\varepsilon,\gamma}(E) \leq \mathcal{F}_{\varepsilon,\gamma}(f_t(E)) + \Lambda |E \triangle f_t(E)|. \quad (3.13)$$

On the one hand, it is standard that for any $|t|$ small enough

$$|E \triangle f_t(E)| \leq 2|t| \left| \int_{\partial E} T \cdot \nu \right| \overset{(2.16)}{\leq} C|t|.$$

On the other hand, for any $|t|$ small enough, we have

$$\mathcal{F}_{\varepsilon,\gamma}(f_t(E)) \leq \mathcal{F}_{\varepsilon,\gamma}(E) + t\big(\delta\mathcal{F}_{\varepsilon,\gamma}(E)[T]\big) + o(t).$$

Hence, by Lemma 2.5, (3.13) implies, for any $|t|$ small enough

$$-t \left[\int_{\partial^* E} \operatorname{div}_E T \, \mathrm{d}\mathcal{H}^{n-1} - 2\gamma \left(\int_{E \times E^c} \operatorname{div} T(x) G_\varepsilon(x - y) \, \mathrm{d}x \, \mathrm{d}y \right.\right.$$
$$\left.\left. + \int_{\partial^* E} \int_E G_\varepsilon(x - y) \left(T(x) - T(y)\right) \cdot \nu_E(y) \, \mathrm{d}x \, \mathrm{d}\mathcal{H}^{n-1}_y \right) \right] \leq C|t|(\Lambda + o(1)).$$

Since this holds for $\pm t$ and for arbitrary small $|t|$, in terms of T this gives

$$\left|\int_{\partial E}\nu\cdot(\nabla T\nu)+2\gamma\left(\int_{E\cap\mathbf{C}_1\cap\{x_n\le-\frac{1}{2}\}}\int_{E^c}\operatorname{div}T(x)G_\varepsilon(x-y)\,\mathrm{d}y\,\mathrm{d}x\right.\right.$$
$$\left.\left.+\int_{\partial E}\int_E G_\varepsilon(x-y)(T_x-T_y)\cdot\nu_y\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y\right)\right|\le C\Lambda. \quad (3.14)$$

Using again that $P(E;\mathbf{C}_2)\le C$ by (2.16) and (3.11) with $A=E\cap\mathbf{C}_1\cap\{x_n\le-\frac{1}{2}\}$, $B=E^c$ and μ the Lebesgue measure (recall that $E^c\cap\mathbf{C}_2\subseteq\{(x',x_n)\ :\ x_n\ge-\frac{1}{4}\}$) we see that in order to prove (3.12) it is enough to show that

$$\left|\int_{\partial E}\int_E G_\varepsilon(x-y)(T_x-T_y)\cdot\nu_y\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y\right.$$
$$\left.-\int_{\partial E\cap\mathbf{C}_2}\int_{E\cap\mathbf{C}_2}G_\varepsilon(x-y)(T_x-T_y)\cdot\nu_y\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y\right|\le CQ\left(\frac{1}{4\varepsilon}\right). \quad (3.15)$$

Recalling that $T=0$ in $\mathbf{C}_1^c$ we write

$$\int_{\partial E}\int_E G_\varepsilon(x-y)(T_x-T_y)\cdot\nu_y\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y-\int_{\partial E\cap\mathbf{C}_2}\int_{E\cap\mathbf{C}_2}G_\varepsilon(x-y)(T_x-T_y)\cdot\nu_y\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y$$
$$=-\int_{\partial E\cap\mathbf{C}_1}\int_{E\setminus\mathbf{C}_2}G_\varepsilon(x-y)T_y\cdot\nu_y\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y+\int_{\partial E\setminus\mathbf{C}_2}\int_{E\cap\mathbf{C}_1}G_\varepsilon(x-y)T_x\cdot\nu_y\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y.$$

Considering the last term on the right-hand side and using integration by parts we have

$$\int_{\partial E\setminus\mathbf{C}_2}\int_{E\cap\mathbf{C}_1}G_\varepsilon(x-y)T_x\cdot\nu_y\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y$$
$$=\int_{E\cap\partial\mathbf{C}_2}\int_{E\cap\mathbf{C}_1}G_\varepsilon(x-y)T_x\cdot\nu_{\mathbf{C}_2}(y)\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y-\int_{E\setminus\mathbf{C}_2}\int_{E\cap\mathbf{C}_1}\nabla G_\varepsilon(x-y)\cdot T_x\,\mathrm{d}x\,\mathrm{d}y.$$

Using Fubini's theorem and integration by parts again leads to

$$\int_{E\setminus\mathbf{C}_2}\int_{E\cap\mathbf{C}_1}\nabla G_\varepsilon(x-y)\cdot T_x\,\mathrm{d}x\,\mathrm{d}y$$
$$=\int_{E\cap\mathbf{C}_1}\int_{E\setminus\mathbf{C}_2}\nabla G_\varepsilon(x-y)\cdot T_x\,\mathrm{d}y\,\mathrm{d}x$$
$$=\int_{\partial E\cap\mathbf{C}_1}\int_{E\setminus\mathbf{C}_2}G_\varepsilon(x-y)T_x\cdot\nu_x\,\mathrm{d}y\,\mathrm{d}\mathcal{H}^{n-1}_x-\int_{E\cap\mathbf{C}_1}\int_{E\setminus\mathbf{C}_2}G_\varepsilon(x-y)\operatorname{div}T(x)\,\mathrm{d}y\,\mathrm{d}x.$$

Putting everything together we find

$$\int_{\partial E}\int_E G_\varepsilon(x-y)(T_x-T_y)\cdot\nu_y\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y-\int_{\partial E\cap\mathbf{C}_2}\int_{E\cap\mathbf{C}_2}G_\varepsilon(x-y)(T_x-T_y)\cdot\nu_y\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y$$
$$=-2\int_{\partial E\cap\mathbf{C}_1}\int_{E\setminus\mathbf{C}_2}G_\varepsilon(x-y)T_y\cdot\nu_y\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y+\int_{E\cap\partial\mathbf{C}_2}\int_{E\cap\mathbf{C}_1}G_\varepsilon(x-y)T_x\cdot\nu_{\mathbf{C}_2}(y)\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y$$
$$+\int_{E\cap\mathbf{C}_1}\int_{E\setminus\mathbf{C}_2}G_\varepsilon(x-y)\operatorname{div}T(x)\,\mathrm{d}y\,\mathrm{d}x.$$

Using (3.11) with either $A=\partial E\cap\mathbf{C}_1$, $B=E\setminus\mathbf{C}_2$, $\mu=\mathcal{H}^{n-1}\llcorner\partial E$ (and $P(E;\mathbf{C}_1)\le C$), $A=E\cap\partial\mathbf{C}_2$, $B=E\cap\mathbf{C}_1$, $\mu=\mathcal{H}^{n-1}\llcorner\partial\mathbf{C}_2$ or $A=E\cap\mathbf{C}_1$, $B=E\setminus\mathbf{C}_2$ and μ the Lebesgue measure we conclude the proof of (3.15). □

We now pass transfer this information to the graph of u.

Lemma 3.5. *Under the assumptions of Theorem 3.1, there exists $C = C(n, \gamma) > 0$ such that for every $\varphi \in C_c^1(D_1)$ we have*

$$\begin{aligned}\left| \int_{\Gamma_u} (\nabla\varphi \cdot \nu_{E_u})(\nu_{E_u} \cdot e_n) + 2\gamma \int_{\Gamma_u} \int_{E_u} G_\varepsilon(x-y)(\varphi(x) - \varphi(y))(\nu_{E_u}(y) \cdot e_n) \, \mathrm{d}x \, \mathrm{d}\mathcal{H}^{n-1}_y \right| \\ \leq C \left(\mathbf{e}_n(4) + Q\left(\frac{1}{4\varepsilon}\right)\right),\end{aligned} \tag{3.16}$$

where

$$E_u := \Big\{ (x', x_n) \ : \ x' \in D_2 \text{ and } x_n < u(x') \Big\}.$$

Proof. As above we may assume without loss of generality that $\|\nabla\varphi\|_{L^\infty} = 1$. To simplify notation we write ν for ν_E and ν^u for ν_{E_u} and will use the convention $T_x = T(x)$. We recall that $M = \partial E \cap \mathbf{C}_2$. Since it is classical (see e.g. the proof of [26, Theorem 23.7]) that

$$\left| \int_{\Gamma_u} (\nabla\varphi \cdot \nu_{E_u})(\nu_{E_u} \cdot e_n) - \int_M (\nabla\varphi \cdot \nu_E)(\nu_E \cdot e_n) \right| \leq C\mathbf{e}_n(4),$$

from (3.9) it is enough to prove that

$$\begin{aligned}\left| \int_{\Gamma_u} \int_{E_u} G_\varepsilon(x-y)(\varphi_x - \varphi_y)(\nu^u_y \cdot e_n) \, \mathrm{d}x \, \mathrm{d}\mathcal{H}^{n-1}_y - \int_M \int_{E \cap \mathbf{C}_2} G_\varepsilon(x-y)(\varphi_x - \varphi_y)(\nu_y \cdot e_n) \, \mathrm{d}x \, \mathrm{d}\mathcal{H}^{n-1}_y \right| \\ \leq C\mathbf{e}_n(4)\|\nabla\varphi\|_{L^\infty}.\end{aligned} \tag{3.17}$$

To this aim we write

$$\begin{aligned}&\int_{\Gamma_u} \int_{E_u} G_\varepsilon(x-y)(\varphi_x - \varphi_y)(\nu^u_y \cdot e_n) \, \mathrm{d}x \, \mathrm{d}\mathcal{H}^{n-1}_y - \int_M \int_{E \cap \mathbf{C}_2} G_\varepsilon(x-y)(\varphi_x - \varphi_y)(\nu_y \cdot e_n) \, \mathrm{d}x \, \mathrm{d}\mathcal{H}^{n-1}_y \\ &= \int_{\Gamma_u} \left(\int_{E_u} G_\varepsilon(x-y)(\varphi_x - \varphi_y) \, \mathrm{d}x - \int_{E \cap \mathbf{C}_2} G_\varepsilon(x-y)(\varphi_x - \varphi_y) \, \mathrm{d}x \right) (\nu^u_y \cdot e_n) \, \mathrm{d}\mathcal{H}^{n-1}_y \\ &\qquad + \int_{\Gamma_u} \int_{E \cap \mathbf{C}_2} G_\varepsilon(x-y)(\varphi_x - \varphi_y)(\nu^u_y \cdot e_n) \, \mathrm{d}x \, \mathrm{d}\mathcal{H}^{n-1}_y \\ &\qquad\quad - \int_M \int_{E \cap \mathbf{C}_2} G_\varepsilon(x-y)(\varphi_x - \varphi_y)(\nu_y \cdot e_n) \, \mathrm{d}x \, \mathrm{d}\mathcal{H}^{n-1}_y.\end{aligned}$$

We claim that

$$\int_{\Gamma_u} \int_{E_u \triangle (E \cap \mathbf{C}_2)} G_\varepsilon(x-y)|\varphi_x - \varphi_y| \, \mathrm{d}x \, \mathrm{d}\mathcal{H}^{n-1}_y \leq C\mathbf{e}_n(4) \tag{3.18}$$

and

$$\begin{aligned}\Bigg| \int_{\Gamma_u} \int_{E \cap \mathbf{C}_2} G_\varepsilon(x-y)(\varphi_x - \varphi_y)(\nu^u_y \cdot e_n) \, \mathrm{d}x \, \mathrm{d}\mathcal{H}^{n-1}_y \\ - \int_M \int_{E \cap \mathbf{C}_2} G_\varepsilon(x-y)(\varphi_x - \varphi_y)(\nu_y \cdot e_n) \, \mathrm{d}x \, \mathrm{d}\mathcal{H}^{n-1}_y \Bigg| \leq C\mathbf{e}_n(4),\end{aligned} \tag{3.19}$$

from which (3.17) would follow. We start with (3.18).
By (ii) of Theorem 3.1, there exists a set $A \subseteq D_2$ such that $\mathcal{H}^{n-1}(A) \leq C\mathbf{e}_n(4)$ and

$$E \cap \big\{(x', t) \ : \ t \in (-2, 2)\big\} = E_u \cap \big\{(x', t) \ : \ t \in (-2, 2)\big\}, \quad \forall x' \in D_2 \setminus A,$$

since

$$\Big\{ y' \in D_2 \ : \ \Pi_n^{-1}(\{y'\}) \cap E \cap \mathbf{C}_2 \neq \Pi_n^{-1}(\{y'\}) \cap \Gamma_u^- \cap \mathbf{C}_2 \Big\} = \Pi_n(M \triangle \Gamma_u),$$

where $\Pi_n : (y', y_n) \mapsto y_n$. Thus, since φ and u are Lipschitz continuous we find

$$\begin{aligned}\int_{\Gamma_u}\int_{E_u\triangle(E\cap\mathbf{C}_2)} & G_\varepsilon(x-y)|\varphi_x-\varphi_y|\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y \\ &\le \int_{\Gamma_u}\int_A\int_{-2}^{2} G_\varepsilon((x',t)-y)|(x',t)-y|\,\mathrm{d}t\,\mathrm{d}x'\,\mathrm{d}\mathcal{H}^{n-1}_y \\ &\le C\int_A\int_{D_2}\int_{\mathbb{R}} G_\varepsilon((x',t)-u_{y'})|(x',t)-u_{y'}|\,\mathrm{d}t\,\mathrm{d}y'\,\mathrm{d}x' \\ &\le C\mathcal{H}^{n-1}(A)\int_{\mathbb{R}^n}|z|G(z)\,\mathrm{d}z \le C\mathcal{H}^{n-1}(A).\end{aligned}$$

We now turn to (3.19). Notice that $\mathcal{H}^{n-1}$-a.e. on $\Gamma_u\cap M$ we have $\nu^u = \pm\nu$. Moreover, setting $\Gamma_1 \coloneqq \Gamma_u\cap M\cap\{\nu^u=\nu\}$ and arguing exactly as in [26, (23.51)], we have $\mathcal{H}^{n-1}((M\cap\Gamma_u)\setminus\Gamma_1)\le C\mathbf{e}_n(4)$. Recalling that by (ii) of Theorem 3.1, $\mathcal{H}^{n-1}(M\triangle\Gamma_u)\le C\mathbf{e}_n(4)$, we find

$$\begin{aligned}\Bigg|\int_{\Gamma_u}\int_{E\cap\mathbf{C}_2} & G_\varepsilon(x-y)(\varphi_x-\varphi_y)(\nu^u_y\cdot e_n)\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y \\ &\qquad - \int_M\int_{E\cap\mathbf{C}_2} G_\varepsilon(x-y)(\varphi_x-\varphi_y)(\nu_y\cdot e_n)\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y\Bigg| \\ &\le \int_{(M\triangle\Gamma_u)\cup((M\cap\Gamma_u)\setminus\Gamma_1)}\int_{E\cap\mathbf{C}_2} G_\varepsilon(x-y)|x-y|\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_y \\ &\le C\mathbf{e}_n(4)\int_{\mathbb{R}^n}|z|G_\varepsilon(z)\,\mathrm{d}z \le C\mathbf{e}_n(4).\end{aligned}$$

□

In order to conclude the proof of Theorem 3.2, we are left with the linearization of (3.16).

