Level-set approximation of noisy functions

Matteo Croci (Basque Center for Applied Mathematics & Ikerbasque)
Joint with: I. C. J. Powell and A.-L. Haji-Ali (Herriot-Watt University)

MCQMC 2026 – 17th International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, University of Edinburgh, 10th June 2026

Overview

1. Motivation & Introduction

Statistical function approximation

Let $P_0(x, \omega)$ be a random QoI with $x \in D \subset \mathbb{R}^d$.

Function estimation: approximate $f(x) = \mathbb{E}[P_0(x, \omega)]$,

Statistical function approximation

Let $P_0(x, \omega)$ be a random QoI with $x \in D \subset \mathbb{R}^d$.

Function estimation: approximate $f(x) = \mathbb{E}[P_0(x, \omega)]$,

Examples: Failure probabilities, density functions, characteristic functions, etc.

Statistical function approximation

Let $P_0(x, \omega)$ be a random QoI with $x \in D \subset \mathbb{R}^d$.

Function estimation: approximate $f(x) = \mathbb{E}[P_0(x, \omega)]$,

Examples: Failure probabilities, density functions, characteristic functions, etc.

Standard approach: Monte Carlo $+$ function approximation. Expensive!

Specific objective: Level-sets of noisy functions

Let $f(x)$ be only accessible via expensive and noisy point evaluations (e.g., due to MC).

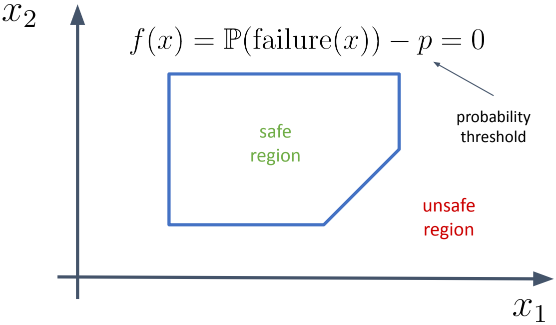
Objective: Approximate the zero level set of f (the set of points where $f = 0$).

Specific objective: Level-sets of noisy functions

Let $f(x)$ be only accessible via expensive and noisy point evaluations (e.g., due to MC).

Objective: Approximate the zero level set of f (the set of points where $f = 0$).

Example: If $f(x)$ is a failure probability, its contours are failure regions, key in design under uncertainty and reliability engineering.



$$f(x) = \mathbb{P}(\text{failure}(x)) - p = 0$$

probability threshold

2. Problem and algorithm overview

Problem statement and general strategy

For any $x \in \overline{D} \subset \mathbb{R}^d$, let $\tilde{f}(x, \omega)$ for $\omega \in \Omega$ be an (expensive, random) estimator of $f(x)$.

Objective. Given (possibly independent) point values of $\tilde{f}$, approximate

$$\mathcal{L}_0 := \{x \in \overline{D} : \quad f(x) = 0\}.$$

Problem statement and general strategy

For any $x \in \overline{D} \subset \mathbb{R}^d$, let $\tilde{f}(x, \omega)$ for $\omega \in \Omega$ be an (expensive, random) estimator of $f(x)$.

Objective. Given (possibly independent) point values of $\tilde{f}$, approximate

$$\mathcal{L}_0 := \{x \in \overline{D} : \quad f(x) = 0\}.$$

General strategy:

Problem statement and general strategy

For any $x \in \overline{D} \subset \mathbb{R}^d$, let $\tilde{f}(x,\omega)$ for $\omega \in \Omega$ be an (expensive, random) estimator of $f(x)$.

Objective. Given (possibly independent) point values of $\tilde{f}$, approximate

$$\mathcal{L}_0 := \{x \in \overline{D} : \quad f(x) = 0\}.$$

General strategy:

1. Construct a *proxy rootfinder*‡ (surrogate).

Problem statement and general strategy

For any $x \in \overline{D} \subset \mathbb{R}^d$, let $\tilde{f}(x, \omega)$ for $\omega \in \Omega$ be an (expensive, random) estimator of $f(x)$.

Objective. Given (possibly independent) point values of $\tilde{f}$, approximate

$$\mathcal{L}_0 := \{x \in \overline{D} : \quad f(x) = 0\}.$$

General strategy:

1. Construct a *proxy rootfinder*‡ (surrogate).
2. Apply state-of-the-art contour approximation techniques from computer vision, e.g., marching cubes, isosurface triangulation techniques, etc.

‡Noise-free proxy rootfinders [Boyd, 2003, 2013], [Townsend and Trefethen, 2013].

Constructing the proxy

Goal-oriented adaptivity: We want to construct an adaptive approximation in space and in terms of MC sampling that specifically targets the level set.