Proof of Theorem 3.2. Since arguing verbatim as in [26, Theorem 23.7] we have

$$\left|\int_{D_1}\nabla u\cdot\nabla\varphi - \int_{\Gamma_u}(\nabla\varphi\cdot\nu_{E_u})(\nu_{E_u}\cdot e_n)\right| \le C\mathbf{e}_n(4),$$

by Lemma 3.5 it is enough to prove that (recall the notation $u_{x'} = u(x')$)

$$\begin{aligned}\Bigg|\int_{D_2}\int_{D_2}\int_{-2}^{u_{x'}} & G_\varepsilon(x'-y', t-u_{y'})(\varphi_{x'}-\varphi_{y'})\,\mathrm{d}t\,\mathrm{d}x'\,\mathrm{d}y' \\ &- \int_{D_2\times D_2}(u_{x'}-u_{y'})(\varphi_{x'}-\varphi_{y'})G_\varepsilon(x'-y',0)\,\mathrm{d}x'\,\mathrm{d}y'\Bigg| \le C\left(\mathbf{e}_n(4)+Q\left(\frac{1}{4\varepsilon}\right)\right).\end{aligned} \tag{3.20}$$

For $x'\neq y'$ we have

$$\begin{aligned}\int_{-2}^{u_{x'}} G_\varepsilon(x'-y', t-u_{y'})\,\mathrm{d}t &= \int_{-2-u_{y'}}^{u_{x'}-u_{y'}} G_\varepsilon(x'-y',s)\,\mathrm{d}s \\ &= \int_{-2-u_{y'}}^{-1} G_\varepsilon(x'-y',s)\,\mathrm{d}s + \int_{-1}^{0} G_\varepsilon(x'-y',s)\,\mathrm{d}s + \int_0^{u_{x'}-u_{y'}} G_\varepsilon(x'-y',s)\,\mathrm{d}s.\end{aligned} \tag{3.21}$$

On the one hand we observe that

$$\int_{D_2\times D_2}(\varphi_{x'}-\varphi_{y'})\int_{-1}^{0} G_\varepsilon(x'-y',s)\,\mathrm{d}s\,\mathrm{d}x'\,\mathrm{d}y' = 0. \tag{3.22}$$

On the other hand, since $\|u\|_{L^\infty} \le 1$, for any y' we have $-2 - u_{y'} < -1$. Thus, using the fact that φ is 1-Lipschitz, we compute

$$\begin{aligned}\left|\int_{D_2}\int_{D_2}\int_{-2-u_{y'}}^{-1} G_\varepsilon(x'-y',s)(\varphi_{x'}-\varphi_{y'})\,\mathrm{d}t\,\mathrm{d}x'\,\mathrm{d}y'\right| \\ \le \int_{D_2}\int_{-2-u_{y'}}^{-1}\int_{D_2} |x'-y'|G_\varepsilon(x'-y',s)\,\mathrm{d}x'\,\mathrm{d}s\,\mathrm{d}y' \\ \le \int_{D_2}\int_{\mathbb{R}^n\setminus B_1} |z|G_\varepsilon(z)\,\mathrm{d}z\,\mathrm{d}y' \le CQ\left(\frac{1}{4\varepsilon}\right).\end{aligned} \tag{3.23}$$

Combining (3.21), (3.22) and (3.23) yields

$$\begin{aligned}\Bigg|\int_{D_2}\int_{D_2}\int_{-2}^{u_{x'}} G_\varepsilon(x'-y',t-u_{y'})(\varphi_{x'}-\varphi_{y'})\,\mathrm{d}t\,\mathrm{d}x'\,\mathrm{d}y' \\ -\int_{D_2\times D_2}\int_0^{u_{x'}-u_{y'}} G_\varepsilon(x'-y',t)(\varphi_{x'}-\varphi_{y'})\,\mathrm{d}t\,\mathrm{d}x'\,\mathrm{d}y'\Bigg| \le CQ\left(\frac{1}{4\varepsilon}\right).\end{aligned} \tag{3.24}$$

Using again that φ is 1-Lipschitz and Fubini's theorem, we estimate

$$\begin{aligned}\left|\int_{D_2\times D_2}\int_0^{u_{x'}-u_{y'}} (G_\varepsilon(x'-y',t)-G_\varepsilon(x'-y',0))(\varphi_{x'}-\varphi_{y'})\,\mathrm{d}t\,\mathrm{d}x'\,\mathrm{d}y'\right| \\ \le \int_0^1\int_{D_2\times D_2}\int_0^{|u_{x'}-u_{y'}|} t\,|x'-y'||\nabla G_\varepsilon(x'-y',st)|\,\mathrm{d}t\,\mathrm{d}x'\,\mathrm{d}y'\,\mathrm{d}s \\ = \int_0^1\int_0^1 t\int_{D_2\times D_2} |u_{x'}-u_{y'}|^2\,|x'-y'||\nabla G_\varepsilon(x'-y',st|u_{x'}-u_{y'}|)|\,\mathrm{d}x'\,\mathrm{d}y'\,\mathrm{d}t\,\mathrm{d}s \\ \le \int_0^1\int_0^1\int_{D_2\times D_2} |u_{x'}-u_{y'}|^2\,|x'-y'||\nabla G_\varepsilon(x'-y',st|u_{x'}-u_{y'}|)|\,\mathrm{d}x'\,\mathrm{d}y'\,\mathrm{d}t\,\mathrm{d}s.\end{aligned} \tag{3.25}$$

Set $\widetilde{G}_\varepsilon := \varepsilon^{-(n+1)}\widetilde{G}(\cdot/\varepsilon)$ where $\widetilde{G} := |\cdot||\nabla G|$ and $\Phi_{stu}(x',y') := (x'-y', st(u_{x'}-u_{y'}))$. Observing that $|\Phi_{stu}(x',y')| \ge |x'-y'|$ we have for every fixed s,t,

$$\begin{aligned}\int_{D_2\times D_2} |u_{x'}-u_{y'}|^2\,|x'-y'||\nabla G_\varepsilon(x'-y',st|u_{x'}-u_{y'}|)|\,\mathrm{d}x'\,\mathrm{d}y' \\ \le \int_{D_2\times D_2} |u_{x'}-u_{y'}|^2\widetilde{G}_\varepsilon(|\Phi_{stu}(x',y')|)\,\mathrm{d}x'\,\mathrm{d}y'.\end{aligned}$$

Observing that $I^1_{\widetilde{G}} = I^2_{|\nabla G|}$, Lemma 3.6 below yields

$$\int_{D_2\times D_2} |u_{x'}-u_{y'}|^2\widetilde{G}_\varepsilon(|\Phi_{stu}(x',y')|)\,\mathrm{d}x'\,\mathrm{d}y' \le CI^2_{|\nabla G|}\int_{D_2}|\nabla u|^2 \le CI^2_{|\nabla G|}\mathbf{e}_n(14),$$

where we used that by (iii) of Theorem 3.1, $\int_{D_2}|\nabla u|^2 \le C\mathbf{e}_n(4)$. Combining this with (3.25) and (3.24) concludes the proof of (3.20).

□

In the proof of Theorem 3.2 above, we used the following technical lemma.

Lemma 3.6. *Let $G : \mathbb{R}^n \mapsto \mathbb{R}^+$ be a radial kernel such that (recall definition (1.1)) $I^1_G < \infty$. For $u \in \mathrm{Lip}(D_2)$, we define the map $\Phi_u : D_2 \times D_2 \to \mathbb{R}^{n-1}\times\mathbb{R}$ by*

$$\Phi_u(x',y') = (x'-y', u(x')-u(y')). \tag{3.26}$$

There exists a constant $C = C(n) > 0$ such that if $\|\nabla u\|_{L^\infty(D_2)} \le \frac{1}{2}$ then

$$\int_{D_2\times D_2} (u(x')-u(y'))^2 G(\Phi_u(x',y'))\,\mathrm{d}x'\,\mathrm{d}y' \le CI^1_G\int_{D_2}|\nabla u|^2. \tag{3.27}$$

Proof. We start by estimating

$$\begin{aligned}
&\int_{D_2\times D_2}(u(x')-u(y'))^2G(\Phi_u(x',y'))\,\mathrm{d}x'\,\mathrm{d}y'\\
&\quad\le\int_0^1\int_{\mathbb{R}^{n-1}\times\mathbb{R}^{n-1}}\mathbf{1}_{D_2}(x')\mathbf{1}_{D_2}(y')|\nabla u(x'+t(y'-x'))|^2|x'-y'|^2G(\Phi_u(x',y'))\,\mathrm{d}x'\,\mathrm{d}y'\,\mathrm{d}t\\
&\quad=\int_0^1\int_{\mathbb{R}^{n-1}\times\mathbb{R}^{n-1}}\mathbf{1}_{D_2}(\hat{x}'-tz')\mathbf{1}_{D_2}(\hat{x}'+(1-t)z')|\nabla u(\hat{x})|^2|z'|^2G(\Phi_u(\hat{x}'-tz',\hat{x}'+(1-t)z'))\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\mathrm{d}\hat{x}'\,\mathrm{d}z'\,\mathrm{d}t\\
&\quad\le\int_0^1\int_{D_2}|\nabla u(x')|^2\left[\int_{\mathbb{R}^{n-1}}|z'|^2G(\Phi_u(x'-tz',x'+(1-t)z'))\,\mathrm{d}z'\right]\mathrm{d}x'\,\mathrm{d}t,
\end{aligned}$$

where we made the change of variables $z'=y'-x'$, $\hat{x}'=x'+tz'$, and used that by convexity of D_2, $D_2(tz')\cap D_2(-(1-t)z')\subseteq D_2$. We finally claim that for every fixed $t\in[0,1]$ and $x'\in D_2$,

$$\int_{\mathbb{R}^{n-1}}|z'|^2G(\Phi_u(x'-tz',x'+(1-t)z'))\,\mathrm{d}z'\le CI_G^1,\tag{3.28}$$

which would conclude the proof of (3.27). For this we set $G(z)=g(|z|)$ for some $g:\mathbb{R}^+\mapsto\mathbb{R}^+$. and write using polar coordinates

$$\begin{aligned}
&\int_{\mathbb{R}^{n-1}}|z'|^2G(\Phi_u(x'-tz',x'+(1-t)z'))\,\mathrm{d}z'\\
&\qquad\qquad=\int_{\mathbb{S}^{n-2}}\int_0^\infty r^n g\big(\sqrt{r^2+|u(x'-tr\sigma)-u(x'+(1-t)r\sigma)|^2}\big)\,\mathrm{d}r\,\mathrm{d}\mathcal{H}^{n-2}_\sigma.
\end{aligned}\tag{3.29}$$

We finally notice that for every fixed $t\in[0,1]$, $x'\in D_2$ and $\sigma\in\mathbb{S}^{n-2}$, the function $\Psi(r):=\sqrt{r^2+|u(x'-tr\sigma)-u(x'+(1-t)r\sigma)|^2}$ is Lipschitz continuous with

$$\frac{\sqrt{5}}{2}r\ge\Psi(r)\ge r\qquad\text{and}\qquad\frac{5}{4}\ge\Psi'(r)\ge\frac{3}{2\sqrt{5}}$$

so that making the change of variables $s=\Psi(r)$ we find

$$\int_{\mathbb{R}^{n-1}}|z'|^2G(\Phi_u(x'-tz',x'+(1-t)z'))\,\mathrm{d}z'\le C\int_0^\infty s^n g(s)\,\mathrm{d}s=CI_G^1.$$

This concludes the proof of (3.28). □

4. Caccioppoli inequality

Let us first introduce the standard notion of *flatness* for sets of finite perimeter.

Definition 4.1 (Flatness). For any set of finite perimeter $E\subseteq\mathbb{R}^n$ we define the flatness of E in $x\in\partial E$ at scale $r>0$ with respect to the direction $\nu\in\mathbb{S}^{n-1}$ by

$$\mathbf{f}(E,x,r,\nu):=\inf_{c\in\mathbb{R}}\frac{1}{r^{n-1}}\int_{\partial^*E\cap\mathbf{C}(x,r,\nu)}\frac{|(y-x)\cdot\nu-c|^2}{r^2}\,\mathrm{d}\mathcal{H}^{n-1}_y.$$

When $\nu=e_n$, we write $\mathbf{f}_n(E,x,r)$ for $\mathbf{f}(E,x,r,e_n)$ and we write $\mathbf{f}_n(E,r)$ for $\mathbf{f}_n(E,0,r)$.

Using the harmonic approximation result given by Proposition 3.3, we will be able to show in Lemma 5.3 that there exists a direction ν such that $\mathbf{f}(E,\lambda r,\nu)\lesssim\lambda^2\mathbf{e}_n(E,r)$ for (Λ,r_0)-minimizers of $\mathcal{F}_{\varepsilon,\gamma}$, as long as r is much larger than ε. To pass this estimate to the excess at scale $\lambda r/2$, we prove in this section a Caccioppoli-type (or Reverse Poincaré) inequality. The key argument is to prove first that for sets which are sufficiently flat, the quasi-minimality condition (2.15) can be upgraded.

To that effect, we need to introduce some notation. For any $t>0$ and $z\in\mathbb{R}^{n-1}$, we define $\mathbf{K}_t(z):=D_t(z)\times(-1,1)$, and we simply write $\mathbf{K}_t$ when $z=0$. For any cylinder $\mathbf{K}_t(z)$, any set of locally finite perimeter E, and any constant $c\in\mathbb{R}$, we define the quantities

$$\mathscr{F}(E,\mathbf{K}_t(z),c):=\int_{\mathbf{K}_t(z)\cap\partial^*E}\frac{(x_n-c)^2}{t^2}\,\mathrm{d}\mathcal{H}^{n-1}\tag{4.1}$$

and

$$\mathscr{E}(E,\mathbf{K}_t(z)) \coloneqq P(E;\mathbf{K}_t(z)) - \mathcal{H}^{n-1}(D_t(z)). \tag{4.2}$$

When $z=0$, we make the abuse of notation $\mathscr{E}(E,t)=\mathscr{E}(E,\mathbf{K}_t(0))$ and $\mathscr{F}(E,t,c)=\mathscr{E}(E,\mathbf{K}_t(0),c)$. Let us point out that these two quantities are respectively linked with the (non-scale-invariant) flatness and excess of E at scale t in the direction e_n. Indeed, if $0\in\partial E$ and if for some $h\in(0,t)$

$$\{(x',x_n)\in\mathbf{K}_t\ :\ x_n<-h\}\subseteq E\cap\mathbf{C}_t\subseteq\{(x',x_n)\in\mathbf{K}_t\ :\ x_n<h\},$$

then

$$\mathbf{f}_n(E,t)=\inf_{c\in\mathbb{R}}\frac{1}{t^{n-1}}\mathscr{F}(E,t,c),$$

and

$$\mathcal{H}^{n-1}(D_t)=\int_{\partial^*E\cap\mathbf{C}_t}\nu_E\cdot e_n$$

(see [26, Lemma 22.11]), thus, for any $t\in(0,1)$,

$$\mathscr{E}(E,t)=\int_{\partial^*E\cap\mathbf{C}_t}(1-\nu_E\cdot e_n)\,\mathrm{d}\mathcal{H}^{n-1}=\frac12\int_{\partial^*E\cap\mathbf{C}_t}|\nu_E-e_n|^2\,\mathrm{d}\mathcal{H}^{n-1}=\left(\frac{t^{n-1}}{2}\right)\mathbf{e}_n(E,t). \tag{4.3}$$

Notice in particular that $\mathscr{E}(E,\cdot)$ is increasing in $(0,1)$. Eventually, recalling the definition of the function Q in (1.3), for any $\theta\in[0,1]$ we define the function $Q_{1-\theta}$ by

$$Q_{1-\theta}(t)\coloneqq Q(t^{1-\theta}),\qquad\forall t>0. \tag{4.4}$$

4.1. **A refined quasi-minimality condition.** We improve the quasi-minimality condition (2.15) for sets which are sufficiently flat. For any $\varepsilon>0$, let us define the "critical" energy functional

$$\mathcal{F}_\varepsilon(E)\coloneqq\mathcal{F}_{1,\varepsilon}(E)=(P-P_\varepsilon)(E).$$

Proposition 4.2. *Assume that G satisfies (H1) and (H2), and let $\varepsilon\in(0,1)$, $\gamma\in(0,1)$, $\theta\in(0,1]$ and $\Lambda>0$ with $4\Lambda\le1-\gamma$. There exists $C=C(n)>0$ such that if E is a $(\Lambda,4)$-minimizer of $\mathcal{F}_{\varepsilon,\gamma}$ with*

$$\left\{x_n<-\frac14\right\}\cap\mathbf{K}_3\subseteq E\cap\mathbf{K}_3\subseteq\left\{x_n<\frac14\right\}\cap\mathbf{K}_3,$$

then the following holds. If $t\in(0,2)$ is such that $\mathcal{H}^{n-1}(\partial\mathbf{K}_t\cap\partial E)=0$ then for any set F of finite perimeter such that $E\triangle F\subseteq\mathbf{K}_t$ and

$$\left\{x_n<-\frac14\right\}\cap\mathbf{K}_t\subseteq F\cap\mathbf{K}_t\subseteq\left\{x_n<\frac14\right\}\cap\mathbf{K}_t,$$

we have

$$\begin{aligned}\mathscr{E}(E,t)\le\left(\frac{1+\gamma}{1-\gamma}\right)\mathscr{E}(F,t)&+\frac{2\gamma}{(1-\gamma)}\big[\mathscr{E}(E,t+\varepsilon^\theta)-\mathscr{E}(E,t)\big]\\&+\frac{C}{(1-\gamma)}Q_{1-\theta}\left(\frac1\varepsilon\right)+\frac{\Lambda}{(1-\gamma)}|E\triangle F|+\frac{1+3\gamma}{(1-\gamma)}\mathcal{H}^{n-1}(\partial^*F\cap\partial\mathbf{K}_t).\end{aligned} \tag{4.5}$$

Proof. To simplify a bit notation set $\eta\coloneqq\mathcal{H}^{n-1}(\partial^*F\cap\partial\mathbf{K}_t)$. Since $\mathcal{H}^{n-1}(\partial\mathbf{K}_t\cap\partial E)=0$ and $E\triangle F\subseteq\mathbf{K}_t$ we have

$$P(E)-P(F)=P(E;\mathbf{K}_t)-P(F;\mathbf{K}_t)-\eta. \tag{4.6}$$

By $(\Lambda,4)$-minimality of E we find

$$(1-\gamma)P(E;\mathbf{K}_t)\le(1-\gamma)P(F;\mathbf{K}_t)+\gamma\big[\mathcal{F}_\varepsilon(F)-\mathcal{F}_\varepsilon(E)\big]+\Lambda|E\triangle F|+(1-\gamma)\eta. \tag{4.7}$$

In the next two steps we prove that

$$\mathcal{F}_\varepsilon(F)-\mathcal{F}_\varepsilon(E)\le2\mathscr{E}(F,t+\varepsilon^\theta)+CQ_{1-\theta}\left(\frac1\varepsilon\right). \tag{4.8}$$

Step 1. In this first step we localize the estimate. Setting for simplicity

$$\widetilde{D}_t\coloneqq D_{t+\varepsilon^\theta},\qquad\widetilde{\mathbf{K}}_t\coloneqq\mathbf{K}_{t+\varepsilon^\theta}$$

and defining the "localized" functional

$$\mathcal{F}_\varepsilon^{\mathrm{loc}}(E)\coloneqq P(E;\widetilde{\mathbf{K}}_t)-\int_{(E\cap\widetilde{\mathbf{K}}_t)\times(E^c\cap\widetilde{\mathbf{K}}_t)}G_\varepsilon(x-y)\,\mathrm{d}x\,\mathrm{d}y,$$

we claim that

$$\mathcal{F}_\varepsilon(F) - \mathcal{F}_\varepsilon(E) \leq \mathcal{F}^{\mathrm{loc}}_\varepsilon(F) - \mathcal{F}^{\mathrm{loc}}_\varepsilon(E) + CQ_{1-\theta}(\varepsilon^{-1}). \tag{4.9}$$

Since $E \triangle F \subseteq \mathbf{K}_t \subset\subset \widetilde{\mathbf{K}}_t$, $P(E) - P(F) = P(E;\widetilde{\mathbf{K}}_t) - P(F;\widetilde{\mathbf{K}}_t)$ and thus in order to prove (4.9), we just need to consider the nonlocal part. Setting $\Phi(A,B) = \int_{A\times B} G_\varepsilon(x-y)\,\mathrm{d}x\,\mathrm{d}y$ and using that $E\triangle F \subseteq \mathbf{K}_t$ we have

$$\begin{aligned}
P_\varepsilon(E) - P_\varepsilon(F) &= \Phi(E,E^c) - \Phi(F,F^c)\\
&= \Big[\Phi(E\cap \widetilde{\mathbf{K}}_t, E^c\cap \widetilde{\mathbf{K}}_t) - \Phi(F\cap \widetilde{\mathbf{K}}_t, F^c\cap \widetilde{\mathbf{K}}_t)\Big] + \Phi(E\cap \mathbf{K}_t, E^c\backslash \widetilde{\mathbf{K}}_t) - \Phi(F\cap \mathbf{K}_t, E^c\backslash \widetilde{\mathbf{K}}_t)\\
&\quad + \Phi(E\backslash \widetilde{\mathbf{K}}_t, E^c\cap \mathbf{K}_t) - \Phi(E\backslash \widetilde{\mathbf{K}}_t, F^c\cap \mathbf{K}_t)\\
&\leq \Big[\Phi(E\cap \widetilde{\mathbf{K}}_t, E^c\cap \widetilde{\mathbf{K}}_t) - \Phi(F\cap \widetilde{\mathbf{K}}_t, F^c\cap \widetilde{\mathbf{K}}_t)\Big] + \Phi(E\cap \mathbf{K}_t, E^c\backslash \widetilde{\mathbf{K}}_t) + \Phi(E\backslash \widetilde{\mathbf{K}}_t, E^c\cap \mathbf{K}_t)\\
&\leq \Big[\Phi(E\cap \widetilde{\mathbf{K}}_t, E^c\cap \widetilde{\mathbf{K}}_t) - \Phi(F\cap \widetilde{\mathbf{K}}_t, F^c\cap \widetilde{\mathbf{K}}_t)\Big] + 2\Phi(\mathbf{K}_t, (\widetilde{\mathbf{K}}_t)^c).
\end{aligned}$$

In order to prove (4.9) it is thus enough to estimate $\Phi(\mathbf{K}_t, (\widetilde{\mathbf{K}}_t)^c)$. For this we write that

$$\begin{aligned}
\Phi(\mathbf{K}_t, (\widetilde{\mathbf{K}}_t)^c) &= \int_{\mathbf{K}_t\times(\widetilde{\mathbf{K}}_t)^c} G_\varepsilon(x-y)\,\mathrm{d}x\,\mathrm{d}y\\
&= \int_{\mathbb{R}^n\times\mathbb{R}^n} \mathbf{1}_{\mathbf{K}_t}(x)\mathbf{1}_{\widetilde{\mathbf{K}}_t^c}(x+z)G_\varepsilon(z)\,\mathrm{d}x\,\mathrm{d}z\\
&= \int_{\mathbb{R}^n\times\mathbb{R}^n} \mathbf{1}_{\mathbf{K}_t}(x)\mathbf{1}_{\widetilde{\mathbf{K}}_t^c}(x+z)\mathbf{1}_{\mathbb{R}^n\backslash B_{\varepsilon^\theta}}(z)G_\varepsilon(z)\,\mathrm{d}x\,\mathrm{d}z\\
&\leq \int_{\mathbb{R}^n\times\mathbb{R}^n} \mathbf{1}_{\mathbf{K}_t}(x)\mathbf{1}_{\mathbf{K}_t^c}(x+z)\mathbf{1}_{\mathbb{R}^n\backslash B_{\varepsilon^\theta}}(z)G_\varepsilon(z)\,\mathrm{d}x\,\mathrm{d}z\\
&\leq \frac{1}{2}\left(\int_{\mathbb{R}^n\backslash B_{\varepsilon^\theta}} |z|G_\varepsilon(z)\,\mathrm{d}z\right)P(\mathbf{K}_t)\\
&\leq CQ_{1-\theta}(\varepsilon^{-1}),
\end{aligned}$$

where we used (2.1) with $K = \mathbf{1}_{\mathbb{R}^n\backslash B_{\varepsilon^\theta}}G_\varepsilon$.

Step 2. In this step we show

$$\mathcal{F}^{\mathrm{loc}}_\varepsilon(F) - \mathcal{F}^{\mathrm{loc}}_\varepsilon(E) \leq 2\mathscr{E}(F, t+\varepsilon^\theta) + CQ_{1-\theta}(\varepsilon^{-1}). \tag{4.10}$$

Together with (4.9) this would conclude the proof of (4.8). To this aim, we will use the slicing techniques introduced in [27, Section 3], rewriting P and P_ε as an average over 1-dimensional slices. Let us set $\rho(t) := \omega_{n-1}|t|^{n-1}g(|t|)$ and $\rho_\varepsilon(t) := \varepsilon^{-2}\rho(\varepsilon^{-1}t)$ for $t\in\mathbb{R}\setminus\{0\}$. For every line segment $L\subseteq\mathbb{R}^n$, we define the one-dimensional nonlocal perimeter functional in L

$$P^{1\mathrm{D}}_\varepsilon(E;L) := \int_{L\times L} |\mathbf{1}_E(x) - \mathbf{1}_E(y)|\rho_\varepsilon(x-y)\,\mathrm{d}\mathcal{H}^1_x\,\mathrm{d}\mathcal{H}^1_y = 2\int_{(E\cap L)\times(E^c\cap L)} \rho_\varepsilon(x-y)\,\mathrm{d}\mathcal{H}^1_x\,\mathrm{d}\mathcal{H}^1_y$$

and the one-dimensional critical energy in L by

$$\mathcal{F}^{1\mathrm{D}}_\varepsilon(E;L) := \mathcal{H}^0(\partial^* E\cap L) - P^{1\mathrm{D}}(E;L).$$

Proceeding as in [27, Proposition 3.1 & Corollary 3.3] (the only difference is the restriction to $\widetilde{\mathbf{K}}_t$) we have

$$\mathcal{F}^{\mathrm{loc}}_\varepsilon(E) = \frac{1}{2\omega_{n-1}}\int_{\mathbb{S}^{n-1}}\int_{\{\sigma\}^\perp} \mathcal{F}^{1\mathrm{D}}_\varepsilon(E;\widetilde{L}_{\sigma,x})\,\mathrm{d}\mathcal{H}^{n-1}_x\,\mathrm{d}\mathcal{H}^{n-1}_\sigma,$$

where we set

$$L_{\sigma,x} := x + \mathbb{R}\sigma, \quad \text{and} \quad \widetilde{L}_{\sigma,x} := L_{\sigma,x}\cap\widetilde{\mathbf{K}}_t.$$

In particular

$$\mathcal{F}^{\mathrm{loc}}_\varepsilon(F) - \mathcal{F}^{\mathrm{loc}}_\varepsilon(E) = \frac{1}{2\omega_{n-1}}\int_{\mathbb{S}^{n-1}}\int_{\{\sigma\}^\perp} (\mathcal{F}^{1\mathrm{D}}_\varepsilon(F;\widetilde{L}_{\sigma,x}) - \mathcal{F}^{1\mathrm{D}}_\varepsilon(E;\widetilde{L}_{\sigma,x}))\,\mathrm{d}x\,\mathrm{d}\mathcal{H}^{n-1}_\sigma. \tag{4.11}$$

Step 2.1. We claim that for every $\sigma\in\mathbb{S}^{n-1}$ and $\mathcal{H}^{n-1}$-a.e $x\in\{\sigma\}^\perp$, we have

$$\mathcal{F}_\varepsilon^{1\mathrm{D}}(F;\widetilde{L}_{\sigma,x}) - \mathcal{F}_\varepsilon^{1\mathrm{D}}(E;\widetilde{L}_{\sigma,x}) \leq \begin{cases} 0 & \text{it } \partial^* F\cap \widetilde{L}_{\sigma,x}=\emptyset, \\ 2\left(\mathcal{H}^0(\partial^* F\cap \widetilde{L}_{\sigma,x})-1\right)+CQ_{1-\theta}(\varepsilon^{-1}) & \text{otherwise.}\end{cases} \tag{4.12}$$

By the standard properties of one-dimensional restrictions of sets of finite perimeter (see e.g. [3]), it is enough to prove (4.12) for $E,F\subseteq\mathbb{R}$ and $\widetilde{L}_{\sigma,x}=L=(0,a)$ for some $a>0$. Notice that since E and F are of finite perimeter in L, they are just a finite union of disjoint intervals.
By [27, Remark 3.2], for any set of finite perimeter $E\subseteq\mathbb{R}$, we have $P_\varepsilon^{1\mathrm{D}}(E;\mathbb{R})\leq\mathcal{H}^0(\partial E)$, which implies $\mathcal{F}_\varepsilon^{1\mathrm{D}}(E;L)\geq\mathcal{F}_\varepsilon^{1\mathrm{D}}(E;\mathbb{R})\geq 0$. Thus, if $\mathcal{H}^0(\partial F\cap L)=0$ (that is, $\partial F\cap L=\emptyset$), then $\mathcal{F}_\varepsilon^{1\mathrm{D}}(F;L)-\mathcal{F}_\varepsilon^{1\mathrm{D}}(E;L)\leq\mathcal{F}_\varepsilon^{1\mathrm{D}}(F;L)\leq -P_\varepsilon^{1\mathrm{D}}(E;L)\leq 0$. If $\mathcal{H}^0(\partial F\cap L)\geq 2$, then

$$\mathcal{F}_\varepsilon^{1\mathrm{D}}(F;L)-\mathcal{F}_\varepsilon^{1\mathrm{D}}(E;L)\leq\mathcal{F}_\varepsilon^{1\mathrm{D}}(F;L)\leq\mathcal{H}^0(\partial F\cap L)\leq 2\left(\mathcal{H}^0(\partial F\cap L)-1\right).$$

There remains to focus on the case where $\mathcal{H}^0(\partial F\cap L)=1$. In this case we claim that

$$\mathcal{F}_\varepsilon^{1\mathrm{D}}(F;L)-\mathcal{F}_\varepsilon^{1\mathrm{D}}(E;L)\leq CQ_{1-\theta}(\varepsilon^{-1}). \tag{4.13}$$

Let t_F be such that $L\cap\partial F=\{t_F\}$ then either $F\cap L=(0,t_F)$ or $F\cap L=(t_F,a)$. Since both cases are similar, we treat only the case $F\cap L=(0,t_F)$. We distinguish two sub-cases.

Case 1: $\mathrm{d}(t_F,L^{\mathrm{c}})\geq\varepsilon^\theta$. In this case we argue somewhat similarly to (4.9). Using the fact that

$$2\int_{-\infty}^{c}\int_{c}^{\infty}\rho_\varepsilon(s-t)\,\mathrm{d}s\,\mathrm{d}t = 2\int_0^\infty t\rho_\varepsilon(t)\,\mathrm{d}t=1,\qquad\forall c\in\mathbb{R}, \tag{4.14}$$

we compute

$$\begin{aligned}\mathcal{F}_\varepsilon^{1\mathrm{D}}(F;L) &= 1-2\int_0^{t_F}\int_{t_F}^{a}\rho_\varepsilon(s-t)\,\mathrm{d}s\,\mathrm{d}t\\ &= 2\int_{-\infty}^{t_F}\int_{t_F}^{\infty}\rho_\varepsilon(s-t)\,\mathrm{d}s\,\mathrm{d}t-2\int_0^{t_F}\int_{t_F}^{a}\rho_\varepsilon(s-t)\,\mathrm{d}s\,\mathrm{d}t\\ &= 2\int_{-\infty}^{t_F}\int_{a}^{\infty}\rho_\varepsilon(s-t)\,\mathrm{d}s\,\mathrm{d}t+2\int_{-\infty}^{0}\int_{t_F}^{a}\rho_\varepsilon(s-t)\,\mathrm{d}s\,\mathrm{d}t\\ &\leq 2\int_{-\infty}^{0}\left(\int_{a-t_F}^{\infty}\rho_\varepsilon(s-t)\,\mathrm{d}s\,\mathrm{d}t+\int_{t_F}^{\infty}\rho_\varepsilon(s-t)\,\mathrm{d}s\right)\mathrm{d}t\\ &\leq 4\int_{-\infty}^{0}\int_{\varepsilon^\theta}^{\infty}\rho_\varepsilon(s-t)\,\mathrm{d}s\,\mathrm{d}t\\ &= 4\int_{\varepsilon^\theta}^{\infty}(t-\varepsilon^\theta)\rho_\varepsilon(t)\,\mathrm{d}t,\end{aligned}$$

thus

$$\mathcal{F}_\varepsilon^{1\mathrm{D}}(F;L)\leq C\int_{\mathbb{R}^n\setminus B_{\varepsilon^\theta}}|z|G_\varepsilon(z)\,\mathrm{d}z=CQ_{1-\theta}(\varepsilon^{-1}),$$

proving (4.13) in this case.
Case 2: $\mathrm{d}(t_F,L^{\mathrm{c}})<\varepsilon^\theta$. Either $0<t_F<\varepsilon^\theta<a$ or $0<a-\varepsilon^\theta<t_F<a$. Since both cases are similar, we treat only the case $0<a-\varepsilon^\theta<t_F<a$.

Notice that $E\triangle F\subseteq\mathbf{K}_t$ implies $t_F\in\partial E$ and

$$F^{\mathrm{c}}\cap(a-\varepsilon^\theta,a)=E^{\mathrm{c}}\cap(a-\varepsilon^\theta,a)=(a-\varepsilon^\theta,a). \tag{4.15}$$

Let us write $E\cap(0,a)=\bigcup_{j=1}^{k}I_j$, where $k\geq 1$ and $I_j\subseteq(0,a)$ are open intervals. Then let $\{s_1,\dots,s_{k_1},t_1,\dots,t_{k_2}\}$ be the elements of $\partial E\cap(0,a)$ such that

- $s_1<s_2<\ldots<s_{k_1}$ and $t_1<t_2<\ldots<t_{k_2}=t_F$;
- for each j, s_j is a left endpoint of some I_i, and t_j is a right endpoint of some I_i.

Note that s_1 may not be the left endpoint of I_1 (if $I_1=(0,t_1)$), so that k_1 may be different from k_2 (see Figure 1). The fact that $t_{k_2}=t_F$ is due to (4.15).

For $1\leq j\leq k_1$, we denote by A_j the connected component of $E^{\mathrm{c}}\cap L$ which is immediately on the left side of s_j (that is, its right endpoint is s_j), and by $\widetilde{B}_j$ the union of all the connected components of $E\cap L$ on the right side of s_j. Similarly, for $1\leq j\leq k_2$, we denote by B_j the connected component of

FIGURE 1. Two examples of the situation in Step 2.1, Case 2 in the proof of Proposition 4.2. On the left, $k_1 = 2$ and $k_2 = 3$, and on the right, $k_1 = k_2 = 3$. The thick segments represent the set $E \cap (0, a)$.

$E \cap L$ which is immediately on the left side of t_j, and by $\widetilde{A}_j$ the union of all the connected components of $E^c \cap L$ on the right side of t_j. See Figure 2.

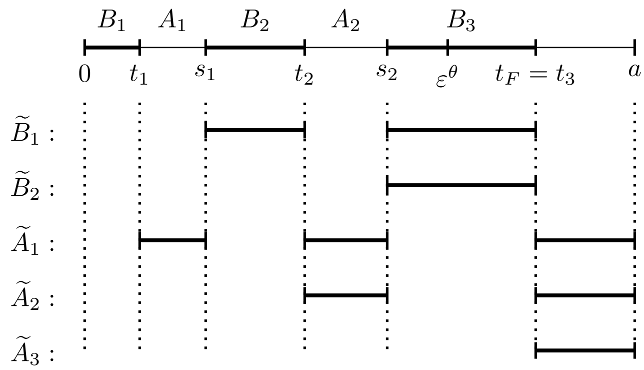

FIGURE 2. An example of the situation in Step 2.1, Case 2 in the proof of Proposition 4.2, when $k_1 = 2$ and $k_2 = 3$, with the representation of the $\widetilde{A}_j$ and the $\widetilde{B}_j$.

Then using that $\mathcal{H}^0(\partial E \cap L) = k_1 + k_2$ and decomposing the domain of integration of $P_\varepsilon^{1\mathrm{D}}(E; L)$ we see that

$$\mathcal{F}_\varepsilon^{1\mathrm{D}}(E; L) = \sum_{j=1}^{k_1} \left[1 - 2 \int_{A_j \times \widetilde{B}_j} \rho_\varepsilon(s - t) \, \mathrm{d}s \, \mathrm{d}t \right] + \sum_{j=1}^{k_2} \left[1 - 2 \int_{B_j \times \widetilde{A}_j} \rho_\varepsilon(s - t) \, \mathrm{d}s \, \mathrm{d}t \right].$$

Each term of each sum is nonnegative by (4.14), and since $\widetilde{A}_{k_2} = (t_F, a)$ by (4.15) and $B_{k_2} \subseteq (0, t_F)$ this implies in particular

$$\mathcal{F}_\varepsilon^{1\mathrm{D}}(E; L) \geq 1 - 2 \int_{(0, t_F) \times (t_F, a)} \rho_\varepsilon(s - t) \, \mathrm{d}s \, \mathrm{d}t = \mathcal{F}_\varepsilon^{1\mathrm{D}}(F; L),$$

concluding the proof of (4.13) in this case as well.

Step 2.2. For $\sigma \in \mathbb{S}^{n-1}$ we define π_σ as the projection on $\{\sigma\}^\perp$. We then set

$$\begin{aligned} \mathrm{Sh}(F; \widetilde{\mathbf{K}}_t) :=& \frac{1}{2\omega_{n-1}} \int_{\mathbb{S}^{n-1}} \mathcal{H}^{n-1}(\pi_{\sigma^\perp}(\partial^* F \cap \widetilde{\mathbf{K}}_t)) \, \mathrm{d}\mathcal{H}^{n-1}_\sigma \\ =& \frac{1}{2\omega_{n-1}} \int_{\mathbb{S}^{n-1}} \int_{\{\sigma\}^\perp} \mathbf{1}_{\{\widetilde{L}_{\sigma,x} \cap \partial^* F \neq \emptyset\}} \, \mathrm{d}\mathcal{H}^{n-1}_x \, \mathrm{d}\mathcal{H}^{n-1}_\sigma. \end{aligned}$$

Since

$$P(F; \Omega) = \frac{1}{2\omega_{n-1}} \int_{\mathbb{S}^{n-1}} \int_{\{\sigma\}^\perp} \mathcal{H}^0(\partial^* F \cap L_{\sigma,x} \cap \Omega) \, \mathrm{d}\mathcal{H}^{n-1}_x \, \mathrm{d}\mathcal{H}^{n-1}_\sigma.$$

we have

$$\begin{aligned} P(F; \widetilde{\mathbf{K}}_t) &= \frac{1}{2\omega_{n-1}} \int_{\mathbb{S}^{n-1}} \int_{\{\sigma\}^\perp} \mathcal{H}^0(\partial^* F \cap \widetilde{L}_{\sigma,x}) \, \mathrm{d}\mathcal{H}^{n-1}_x \, \mathrm{d}\mathcal{H}^{n-1}_\sigma \\ &\geq \frac{1}{2\omega_{n-1}} \int_{\mathbb{S}^{n-1}} \int_{\{\sigma\}^\perp} \mathbf{1}_{\{\widetilde{L}_{\sigma,x} \cap \partial^* F \neq \emptyset\}} \, \mathrm{d}\mathcal{H}^{n-1}_x \, \mathrm{d}\mathcal{H}^{n-1}_\sigma. \end{aligned} \tag{4.16}$$

Thus, inserting (4.12) into (4.11) and using the fact that

$$\int_{\mathbb{S}^{n-1}}\int_{\{\sigma\}^\perp}\mathbf{1}_{\{\widetilde{L}_{\sigma,x}\neq\emptyset\}}\,\mathrm{d}\mathcal{H}^{n-1}_x\,\mathrm{d}\mathcal{H}^{n-1}_\sigma\le C$$

gives

$$\mathcal{F}^{\mathrm{loc}}_\varepsilon(F)-\mathcal{F}^{\mathrm{loc}}_\varepsilon(E)\le 2(P(F;\widetilde{\mathbf{K}}_t)-\mathrm{Sh}(F;\widetilde{\mathbf{K}}_t))+CQ_{1-\theta}(\varepsilon^{-1}). \tag{4.17}$$

By Lemma 4.3 below, Sh is minimal when $F\cap\widetilde{\mathbf{K}}_t=\mathbb{R}^n_+\cap\widetilde{\mathbf{K}}_t$, which gives

$$\mathrm{Sh}(F;\widetilde{\mathbf{K}}_t)\ge\mathcal{H}^{n-1}(\widetilde{D}_t).$$

In view of (4.17), this concludes the proof of (4.10).

Step 3. We may now conclude the proof of (4.5). By (4.8) and (4.7), we find

$$(1-\gamma)P(E;\mathbf{K}_t)\le(1-\gamma)P(F;\mathbf{K}_t)+\gamma\left[2\mathscr{E}(F,t+\varepsilon^\theta)+CQ_{1-\theta}(\varepsilon^{-1})\right]+\Lambda|E\triangle F|+(1-\gamma)\eta.$$

Subtracting $(1-\gamma)\mathcal{H}^{n-1}(D_t)$ from the previous inequality and using that by (4.6)

$$\mathscr{E}(F,t+\varepsilon^\theta)-\mathscr{E}(F,t)=\mathscr{E}(E,t+\varepsilon^\theta)-\mathscr{E}(E,t)+\eta,$$

yields

$$\begin{aligned}(1-\gamma)\mathscr{E}(E,t)&\le(1-\gamma)\mathscr{E}(F,t)+2\gamma\mathscr{E}(F,t+\varepsilon^\theta)+C\gamma Q_{1-\theta}(\varepsilon^{-1})+\Lambda|E\triangle F|+(1+\gamma)\eta\\&=(1+\gamma)\mathscr{E}(F,t)+2\gamma\big[\mathscr{E}(F,t+\varepsilon^\theta)-\mathscr{E}(F,t)\big]+\Lambda|E\triangle F|+(1+\gamma)\eta\\&=(1+\gamma)\mathscr{E}(F,t)+2\gamma\big[\mathscr{E}(E,t+\varepsilon^\theta)-\mathscr{E}(E,t)\big]+\Lambda|E\triangle F|+(1+3\gamma)\eta.\end{aligned}$$

Dividing by $(1-\gamma)$ concludes the proof of (4.5). □

Let π_V denote the orthogonal projection on a vector space $V\subseteq\mathbb{R}^n$. We now prove that among sufficiently flat sets, the quantity

$$\mathrm{Sh}(E;\mathbf{K}_t)=\frac{1}{2\omega_{n-1}}\int_{\mathbb{S}^{n-1}}\mathcal{H}^{n-1}(\pi_{\sigma^\perp}(\partial^*E\cap\mathbf{K}_t))\,\mathrm{d}\mathcal{H}^{n-1}_\sigma \tag{4.18}$$

is minimal when E is flat.

Lemma 4.3. *For any $t>0$, and any set of finite perimeter $E\subseteq\mathbb{R}^n$ such that*

$$\{x_n<-1/4\}\cap\mathbf{K}_t\ \subseteq\ E\cap\mathbf{K}_t\ \subseteq\ \{x_n<1/4\}\cap\mathbf{K}_t, \tag{4.19}$$

we have

$$\mathrm{Sh}(E;\mathbf{K}_t)\ge\mathrm{Sh}(D_t\times(-1,0);\mathbf{K}_t)=\mathcal{H}^{n-1}(D_t). \tag{4.20}$$

Notice that the equality $\mathrm{Sh}(D_t\times(-1,0);\mathbf{K}_t)=\mathcal{H}^{n-1}(D_t)$ follows arguing as for (4.16).

Proof of Lemma 4.3. We start by fixing some notation. We denote by $S:\mathbb{R}^n\to\mathbb{R}^n$ the symmetry with respect to the vertical line $\{0_{\mathbb{R}^{n-1}}\}\times\mathbb{R}$, that is, for $\xi=(\xi',\xi_n)$,

$$S\xi\coloneqq(-\xi',\xi_n).$$

We write

$$\mathrm{Sh}(E;\mathbf{K}_t)=\frac{1}{4\omega_{n-1}}\int_{\mathbb{S}^{n-1}}\big[\mathcal{H}^{n-1}(\pi_{\sigma^\perp}(\partial^*E\cap\mathbf{K}_t))+\mathcal{H}^{n-1}(\pi_{(S\sigma)^\perp}(\partial^*E\cap\mathbf{K}_t))\big]\,\mathrm{d}\mathcal{H}^{n-1}_\sigma. \tag{4.21}$$

We claim that for every $\sigma\in\mathbb{S}^{n-1}$, the integrand is minimal when ∂^*E is horizontal in $\mathbf{K}_t$, that is,

$$\mathcal{H}^{n-1}(\pi_{\sigma^\perp}(\partial^*E\cap\mathbf{K}_t))+\mathcal{H}^{n-1}(\pi_{(S\sigma)^\perp}(\partial^*E\cap\mathbf{K}_t))\ge\mathcal{H}^{n-1}(\pi_{\sigma^\perp}(D_t))+\mathcal{H}^{n-1}(\pi_{(S\sigma)^\perp}(D_t)). \tag{4.22}$$

After integration this would conclude the proof of (4.20). The proof of (4.22) is done in two steps. In the first step we prove it for $n=2$ and in the second step we use slicing to reduce ourselves to the two-dimensional situation.

Step 1. We first prove (4.22) for $n=2$.

Step 1.1. By [2], we may decompose the set of finite perimeter $\widetilde{E}\coloneqq E\cap\mathbf{K}_t$ into its (measure theoretic) connected components. By assumption (4.19) one of these components denoted $\widetilde{E}_1$ contains $(-t,t)\times(-1,-\frac14)$. Its external boundary is a Jordan curve $\gamma\in C^0([0,1),[-t,t]\times[-1,\frac14])$ and we have $\gamma([0,1))\subseteq\partial^M\widetilde{E}_1\subseteq\partial E$ up to a $\mathcal{H}^1$-negligible set, where $\partial^M\widetilde{E}_1$ is the essential boundary and $\partial\widetilde{E}_1$ the usual topological boundary of $\widetilde{E}_1$. Moreover, by (4.19) we may assume up to a reparameterization that

$\gamma_{|[0,\frac{1}{2}]}$ is a parameterization of the broken line made of the three oriented segments joining $(-t,-\frac{1}{4})$ to $(-t,-1)$, then $(-t,-1)$ to $(t,-1)$ and $(t,-1)$ to $(t,-\frac{1}{4})$.

Denoting

$$s_1 := \max\{s \in [\tfrac{1}{2},1) : \gamma(s) \in \{t\}\times\mathbb{R}\},$$
$$s_2 := \min\{s \in (s_1,1) : \gamma(s) \in \{-t\}\times\mathbb{R}\},$$

we obtain the parameterization of an arc $\gamma : [s_1,s_2] \to [-t,t]\times[-\frac{1}{4},\frac{1}{4}]$ with

$$\gamma(s_1) \in \{t\}\times\mathbb{R}, \qquad \gamma(s_2) \in \{-t\}\times\mathbb{R}, \qquad \gamma((s_1,s_2)) \subseteq (-t,t)\times[-\tfrac{1}{4},\tfrac{1}{4}].$$

Let I be the segment $[\gamma(s_1),\gamma(s_2)]$. Obviously, any straight line intersecting I also intersects the arc $\gamma([s_1,s_2])$, hence for every $\sigma\in\mathbb{S}^1$, $\pi_{\sigma^\perp}(\gamma([s_1,s_2]))$ contains $\pi_{\sigma^\perp}(I)$, so that

$$\mathcal{H}^1(\pi_{\sigma^\perp}(I)) \le \mathcal{H}^1(\pi_{\sigma^\perp}(\gamma([s_1,s_2]))) \le \mathcal{H}^1(\pi_{\sigma^\perp}(\partial^M E\cap\mathbf{K}_t)) = \mathcal{H}^1(\pi_{\sigma^\perp}(\partial^* E\cap\mathbf{K}_t)). \tag{4.23}$$

Step 1.2. For $\sigma\in\mathbb{S}^1$, let $\theta\in[0,\frac{\pi}{2}]$ be such that

$$\{\mathbb{R}\sigma,\mathbb{R}(S\sigma)\} = \left\{\mathbb{R}\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}, \mathbb{R}\begin{pmatrix}-\cos\theta\\ \sin\theta\end{pmatrix}\right\}$$

and $\varphi\in(-\frac{\pi}{2},\frac{\pi}{2})$ be such that I has direction $\binom{\cos\varphi}{\sin\varphi}$. We compute

$$\begin{aligned}\mathcal{H}^1(\pi_{\sigma^\perp}(I)) + \mathcal{H}^1(\pi_{(S\sigma)^\perp}(I)) &= \frac{2t}{\cos\varphi}\left(|\sin(\theta-\varphi)| + |\sin(\theta+\varphi)|\right)\\ &= \begin{cases} 4t|\tan\varphi|\cos\theta & \text{if } 0\le\theta\le|\varphi|<\frac{\pi}{2},\\ 4t\sin\theta & \text{if } |\varphi|<\theta\le\frac{\pi}{2}.\end{cases}\end{aligned}$$

Since $\tan\varphi$ is increasing in $(\theta,\frac{\pi}{2})$ we have $|\tan\varphi|\cos\theta\ge\sin\theta$ if $\theta\le|\varphi|<\frac{\pi}{2}$ and thus

$$\mathcal{H}^1(\pi_{\sigma^\perp}(I)) + \mathcal{H}^1(\pi_{(S\sigma)^\perp}(I)) \ge 4t\sin\theta = \mathcal{H}^1(\pi_{\sigma^\perp}(I_t)) + \mathcal{H}^1(\pi_{(S\sigma)^\perp}(I_t)).$$

Together with (4.23) applied both to σ and $S\sigma$, this proves the (4.22) when $n=2$.

Step 2. The case $n>2$. There exists $e\in\mathbb{S}^{n-1}\cap[\mathbb{R}^{n-1}\times\{0\}]\sim\mathbb{S}^{n-2}$ such that

$$\sigma = (\sigma\cdot e)e + \sigma_n e_n, \qquad S\sigma = -(\sigma\cdot e)e + \sigma_n e_n.$$

Denoting $P := \operatorname{Span}\{e,e_n\}$, $V := P^\perp$ and $P_y := y+P$ for $y\in V$, we have

$$\mathcal{H}^{n-1}(\pi_{\sigma^\perp}(\partial^* E\cap\mathbf{K}_t)) = \int_{V\cap B_t}\mathcal{H}^1(\pi_{\sigma^\perp}(\partial^* E\cap\mathbf{K}_t\cap P_y))\,d\mathcal{H}^{n-2}_y, \tag{4.24}$$

$$\mathcal{H}^{n-1}(\pi_{(S\sigma)^\perp}(\partial^* E\cap\mathbf{K}_t)) = \int_{V\cap B_t}\mathcal{H}^1(\pi_{(S\sigma)^\perp}(\partial^* E\cap\mathbf{K}_t\cap P_y))\,d\mathcal{H}^{n-2}_y. \tag{4.25}$$

Next, for almost every $y\in V$, $E\cap P_y$ is a set with finite perimeter in the plane P_y and up to a $\mathcal{H}^1$-negligible set, $\partial^*_{P_y}(E\cap P_y) = (\partial^*_{\mathbb{R}^n}E)\cap P_y$.

Noticing that for $|y|\ge t$, $\mathbf{K}_t\cap P_y=\emptyset$ and that for $|y|<t$, $\mathbf{K}_t\cap P_y = y+\{x_1e+x_2e_n \ : \ |x_1|<\sqrt{t^2-|y|^2},\ |x_2|<1\}\sim(-\sqrt{t^2-|y|^2},\sqrt{t^2-|y|^2})\times(-1,1)$ and using *Step 1* in $\mathbf{K}_t\cap P_y$ concludes the proof. □

Remark 4.4. In Proposition 4.2, we introduced a parameter $\theta\in(0,1]$ to find a proper balance between the terms

$$\left[\mathscr{E}(E,t+\varepsilon^\theta) - \mathscr{E}(E,t)\right] \quad\text{and}\quad Q_{1-\theta}\left(\frac{1}{\varepsilon}\right).$$

As we will see, through an averaging argument, we can roughly estimate

$$\left[\mathscr{E}(E,t+\varepsilon^\theta) - \mathscr{E}(E,t)\right] \lesssim \varepsilon^\theta\mathbf{e}(E,2t) \tag{4.26}$$

the first quantity gets smaller the closer θ is to 1. However, $Q_{1-\theta}(\varepsilon^{-1})$ gets larger as θ goes to 1. In particular when $\theta=1$, $Q_{1-\theta}(r/\varepsilon)=Q(1)$ is a constant (non-zero unless G is compactly supported in B_1), which would prevent us to obtain a decay of the excess through iteration. We can choose later θ small enough so that $Q_{1-\theta}(r/\varepsilon)$ stays sufficiently small down to any scale $\varepsilon^{1-\beta}$ with $\beta\in(0,1)$. As long as θ is non-zero, (4.26) will be sufficient to proceed with the iteration.

4.2. **A Caccioppoli-type inequality.** From the improved quasi-minimality condition given by Proposition 4.2, we first obtain an intermediate weaker form of a Caccioppoli inequality. We refer to (4.1), (4.2) and (4.4) for the definitions of $\mathscr{F}$, $\mathscr{E}$ and $Q_{1-\theta}$.

Proposition 4.5. *Assume that G satisfies (H1) and (H2), and let $\varepsilon \in (0,1)$, $\gamma \in (0,1)$, $\theta \in (0,1]$ and $\Lambda > 0$ such that $\varepsilon^\theta \in (0, \frac{1}{4})$ and $4\Lambda \le 1-\gamma$. Then for every $(\Lambda, 4)$-minimizer E of $\mathcal{F}_{\varepsilon,\gamma}$ with*

$$\left\{x_n < -\frac{1}{8}\right\} \cap \mathbf{K}_3 \subseteq E \cap \mathbf{K}_3 \subseteq \left\{x_n < \frac{1}{8}\right\} \cap \mathbf{K}_3,$$

the following holds. For every $c \in \mathbb{R}$ such that $|c| < \frac{1}{4}$ and every $t \in (4\varepsilon^\theta, 1)$, we have

$$\mathscr{E}(E, t/2) \le C\left(\big(\mathscr{E}(E,t)\mathscr{F}(E,t,c)\big)^{\frac{1}{2}} + \left(\frac{\varepsilon^\theta}{t}\right)\mathscr{E}(E,t) + Q_{1-\theta}(\varepsilon^{-1}) + \Lambda t^{n-1}\right), \tag{4.27}$$

where C depends only on n and γ.

Proof. For almost every $t \in (4\varepsilon^\theta, 1)$, we have

$$\mathcal{H}^{n-1}(\partial \mathbf{K}_t \cap \partial E) = 0. \tag{4.28}$$

Let us fix such a t. If $\mathscr{F}(E,t,c) \ge \frac{1}{16}\mathscr{E}(E,t)$, then using the fact that $\mathscr{E}(E,\cdot)$ is nondecreasing, we have

$$\mathscr{E}(E,t/2) \le \mathscr{E}(E,t) \le \sqrt{\mathscr{E}(E,t)}\sqrt{\mathscr{E}(E,t)} \le 4\big(\mathscr{E}(E,t)\mathscr{F}(E,t,c)\big)^{\frac{1}{2}}$$

thus (4.27) holds. Hence, we now assume $\mathscr{F}(E,t,c) < \frac{1}{16}\mathscr{E}(E,t)$, and set $\lambda := \sqrt{\frac{\mathscr{F}(E,t,c)}{\mathscr{E}(E,t)}} \in (0, \frac{1}{4})$. As in [26, Lemma 24.9] we want to use for $s \in (0,1)$ the construction of [26, Lemma 24.6] as competitor inside $\mathbf{K}_{st}$ for Proposition 4.2. To this aim using for instance [26, Theorem 13.8], we approximate E by smooth sets E_k with $|E \triangle E_k| \to 0$, $P(E_k) \to P(E)$ and

$$\left\{x_n < -\frac{1}{4}\right\} \cap \mathbf{K}_3 \subseteq E_k \cap \mathbf{K}_3 \subseteq \left\{x_n < \frac{1}{4}\right\} \cap \mathbf{K}_3. \tag{4.29}$$

By the Morse–Sard lemma, for almost every $s \in (0,1)$,

$$\partial \mathbf{K}_{st} \cap \partial E_k \text{ is a } (n-2)\text{-dimensional hypersurface.} \tag{4.30}$$

For every such s we may apply [26, Lemma 24.6] with $a = (1-\lambda)st$ and $b = st$, and use the inequalities $\sqrt{1+t^2} \le 1+t^2$ and $1-(1-\lambda)^{n-1} \le (n-1)\lambda$, to construct an open set of finite perimeter F_s such that (4.29) holds for F_s,

$$F_s \cap \partial \mathbf{K}_{st} = E_k \cap \partial \mathbf{K}_{st}, \tag{4.31}$$

and

$$\mathscr{E}(F_s, st) \le C\left(\lambda st V_{\mathscr{E}}(st) + \frac{1}{\lambda st} V_{\mathscr{F}}(st)\right), \tag{4.32}$$

where we have set

$$V_{\mathscr{E}}(a) := \frac{\mathrm{d}}{\mathrm{d}a}(\mathscr{E}(E_k, a)) = \mathcal{H}^{n-2}(\partial \mathbf{K}_a \cap \partial E_k) - \mathcal{H}^{n-2}(\partial D_a)$$

and

$$V_{\mathscr{F}}(a) := \frac{\mathrm{d}}{\mathrm{d}a}(a^2 \mathscr{F}(E_k, a, c)) = \int_{\partial \mathbf{K}_a \cap \partial E_k} (x_n - c)^2 \,\mathrm{d}\mathcal{H}^{n-2}.$$

Applying Proposition 4.2 with, $F = (F_s \cap \mathbf{K}_{st}) \cup (E \backslash \mathbf{K}_{st})$ and noticing that by (4.31) and [26, Theorem 16.16], for a.e. s, $\mathcal{H}^{n-1}(\partial^* F_s \cap \partial \mathbf{K}_{st}) = \mathcal{H}^{n-1}((E \triangle E_k) \cap \partial \mathbf{K}_{st})$ we find for such s,

$$\begin{aligned}
\mathscr{E}(E, st) &\le C\Big(\mathscr{E}(F_s, st) + \big[\mathscr{E}(E, st+\varepsilon^\theta) - \mathscr{E}(E, st)\big] + \Lambda |\mathbf{K}_{st}| + Q_{1-\theta}(\varepsilon^{-1}) \\
&\qquad + \mathcal{H}^{n-1}((E \triangle E_k) \cap \partial \mathbf{K}_{st})\Big) \\
&\overset{(4.32)}{\le} C\left(\lambda st V_{\mathscr{E}}(st) + \frac{1}{\lambda st} V_{\mathscr{F}}(st) + \big[\mathscr{E}(E, st+\varepsilon^\theta) - \mathscr{E}(E, st)\big] + \Lambda t^{n-1} + Q_{1-\theta}(\varepsilon^{-1})\right. \\
&\qquad \left. + \mathcal{H}^{n-1}((E \triangle E_k) \cap \partial \mathbf{K}_{st})\right).
\end{aligned} \tag{4.33}$$

We now integrate (4.33) for s between $1/2$ and $3/4$. First, since $\mathscr{E}(E,\cdot)$ is nondecreasing, we have

$$\frac{1}{4}\mathscr{E}(E,t/2) \le \int_{\frac{1}{2}}^{\frac{3}{4}} \mathscr{E}(E,st)\,\mathrm{d}s. \tag{4.34}$$

Second, we compute

$$\begin{aligned}\int_{\frac{1}{2}}^{\frac{3}{4}} [\mathscr{E}(E,st+\varepsilon)-\mathscr{E}(E,st)]\,\mathrm{d}s &= \frac{1}{t}\int_{\frac{t}{2}}^{\frac{3t}{4}} [\mathscr{E}(E,a+\varepsilon^\theta)-\mathscr{E}(E,a)]\,\mathrm{d}a \\ &= \frac{1}{t}\left(\int_{\frac{t}{2}+\varepsilon^\theta}^{\frac{3t}{4}+\varepsilon^\theta} \mathscr{E}(E,a)\,\mathrm{d}a - \int_{\frac{t}{2}}^{\frac{3t}{4}} \mathscr{E}(E,a)\,\mathrm{d}a\right) \le \frac{1}{t}\int_{\frac{3t}{4}}^{\frac{3t}{4}+\varepsilon^\theta} \mathscr{E}(E,a)\,\mathrm{d}a \le \left(\frac{\varepsilon^\theta}{t}\right)\mathscr{E}(E,t),\end{aligned} \tag{4.35}$$

where we used the fact that $\mathscr{E}(E,\cdot)$ is nondecreasing for the last inequality. Third,

$$\int_{\frac{1}{2}}^{\frac{3}{4}} stV_{\mathscr{E}}(st)\,\mathrm{d}s \le \frac{3}{4}\int_{\frac{t}{2}}^{\frac{3t}{4}} V_{\mathscr{E}}(a)\,\mathrm{d}a = \frac{3}{4}\left(\mathscr{E}\left(E_k,\frac{3t}{4}\right)-\mathscr{E}\left(E_k,\frac{t}{2}\right)\right) \le \frac{3}{4}\mathscr{E}(E_k,t). \tag{4.36}$$

Finally, with a similar argument using that $a \mapsto a^2\mathscr{F}(E_k,a,c)$ is nondecreasing, we have

$$\int_{\frac{1}{2}}^{\frac{3}{4}} \frac{1}{st}V_{\mathscr{F}}(st)\,\mathrm{d}s \le 2\mathscr{F}(E_k,t,c). \tag{4.37}$$

Inserting (4.34), (4.35), (4.36) and (4.37) into (4.33) yields

$$\mathscr{E}(E,t/2) \le C\left[\lambda\mathscr{E}(E_k,t)+\frac{1}{\lambda}\mathscr{F}(E_k,t,c)+\left(\frac{\varepsilon^\theta}{t}\right)\mathscr{E}(E,t)+\Lambda t^{n-1}+Q_{1-\theta}(\varepsilon^{-1})+|E\triangle E_k|\right]. \tag{4.38}$$

By (4.28) we can send $k\to\infty$ to obtain

$$\mathscr{E}(E,t/2) \le C\left[\lambda\mathscr{E}(E,t)+\frac{1}{\lambda}\mathscr{F}(E,t,c)+\left(\frac{\varepsilon^\theta}{t}\right)\mathscr{E}(E,t)+\Lambda t^{n-1}+Q_{1-\theta}(\varepsilon^{-1})\right].$$

Recalling that $\lambda = \sqrt{\frac{\mathscr{F}(E,t,c)}{\mathscr{E}(E,t)}}$ concludes the proof of (4.27) for a.e. $t\in(4\varepsilon^\theta,1)$. By the left-continuity of $\mathscr{E}(E,\cdot)$ and $\mathscr{F}(E,\cdot,c)$ this actually holds for every $t\in(4\varepsilon^\theta,1)$. □

We now post-process (4.27) to obtain the desired stronger Caccioppoli inequality. The main difference with [26, Theorem 24.1] is that in our case we cannot apply (4.27) at scales which are smaller than ε^θ.

Proposition 4.6 (Caccioppoli inequality). *Assume that G satisfies (H1) and (H2), and let $\varepsilon\in(0,1)$, $\gamma\in(0,1)$, $\Lambda>0$ and $r_0>0$ with $\Lambda r_0\le 1-\gamma$. There exist constants $\tau_{\mathrm{cac}}=\tau_{\mathrm{cac}}(n)>0$ and $M_{\mathrm{cac}}>1$ such that the following holds. Let E be a (Λ,r_0)-minimizer of $\mathcal{F}_{\varepsilon,\gamma}$ and assume that $0\in\partial E$. If for some $\nu\in\mathbb{S}^{n-1}$, $r_0>M_{\mathrm{cac}}r>0$ and $\theta\in(0,1]$,*

$$\mathbf{e}(E,M_{\mathrm{cac}}r,\nu)+\left(\frac{\varepsilon}{r}\right)^\theta \le \tau_{\mathrm{cac}},$$

then

$$\mathbf{e}(E,r/2,\nu) \le C\left(\mathbf{f}(E,r,\nu)+\left(\frac{\varepsilon}{r}\right)^\theta\mathbf{e}(E,r,\nu)+\Lambda r+Q_{1-\theta}\left(\frac{r}{\varepsilon}\right)\right), \tag{4.39}$$

where $C=C(n,G,\gamma)$.

Proof. Up to a rotation and rescaling, we may assume that $\nu=e_n$ and $r=1$. Therefore, up to choosing M_{cac} large enough and τ_{cac} small enough, E is a $(\Lambda,4)$-minimizer of $\mathcal{F}_{\varepsilon,\gamma}$ with $4\Lambda\le 1-\gamma$ and $16\varepsilon^\theta<1$. Thus, (4.39) amounts to proving

$$\mathbf{e}_n(E,\tfrac{1}{2}) \le C\left(\mathbf{f}_n(E,1)+\varepsilon^\theta\mathbf{e}_n(E,1)+\Lambda+Q_{1-\theta}\left(\frac{1}{\varepsilon}\right)\right). \tag{4.40}$$

By Proposition 2.14, choosing M_{cac} even larger if necessary and $\tau_{\mathrm{cac}}=\tau_{\mathrm{cac}}(n)$ small enough, we have

$$\left\{(x',x_n)\in\mathbf{C}_4 \;:\; x_n<-\frac{1}{8}\right\} \subseteq E\cap\mathbf{C}_4 \subseteq \left\{(x',x_n)\in\mathbf{C}_4 \;:\; x_n<\frac{1}{8}\right\}. \tag{4.41}$$

Thus for every $z\in D_1$, we have

$$\left\{(x',x_n)\in\mathbf{K}_3(z) \;:\; x_n<-\frac{1}{8}\right\} \subseteq E\cap\mathbf{K}_3(z) \subseteq \left\{(x',x_n)\in\mathbf{K}_3(z) \;:\; x_n<\frac{1}{8}\right\},$$

so we can apply Proposition 4.5 to $E+(z,0)$ with $2s$ for every $s\in(2\varepsilon^\theta,\frac12)$. For every $s\in(2\varepsilon^\theta,\frac12)$ such that $D_{2s}(z)\subseteq D_1$, we get that

$$\mathscr{E}(E,\mathbf{K}_s(z))\le C\left(\left(\mathscr{E}(E,\mathbf{K}_{2s}(z))\mathscr{F}(E,\mathbf{K}_{2s}(z),c)\right)^{\frac12}+\left(\frac{\varepsilon}{s}\right)^{\theta}\mathscr{E}(E,\mathbf{K}_{2s}(z))\;+Q_{1-\theta}\left(\frac{1}{\varepsilon}\right)+\Lambda s^{n-1}\right). \tag{4.42}$$

Setting

$$h:=\inf_{|c|<\frac14}\int_{\mathbf{C}_1\cap\partial^*E}(x_n-c)^2\,\mathrm{d}\mathcal{H}^{n-1}$$

multiplying (4.42) by s^2 and taking the infimum over $|c|<\frac14$, using the fact that

$$s^2\mathscr{F}(E,\mathbf{K}_{2s}(z),c)\le\frac{\mathscr{F}(E,\mathbf{K}_1,c)}{4}\le\frac{h}{4}$$

for every $s\in(2\varepsilon^\theta,\frac12)$ with $D_{2s}(z)\subseteq D_1$, we find

$$s^2\mathscr{E}(E,\mathbf{K}_s(z))\le C\left(\left(s^2\mathscr{E}(E,\mathbf{K}_{2s}(z))h\right)^{\frac12}+\varepsilon^\theta\mathscr{E}(E,\mathbf{K}_{2s}(z))+s^2Q_{1-\theta}\left(\frac{1}{\varepsilon}\right)+\Lambda\right). \tag{4.43}$$

Set

$$\Psi:=\sup\left\{s^2\mathscr{E}(E,\mathbf{K}_s(z))\ :\ D_{2s}(z)\subseteq D_1\text{ and }s\in\left(4\varepsilon^\theta,\frac12\right)\right\}.$$

If

$$\Psi=\sup\left\{s^2\mathscr{E}(E,\mathbf{K}_s(z))\ :\ D_{2s}(z)\subseteq D_1\text{ and }s\in\left(4\varepsilon^\theta,8\varepsilon^\theta\right)\right\}$$

then (4.40) holds. Indeed, using the left-continuity of $t\mapsto\mathscr{E}(E,\mathbf{K}_t)$ and the fact that $\mathscr{E}(E,\mathbf{K}_s(z))\le\mathscr{E}(E,\mathbf{K}_1)$ whenever $D_{2s}(z)\subseteq D_1$, in that case we find

$$\mathscr{E}(E,\mathbf{K}_{\frac12})\le\frac{\Psi}{4}\le C\varepsilon^{2\theta}\mathscr{E}(E,\mathbf{K}_1),$$

which gives (4.40) recalling that $\mathbf{e}_n(E,\frac12)=2\mathscr{E}(E,\mathbf{K}_{\frac12})$ (see (4.3)). We can thus take the supremum over $s>4\varepsilon^\theta$ or $s>8\varepsilon^\theta$ for Ψ. For any z and s such that $D_{2s}(z)\subseteq D_1$ and $s\in(8\varepsilon^\theta,\frac12)$, we cover $D_s(z)$ by $N=N(n)$ balls $D_{\frac s4}(z_k)$ with centers $z_k\in D_s(z)$. Then since $\frac s4>2\varepsilon^\theta$ and $D_{\frac s2}(z_k)\subseteq D_1$, we can apply (4.43) to each $\left(\frac s4\right)^2\mathscr{E}(E,\mathbf{K}_{\frac s4}(z_k))$. Thus, by the subadditivity of $\mathscr{E}$, and by definition of Ψ, for such z and $s\in(8\varepsilon^\theta,\frac12)$, we deduce

$$\begin{aligned}s^2\mathscr{E}(E,\mathbf{K}_s(z))&\le\frac{1}{16}\sum_{k=1}^N\left(\frac s4\right)^2\mathscr{E}(E,\mathbf{K}_{\frac s4}(z_k))\\&\le C\sum_{k=1}^N\left(\left(s^2\mathscr{E}(E,\mathbf{K}_{\frac s2}(z_k))h\right)^{\frac12}+\varepsilon^\theta\mathscr{E}(E,\mathbf{K}_{\frac s2}(z_k))+s^2Q_{1-\theta}\left(\frac1\varepsilon\right)+\Lambda\right)\\&\le C\left(\sqrt{\Psi h}+\varepsilon^\theta\Psi+Q_{1-\theta}\left(\frac1\varepsilon\right)+\Lambda\right).\end{aligned} \tag{4.44}$$

Recall that Ψ is in fact obtained by taking the supremum over the s, z such that $D_{2s}(z)\subseteq D_1$ and $s\in(8\varepsilon^\theta,\frac12)$ by the above discussion. Therefore, (4.44) yields

$$\Psi\le C\left(\sqrt{\Psi h}+\varepsilon^\theta\Psi+Q_{1-\theta}\left(\frac1\varepsilon\right)+\Lambda\right). \tag{4.45}$$

If

$$\varepsilon^\theta\Psi+Q_{1-\theta}\left(\frac1\varepsilon\right)+\Lambda<\sqrt{\Psi h},$$

then (4.45) implies $\Psi\le Ch$. Otherwise, (4.45) implies

$$\Psi\le C\left(\varepsilon^\theta\Psi+Q_{1-\theta}\left(\frac1\varepsilon\right)+\Lambda\right)\le C\left(\varepsilon^\theta\mathscr{E}(E,\mathbf{K}_1)+Q_{1-\theta}\left(\frac1\varepsilon\right)+\Lambda\right).$$

Recalling $\mathbf{e}_n(E,\frac12)=2\mathscr{E}(E,\mathbf{K}_{\frac12})$, the left-continuity of $t\mapsto\mathscr{E}(E,\mathbf{K}_t)$ and the definition of h, combining the different cases yields (4.40). This concludes the proof. ☐

5. Uniform regularity

5.1. **Excess decay for $r \lesssim \varepsilon$.** If G satisfies assumptions (H1), (H2) and (H4), by Proposition 2.10 and Remark 2.11, it is standard to obtain power decay of the excess at small scales. Let us recall the following well-known result.

Proposition 5.1. *Let $\omega > 0$, $\alpha \in (0,1)$ and $r_0 > 0$ be fixed. Then, there exists $\tau = \tau(n,\alpha) > 0$ and $C = C(n,\alpha)$ such that the following holds. If E is such that for every $r \le r_0$ and every $E \triangle F \subset B_r$,*

$$P(E; B_r) \le P(F; B_r) + \omega r^{n-1+\alpha},$$

assuming that $0 \in \partial E$ and

$$\mathbf{e}(E,R) + \omega R^\alpha \le \tau$$

for some $R \le r_0$, then

$$\mathbf{e}(E,r) \le C\left(R^{-\alpha}\mathbf{e}(E,R) + \omega\right) r^\alpha \qquad \forall r \in (0,R).$$

Proof. Although this result is standard and can be reconstructed from e.g. [31], we provide a short proof for the convenience of the reader since this precise statement is not easily accessible in the literature. By scaling we may assume without loss of generality that $R = 1$ (replacing ω by ωR^α). We thus want to prove that there exist $\tau = \tau(n) > 0$ and $C = C(n,\alpha) > 0$ such that provided

$$\mathbf{e}(E,1) + \omega \le \tau \tag{5.1}$$

then

$$\mathbf{e}(E,r) \le C\left(\mathbf{e}(E,1) + \omega\right) r^\alpha \qquad \forall r \in (0,1). \tag{5.2}$$

We recall that by the tilt Lemma (see for instance [20, Lemma 4.6] applied to $\Lambda = \omega r^\alpha$ and $\ell = 0$ combined with the beginning of the proof of [20, Proposition 4.1]), for every λ small enough, there exists $\tau_{\mathrm{tilt}} = \tau_{\mathrm{tilt}}(n,\lambda) > 0$, $C_{\mathrm{tilt}} = C_{\mathrm{tilt}}(n) > 0$ and $C_\lambda > 0$ such that provided

$$\mathbf{e}(E,r) + \omega r^\alpha \le \tau_{\mathrm{tilt}}, \tag{5.3}$$

we have

$$\mathbf{e}(E,\lambda r) \le C_{\mathrm{tilt}}\lambda^2 \mathbf{e}(E,r) + C_\lambda \omega r^\alpha. \tag{5.4}$$

We now choose λ such that $C_{\mathrm{tilt}}\lambda^2 \le \lambda^\alpha/2$ and then set $C = 2C_\lambda/\lambda^\alpha$. Letting for $k \ge 0$, $r_k = \lambda^k$, we claim that for every $k \ge 0$,

$$\mathbf{e}(E,r_k) \le C(\mathbf{e}(E,1) + \omega) r_k^\alpha. \tag{5.5}$$

By the scaling properties of the excess, see Proposition 2.12, this would conclude the proof of (5.2).

We start by noting that the statement holds for $k = 0$. Now if it holds for $k-1$, up to choosing τ small enough, (5.3) holds for $r = r_{k-1}$ by (5.1). Therefore, by (5.4) and the induction hypothesis,

$$\mathbf{e}(E,r_k) \le C_{\mathrm{tilt}}\lambda^2 \mathbf{e}(E,r_{k-1}) + C_\lambda \omega r_{k-1}^\alpha \le \frac{C}{2}(\mathbf{e}(E,1)+\omega) r_k^\alpha + \frac{C_\lambda}{\lambda^\alpha}\omega r_k^\alpha \le C(\mathbf{e}(E,1)+\omega) r_k^\alpha.$$

This proves (5.5). □

As a consequence, we have the following power decay of the excess for small scales.

Proposition 5.2 (Excess decay at small scales)**.** *Assume that G satisfies (H1), (H2) and (H4), and let $\gamma \in (0,1)$, $\varepsilon > 0$, $\Lambda > 0$ and $r_0 > 0$ with $\Lambda(r_0 + \varepsilon) \le 1 - \gamma$. There exist positive constants $\tau_{\mathrm{dec}}^{\mathbf{s}} = \tau_{\mathrm{dec}}^{\mathbf{s}}(n,G,\gamma)$ and $C = C(n,G,\gamma)$ such that the following holds. If E is a (Λ, r_0)-minimizer of $\mathcal{F}_{\varepsilon,\gamma}$ with $0 \in \partial E$ and such that for some $R \le r_0$,*

$$\mathbf{e}(E,R) + \frac{R}{\varepsilon} \le \tau_{\mathrm{dec}}^{\mathbf{s}}, \tag{5.6}$$

then we have (recall that s_0 is given by (H4))

$$\mathbf{e}(E,r) \le C\left(\frac{r}{R}\right)^{1-s_0}\left(\mathbf{e}(E,R) + \left(\frac{R}{\varepsilon}\right)^{1-s_0}\right) \qquad \forall r \in (0,R). \tag{5.7}$$

Proof. By scaling (recall Proposition 2.7) we can assume that $\varepsilon = 1$. By Proposition 2.10, we have for $r \le R$,

$$P(E; B_r) \le P(F; B_r) + C\left(r^{n-s_0} + \Lambda r^n\right) \qquad \forall E\triangle F \subset B_r.$$

Since on the one hand, $\Lambda \le 1$ and on the other hand, up to choosing $\tau^{\mathbf{s}}_{\text{dec}}$ small enough, $R \le 1$ by (5.6), this reduces to

$$P(E; B_r) \le P(F; B_r) + Cr^{n-s_0} \qquad \forall E\triangle F \subset B_r.$$

We can then apply Proposition 5.1 with $\alpha = 1 - s_0$ to conclude the proof. □

5.2. **Excess decay for $r \gg \varepsilon$.** Starting with a small excess at a given scale much larger than ε, we show that the excess is smaller at a smaller scale, up to tilting the direction.

Lemma 5.3 (Tilt lemma)**.** *Assume that G satisfies (H1), (H2) and (H3), and let $\gamma \in (0,1)$, $\varepsilon > 0$, $\Lambda > 0$ and $r_0 > 0$ with $\Lambda r_0 \le 1-\gamma$. Then, there exists a positive constant λ_{tilt} such that for every $\lambda \in (0, \lambda_{\text{tilt}})$, there exists $\tau_{\text{tilt}} = \tau_{\text{tilt}}(n, G, \gamma, \lambda) > 0$ such that the following holds. If E is a (Λ, r_0)-minimizer of $\mathcal{F}_{\varepsilon,\gamma}$ with $0 \in \partial E$ which satisfies, for some $0 < r \le r_0$ and $\theta \in (0, 1]$,*

$$\mathbf{e}_n(E, r) + \Lambda r + \left(\frac{\varepsilon}{r}\right)^{\theta} \le \tau_{\text{tilt}}, \tag{5.8}$$

then there exists $\nu \in \mathbb{S}^{n-1}$ such that (recall the definition (4.4) of $Q_{1-\theta}$)

$$\mathbf{e}(E, \lambda r, \nu) \le C\left(\lambda^2 \mathbf{e}_n(E, r) + \lambda\Lambda r + Q_{1-\theta}\left(\frac{\lambda r}{\varepsilon}\right)\right), \tag{5.9}$$

where $C = C(n, G, \gamma)$.

Proof. We follow relatively closely the proof of [26, Theorem 25.3]. Let $\lambda \in (0, \lambda_{\text{tilt}})$, with λ_{tilt} and τ_{tilt} to be chosen later. Up to rescaling, we may assume that $r = 4$, $\mathbf{e}_n(E, 4) + 4\Lambda + \varepsilon^{\theta} \le \tau_{\text{tilt}}$, and E is a $(\Lambda, 4)$-minimizer of $\mathcal{F}_{\varepsilon,\gamma}$ with $4\Lambda \le 1-\gamma$. In the rest of the proof, we shall write $\mathbf{e}_n(r)$ for $\mathbf{e}_n(E, r)$ and $\mathbf{f}(r, \nu)$ for $\mathbf{f}(E, r, \nu)$.

Assuming that $\tau_{\text{tilt}} \le \tau_{\text{lip}}$, we can apply Theorem 3.1 with $r = 1$. Let $C_1 = C_1(n, G, \gamma)$ be a large constant, and set

$$L := C_1\left(\mathbf{e}_n(4) + Q\left(\frac{1}{\varepsilon}\right) + \Lambda\right).$$

We proceed in two steps.

Step 1. We claim that if

$$L \le \min(\lambda^{(n-1)(n+3)}, \sigma^2), \tag{5.10}$$

where $\sigma(n, G, \gamma, \lambda)$ is the constant given by Proposition 3.3 with $\tau = \lambda^{n+3}$, then there exists $\nu \in \mathbb{S}^{n-1}$ such that

$$\mathbf{f}(\lambda, \nu) \le C\lambda^2 L, \tag{5.11}$$

where $C = C(n, G, \gamma)$. Let us assume that (5.10) holds, and let us set $u_0 := \frac{u}{\sqrt{L}}$. By Theorem 3.1, u_0 satisfies

$$\int_{D_2} |\nabla u_0|^2 \le 1$$

and, choosing C_1 large enough, for all $\varphi \in C^\infty_c(D_1)$,

$$\int_{D_1} \nabla u_0 \cdot \nabla\varphi - \gamma\int_{D_2\times D_2} (u_0(x') - u_0(y'))(\varphi(x') - \varphi(y'))G_\varepsilon(x' - y', 0)\,\mathrm{d}x'\,\mathrm{d}y' \le \sqrt{L}\|\nabla\varphi\|_{L^\infty}.$$

Assuming $\tau_{\text{tilt}} \le \varepsilon_{\text{harm}}$, then since $\sqrt{L} \le \sigma$ by assumption, Proposition 3.3 gives the existence of a harmonic function $v_0 \in H^1(D_1)$ such that

$$\int_{D_1} |\nabla v_0|^2 \le 1 \quad \text{and} \quad \int_{D_1} |u_0 - v_0|^2 \le \lambda^{n+3}.$$

Setting $v := \sqrt{L}v_0$, v is a harmonic function in D_1 such that

$$\int_{D_1} |\nabla v|^2 \le L \quad \text{and} \quad \int_{D_1} |u - v|^2 \le \lambda^{n+3} L. \tag{5.12}$$

Consider $w(z) := v(0) + \nabla v(0)\cdot z$ the tangent map of v at the origin. Then since v is harmonic, up to choosing λ_{tilt} small enough we have

$$\|v - w\|^2_{L^\infty(D_\lambda)} \le C\lambda^4\|\nabla v\|^2_{L^2(D_1)} \le C\lambda^4 L,$$

thus with (5.12), this implies

$$\frac{1}{\lambda^{n+1}}\int_{D_\lambda}|u-w|^2\leq C\lambda^2 L. \tag{5.13}$$

Defining the new direction

$$\nu:=\frac{(-\nabla v(0),1)}{\sqrt{1+|\nabla v(0)|^2}},$$

using (5.12) and (5.13) and the consequences of Theorem 3.1, proceeding exactly as in Step 1 of the proof of [26, Theorem 25.3, pp 343], we obtain the claim (5.11).

Step 2. For λ fixed, we can assume that $\tau_{\rm tilt}$ is chosen small enough depending on n, G, γ and λ to enforce (5.10). Then, a key observation is that with that choice of ν, we have

$$|\nu-e_n|^2\leq C\int_{D_1}|\nabla v|^2\leq CL.$$

Thus, since $\mathbf{C}(0,r,\nu)\subseteq\mathbf{C}(0,\sqrt{2}r,e_n)$, by Propositions 2.12 and 2.13, if $\lambda_{\rm tilt}$ is small enough so that $M_{\rm cac}\sqrt{2}\lambda<4$,

$$\mathbf{e}(M_{\rm cac}\lambda,\nu)\leq C\left(\mathbf{e}_n(M_{\rm cac}\sqrt{2}\lambda)+|\nu-e_n|^2\right)\leq C\left(\frac{1}{\lambda^{n-1}}\mathbf{e}_n(4)+L\right)\leq C\frac{L}{\lambda^{n-1}}. \tag{5.14}$$

Whence, by (5.10), up to choosing $\lambda_{\rm tilt}$ even smaller if necessary

$$\mathbf{e}(M_{\rm cac}\lambda,\nu)\leq C\lambda^{(n-1)(n+2)}\leq\tau_{\rm cac}.$$

As a consequence, we can apply Proposition 4.6, which yields (recall that $Q(1/\varepsilon)\leq Q_{1-\theta}(\lambda/\varepsilon)$)

$$\mathbf{e}(\lambda/2,\nu)\leq C_2\left(\mathbf{f}(\lambda,\nu)+\left(\frac{\varepsilon}{\lambda}\right)^\theta\mathbf{e}(\lambda,\nu)+\lambda\Lambda+Q_{1-\theta}\left(\frac{\lambda}{\varepsilon}\right)\right), \tag{5.15}$$

where $C_2=C_2(n,G,\gamma)$. Since $\varepsilon^\theta\leq\tau_{\rm tilt}$, up to choosing $\tau_{\rm tilt}$ even smaller if necessary depending on λ, we have

$$\left(\frac{\varepsilon}{\lambda}\right)^\theta\mathbf{e}(\lambda,\nu)\leq C\left(\frac{\varepsilon}{\lambda}\right)^\theta\mathbf{e}(M_{\rm cac}\lambda,\nu)\overset{(5.14)}{\leq}\frac{L\tau_{\rm tilt}}{\lambda^{n-1+\theta}}\leq\lambda^2L.$$

Thus, for $\lambda_{\rm tilt}$ small enough, (5.11) and (5.15) give

$$\mathbf{e}(\lambda/2,\nu)\leq C\left(\lambda^2L+\lambda\Lambda+Q_{1-\theta}\left(\frac{\lambda}{\varepsilon}\right)\right).$$

Since Q is nonincreasing, this gives (5.9) with $r=4$ and $\lambda/2$ in place of λ, which concludes the proof. □

As a corollary, iterating properly Lemma 5.3, we get the following power decay of the excess down to scales which are large compared to ε (the constant η below is typically large).

Proposition 5.4. *Assume that G satisfies (H1), (H2) and (H3). Let $\gamma\in(0,1)$, $\varepsilon>0$ $\Lambda>0$, $r_0>0$ and $\eta\geq1$ with $\Lambda r_0\leq1-\gamma$. Given any $\theta\in(0,1)$, there exists a positive constant $\tau^{\boldsymbol{\ell}}_{\rm dec}=\tau^{\boldsymbol{\ell}}_{\rm dec}(n,G,\gamma,\theta)$ such that the following holds. If E is a (Λ,r_0)-minimizer of $\mathcal{F}_{\varepsilon,\gamma}$ with $0\in\partial E$ satisfying, for some $\eta\varepsilon\leq R\leq r_0$,*

$$\mathbf{e}(E,R)+\Lambda R+\eta^{-1}\leq\tau^{\boldsymbol{\ell}}_{\rm dec} \tag{5.16}$$

then, for all $r\in[\eta\varepsilon,R]$, we have

$$\mathbf{e}(E,r)\leq C\left[\frac{r}{R}\big(\mathbf{e}(E,R)+\Lambda R\big)+Q_{1-\theta}\left(\frac{r}{\lambda\varepsilon}\right)\right] \tag{5.17}$$

where λ and C depend only on n,G,γ.

Proof. With Lemma 5.3 at hand, the proof of (5.17) is very similar to the proof of Proposition 5.2. By scaling (recall Proposition 2.7) we may assume that $R=1$. Arguing as in [20, Proposition 4.1] we may use the scaling properties of the excess (see Proposition 2.12) to post-process Lemma 5.3 and replace the cylindrical excess by the spherical excess both in the hypothesis (5.8) and the conclusion (5.9).

Let $\lambda=\lambda(n,G,\gamma)\leq1$ to be chosen later and set for $k\geq0$, $r_k=\lambda^k$. By the scaling properties of the excess (see Proposition 2.12) and the monotonicity of $Q_{1-\theta}$, in order to prove (5.17), it is enough to prove that there exists $C>0$ such that if $r_k\geq\eta\varepsilon$ then

$$\mathbf{e}(E,r_k)\leq C\left(r_k(\mathbf{e}(E,1)+\Lambda)+Q_{1-\theta}\left(\frac{r_k}{\varepsilon}\right)\right). \tag{5.18}$$

Let $C_{\text{tilt}} > 0$ be the constant given in (5.9). We then choose λ such that $C_{\text{tilt}}\lambda^2 \le \lambda/2$ and set $C = 2C_{\text{tilt}}$. Since (5.18) holds for $k = 0$ it is enough to show that provided it holds for $k-1$ then it also holds for k.
By (5.16), the induction hypothesis and the fact that Q vanishes at infinity, up to choosing $\tau_{\text{dec}}^{\boldsymbol{\ell}}$ small enough, (5.8) is satisfied for $r = r_{k-1}$ (notice that $(\varepsilon/r_{k-1})^\theta \le \eta^{-\theta}$). Therefore, by (5.9) and the monotonicity of $Q_{1-\theta}$,

$$\begin{aligned}\mathbf{e}(E, r_k) &\le C_{\text{tilt}}\left(\lambda^2\mathbf{e}(E, r_{k-1}) + \Lambda r_k + Q_{1-\theta}\left(\frac{r_k}{\varepsilon}\right)\right)\\ &\overset{(5.18)}{\le} \frac{C}{2}\left(r_k(\mathbf{e}(E,1)+\Lambda) + Q_{1-\theta}\left(\frac{r_k}{\varepsilon}\right)\right) + C_{\text{tilt}}\left(\Lambda r_k + Q_{1-\theta}\left(\frac{r_k}{\varepsilon}\right)\right)\\ &\qquad\qquad \le C\left(r_k(\mathbf{e}(E,1)+\Lambda) + Q_{1-\theta}\left(\frac{r_k}{\varepsilon}\right)\right).\end{aligned}$$

This concludes the proof. □

5.3. **$C^{1,\alpha}$-regularity.** Eventually, combining Propositions 5.2 and 5.4 we obtain power decay of the excess down to arbitrary small scales.

Theorem 5.5. *Assume that G satisfies (H1) to (H5), and let $\gamma \in (0,1)$, $\Lambda > 0$ and $r_0 > 0$ with $\Lambda r_0 \le 1-\gamma$. Then, for every $\alpha \in (0, \alpha_*)$ with*

$$\alpha_* = \frac{1}{2}\frac{p_0}{(n-s_0)(n+p_0)+p_0}(1-s_0),$$

there exist $\beta = \beta(n, G, \alpha)$, $\tau_{\text{reg}} = \tau_{\text{reg}}(n, G, \gamma, \alpha)$ and $\varepsilon_{\text{reg}} = \varepsilon_{\text{reg}}(n, G, \gamma, \alpha, \Lambda)$ such that the following holds. If E is a (Λ, r_0)-minimizer of $\mathcal{F}_{\varepsilon,\gamma}$ with $\varepsilon \in (0, \varepsilon_{\text{reg}})$ and $0 \in \partial E$ satisfying, for some $\varepsilon^{1-\beta} \le R \le r_0$,

$$\mathbf{e}(E, R) + \Lambda R \le \tau_{\text{dec}}$$

then

$$\mathbf{e}(E, r) \le C\left(\frac{r}{R}(\mathbf{e}(E,R) + \Lambda R) + r^{2\alpha}\right) \qquad \text{for } 0 < r \le R, \tag{5.19}$$

where $C = C(n, G, \gamma, \alpha)$.

Proof. Starting from a scale R, the idea of the proof is to use (5.17) to obtain the decay of the excess up to a scale r_+. Then, in order to use (5.7) we use the scaling properties of the excess (see Proposition 2.7) to jump to a scale r_-. Setting $L = \mathbf{e}(E,R) + \Lambda R$, since we want that in particular $\mathbf{e}(E, r_+) \le C((r_+/R)L + r_+^{2\alpha})$, in light of (5.17) we need to take $r_+ = \varepsilon^{1-\beta'}$ for some $\beta' \in (0,1)$. Similarly, since we want $\mathbf{e}(E, r_-) \le Cr_-^{2\alpha}$, (5.7) imposes $r_- = \varepsilon^{1+\beta''}$ for some $\beta'' > 0$. Therefore, when relying on Proposition 2.7 to pass from r_+ to r_- we will lose a factor in the estimates. To compensate that we need to assume that the starting scale R is much larger than r_+ i.e. $R = \varepsilon^{1-\beta}$ for some $1 > \beta > \beta'$.

We start by choosing β arbitrarily close to 1, $p \in (0, p_0)$ arbitrarily close to p_0, and $\theta \in (0,1)$ such that

$$(1-\theta)(n-1+p_0) = (n-1+p).$$

For β' to be chosen later, if $r_+ = \varepsilon^{1-\beta'}$ we have that for ε small enough depending on β' that (5.16) is satisfied (with $\eta = \varepsilon^{-\beta'}$) so that (5.17) together with (H5) yield

$$\mathbf{e}(E, r) \le C\left(\frac{r}{R}L + \left(\frac{\varepsilon}{r}\right)^{n-1+p}\right) \qquad \text{for } R \ge r \ge r_+.$$

This gives (5.19) provided

$$\left(\frac{\varepsilon}{r}\right)^{n-1+p} \le r^{2\alpha} \qquad \text{for } R \ge r \ge r_+$$

which is equivalent to

$$\left(\frac{\varepsilon}{r_+}\right)^{n-1+p} \le r_+^{2\alpha}.$$

Since $r_+ = \varepsilon^{1-\beta'}$, this gives the condition

$$2\alpha \le \frac{\beta'}{1-\beta'}(n-1+p). \tag{5.20}$$

Thanks to Proposition 2.7 and $L \leq 1$, we then have

$$\mathbf{e}(E,r) \leq \left(\frac{r_+}{r}\right)^{n-1} \mathbf{e}(E,r_+) \leq C\left(\frac{r_+}{r}\right)^{n-1} \left(\frac{r_+}{R} + \left(\frac{\varepsilon}{r_+}\right)^{n-1+p}\right) \qquad \text{for } r_+ \geq r \geq r_-. \tag{5.21}$$

In order to have $\mathbf{e}(E,r) \leq Cr^{2\alpha}$ in this range it is enough to have this estimate for $r = r_-$ which means

$$\varepsilon^{-(\beta'+\beta'')(n-1)}(\varepsilon^{\beta-\beta'} + \varepsilon^{\beta'(n-1+p)}) \leq \varepsilon^{2\alpha(1+\beta'')}.$$

In particular, the optimal choice is

$$\beta' = \frac{\beta}{n+p}, \tag{5.22}$$

for which the condition becomes

$$2\alpha \leq \frac{\beta' p - (n-1)\beta''}{1+\beta''}. \tag{5.23}$$

Let us point out that (5.23) is stronger than (5.20). Notice that under (5.23), we have in particular that (5.6) is satisfied at $R = r_-$ provided ε is small enough. We may thus use (5.7) to obtain

$$\mathbf{e}(E,r) \leq C\left(\frac{r}{r_-}\right)^{1-s_0} \left(\mathbf{e}(E,r_-) + \left(\frac{r_-}{\varepsilon}\right)^{1-s_0}\right) \qquad \text{for } r_- \geq r > 0.$$

By (5.21) (with the choice (5.22)) this reduces to

$$\mathbf{e}(E,r) \leq C\left(\frac{r}{r_-}\right)^{1-s_0} \left(\varepsilon^{\beta' p-(n-1)\beta''} + \varepsilon^{(1-s_0)\beta''}\right) \qquad \text{for } r_- \geq r > 0.$$

In particular $\mathbf{e}(E,r) \leq Cr^{2\alpha}$ for $r \leq r_-$ provided it holds for $r = r_-$ i.e.

$$\varepsilon^{\beta' p-(n-1)\beta''} + \varepsilon^{\beta''(1-s_0)} \leq \varepsilon^{2\alpha(1+\beta'')}$$

which gives the condition

$$2\alpha \leq \min\left(\frac{\beta' p-(n-1)\beta''}{1+\beta''}, \frac{\beta''}{1+\beta''}(1-s_0)\right). \tag{5.24}$$

This constraint implies (5.23) (which itself implies (5.20)). Since the first term is decreasing in β'' and the second one is increasing, we see that the optimal choice of β'' is obtained when both terms are equal, that is

$$\beta'' = \frac{\beta p}{(n-s_0)(n+p)}$$

which gives

$$2\alpha \leq \frac{\beta p}{(n-s_0)(n+p)+\beta p}(1-s_0).$$

Since p can be taken arbitrarily close to p_0 and β arbitrarily close to 1. This concludes the proof.

□

Acknowledgments. M. Goldman was partially supported by the ANR grant SHAPO. B. Merlet and M. Pegon are partially supported by the Labex CEMPI (ANR-11-LABX-0007-01).

References

[1] S. Alama, L. Bronsard, I. Topaloglu, A. Zuniga : A nonlocal isoperimetric problem with density perimeter, *Calc. Var. Partial Differ. Equ.* **60.1** (2021), pp. 1.

[2] L. Ambrosio, V. Caselles, S. Masnou, J.-M. Morel : Connected components of sets of finite perimeter and applications to image processing, *J. Eur. Math. Soc.* **3.1** (2001), pp. 39–92.

[3] L. Ambrosio, N. Fusco, D. Pallara : *Functions of bounded variation and free discontinuity problems*, Oxford mathematical monographs, Clarendon Press, (2000).

[4] L. Caffarelli, E. Valdinoci : Regularity properties of nonlocal minimal surfaces via limiting arguments, *Advances in Mathematics* **248** (2013), pp. 843–871.

[5] D. Carazzato, N. Fusco, A. Pratelli : Minimality of balls in the small volume regime for a general Gamow-type functional, *Advances in Calculus of Variations* **0.0** (2021).

[6] A. Cesaroni, M. Novaga : Second-order asymptotics of the fractional perimeter as $s \to 1$, *Mathematics in Engineering* **2.3** (2020), pp. 512–526.

[7] R. Choksi, C. B. Muratov, I. Topaloglu : An old problem resurfaces nonlocally: Gamow's liquid drops inspire today's research and applications, *Notices Amer. Math. Soc.* **64.11** (2017), pp. 1275–1283.

[8] M. Cicalese, G. P. Leonardi : A selection principle for the sharp quantitative isoperimetric inequality, *Arch Rational Mech Anal* **206.2** (2012), pp. 617–643.

[9] G. David, S. Semmes : Surfaces quasiminimales de codimension 1 et domaines de John, *Séminaire Équations aux dérivées partielles (Polytechnique)* (1995–1996).

[10] J. Dávila : On an open question about functions of bounded variation, *Calc. Var. Partial Differ. Equ.* **15.4** (2002), pp. 519–527.

[11] L. De Luca, A. Kubin, M. Ponsiglione : The core-radius approach to supercritical fractional perimeters, curvatures and geometric flows, *Nonlinear Analysis* **214** (2022), pp. 112585.

[12] G. De Philippis, J. Hirsch, G. Vescovo : Regularity of minimizers for a model of charged droplets, (2019), URL: http://arxiv.org/abs/1901.02546.

[13] A. Di Castro, M. Novaga, B. Ruffini, E. Valdinoci : Nonlocal quantitative isoperimetric inequalities, *Calc. Var.* **54.3** (2015), pp. 2421–2464.

[14] A. Figalli, N. Fusco, F. Maggi, V. Millot, M. Morini : Isoperimetry and stability properties of balls with respect to nonlocal energies, *Commun. Math. Phys.* **336.1** (2015), pp. 441–507.

[15] R. L. Frank, P. T. Nam : Existence and nonexistence in the liquid drop model, *Calc. Var.* **60.6** (2021), pp. 223.

[16] N. Fusco : The quantitative isoperimetric inequality and related topics, *Bull. Math. Sci.* **5.3** (2015), pp. 517–607.

[17] N. Fusco, L. Esposito : A remark on a free interface problem with volume constraint, *J. Convex Anal.* **18.2** (2011), pp. 417–426.

[18] M. Goldman, M. Novaga : Volume-constrained minimizers for the prescribed curvature problem in periodic media, *Calc. Var. Partial Differ. Equ.* **44.3-4** (2012), pp. 297–318.

[19] M. Goldman, M. Novaga, B. Ruffini : Existence and stability for a non-local isoperimetric model of charged liquid drops, *Arch Rational Mech Anal* **217.1** (2015), pp. 1–36.

[20] M. Goldman, M. Novaga, B. Ruffini : Reifenberg flatness for almost-minimizers of the perimeter under minimal assumptions, *Proc. Am. Math. Soc.* (2021).

[21] M. Goldman, M. Novaga, B. Ruffini : *Rigidity of the ball for an isoperimetric problem with strong capacitary repulsion*, 2022.

[22] H. Knüpfer, C. B. Muratov : On an isoperimetric problem with a competing nonlocal term II: the general case, *Comm. Pure Appl. Math.* **67.12** (2014), pp. 1974–1994.

[23] H. Knüpfer, C. B. Muratov, M. Novaga : Emergence of nontrivial minimizers for the three-dimensional Ohta–Kawasaki energy, *Pure Appl. Analysis* **2.1** (2020), pp. 1–21.

[24] H. Knüpfer, W. Shi : Second order expansion for the nonlocal perimeter functional, (2021), URL: http://arxiv.org/abs/2110.12378.

[25] J. Lu, F. Otto : Nonexistence of a minimizer for Thomas-Fermi-Dirac-von Weizsäcker model, *Comm. Pure Appl. Math.* **67.10** (2014), pp. 1605–1617.

[26] F. Maggi : *Sets of Finite Perimeter and Geometric Variational Problems: An Introduction to Geometric Measure Theory*, Cambridge University Press, (2012).

[27] B. Merlet, M. Pegon : Large mass rigidity for a liquid drop model in 2D with kernels of finite moments, *Journal de l'École polytechnique — Mathématiques* **9** (2021), pp. 63–100.

[28] C. B. Muratov, T. M. Simon : A nonlocal isoperimetric problem with dipolar repulsion, *Commun. Math. Phys.* **372.3** (2019), pp. 1059–1115.

[29] M. Pegon : Large mass minimizers for isoperimetric problems with integrable nonlocal potentials, *Nonlinear Analysis* **211** (2021), pp. 112395.

[30] S. Rigot : Ensembles quasi-minimaux avec contrainte de volume et rectifiabilité uniforme, *Mémoires Soc. Math. France* **1** (2000), pp. 1–104.

[31] I. Tamanini : Boundaries of Caccioppoli sets with Hölder-continuous normal vector, *Journal für die reine und angewandte Mathematik* **1982.334** (1982), pp. 27–39.

CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France
Email address: michael.goldman@cnrs.fr

Univ. Lille, CNRS, UMR 8524, Inria - Laboratoire Paul Painlevé, F-59000 Lille
Email address: benoit.merlet@univ-lille.fr
Email address: marc.pegon@univ-lille.fr