Constructing the proxy

Goal-oriented adaptivity: We want to construct an adaptive approximation in space and in terms of MC sampling that specifically targets the level set.

What we need: Local approximations and local error bounds.

Constructing the proxy

Goal-oriented adaptivity: We want to construct an adaptive approximation in space and in terms of MC sampling that specifically targets the level set.

What we need: Local approximations and local error bounds.

Algorithm sketch

Constructing the proxy

Goal-oriented adaptivity: We want to construct an adaptive approximation in space and in terms of MC sampling that specifically targets the level set.

What we need: Local approximations and local error bounds.

Algorithm sketch

Start from a (quasi-)uniform mesh.

Constructing the proxy

Goal-oriented adaptivity: We want to construct an adaptive approximation in space and in terms of MC sampling that specifically targets the level set.

What we need: Local approximations and local error bounds.

Algorithm sketch

Start from a (quasi-)uniform mesh. Repeat until convergence:

Constructing the proxy

Goal-oriented adaptivity: We want to construct an adaptive approximation in space and in terms of MC sampling that specifically targets the level set.

What we need: Local approximations and local error bounds.

Algorithm sketch

Start from a (quasi-)uniform mesh. Repeat until convergence:

- Over each cell $\Box$, evaluate $\tilde{f}$ at given nodes (e.g., via MC).

Constructing the proxy

Goal-oriented adaptivity: We want to construct an adaptive approximation in space and in terms of MC sampling that specifically targets the level set.

What we need: Local approximations and local error bounds.

Algorithm sketch

Start from a (quasi-)uniform mesh. Repeat until convergence:

- Over each cell $\square$, evaluate $\tilde{f}$ at given nodes (e.g., via MC).
- Use an approximation operator $I^{\square}$ to construct a local approximation $\hat{f}^{\square} = I^{\square}\tilde{f}$ from these evaluations. Think interpolation, projection, or regression.

Constructing the proxy

Goal-oriented adaptivity: We want to construct an adaptive approximation in space and in terms of MC sampling that specifically targets the level set.

What we need: Local approximations and local error bounds.

Algorithm sketch

Start from a (quasi-)uniform mesh. Repeat until convergence:

- Over each cell $\Box$, evaluate $\tilde{f}$ at given nodes (e.g., via MC).
- Use an approximation operator $I^{\Box}$ to construct a local approximation $\hat{f}^{\Box} = I^{\Box}\tilde{f}$ from these evaluations. Think interpolation, projection, or regression.
- Refine cells and estimator $\tilde{f}$ close to the level set based on information from $\hat{f}^{\Box}$.

Constructing the proxy

Goal-oriented adaptivity: We want to construct an adaptive approximation in space and in terms of MC sampling that specifically targets the level set.

What we need: Local approximations and local error bounds.

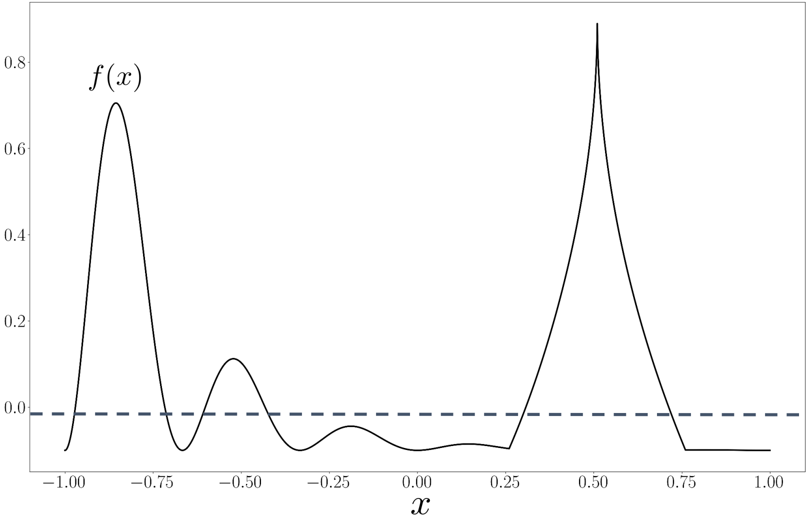
Algorithm sketch

Start from a (quasi-)uniform mesh. Repeat until convergence:

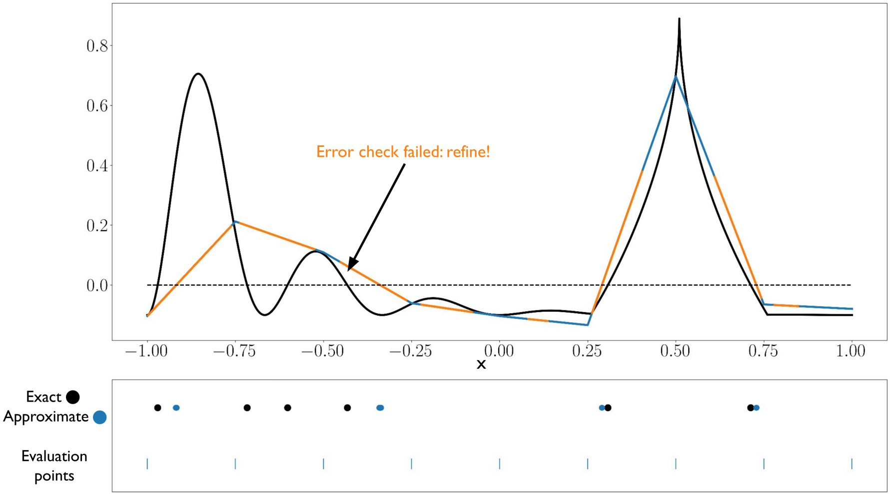
- Over each cell $\square$, evaluate $\tilde{f}$ at given nodes (e.g., via MC).
- Use an approximation operator $I^{\square}$ to construct a local approximation $\hat{f}^{\square} = I^{\square}\tilde{f}$ from these evaluations. Think interpolation, projection, or regression.
- Refine cells and estimator $\tilde{f}$ close to the level set based on information from $\hat{f}^{\square}$.

Note: We use piecewise-polynomial interpolation in numerical experiments.

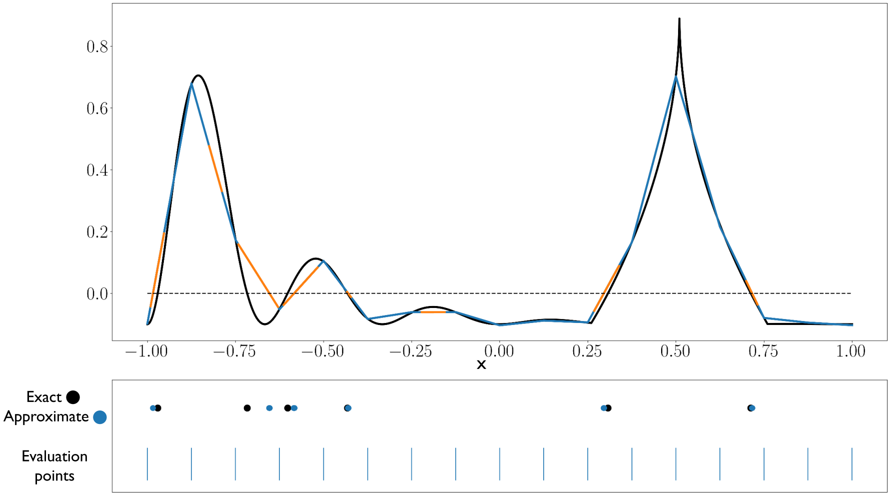
How the algorithm works in practice



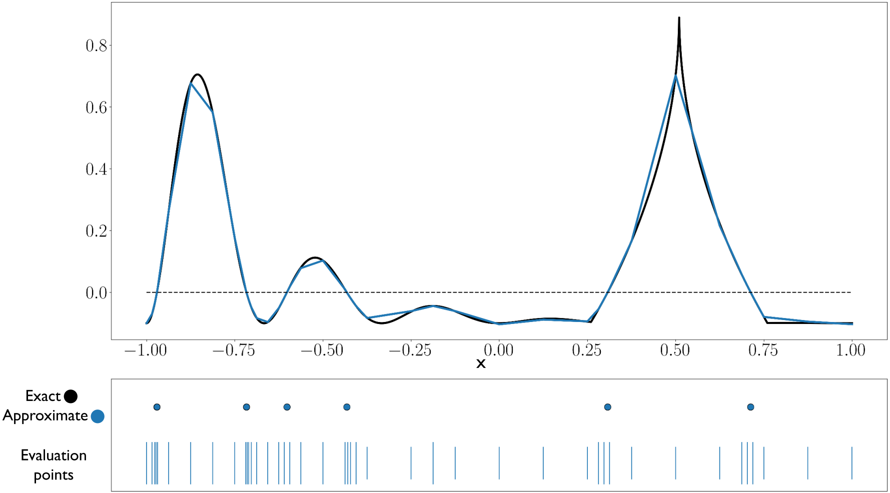
How the algorithm works in practice


Error check failed: refine!

How the algorithm works in practice



How the algorithm works in practice



3. Analysis

Assumptions on f

Regularity of f

Let $p \geq 1$ and $\alpha > 1 + 1/p$. f is Lipschitz-continuous over $\overline{D}$ and $f \in W^{\alpha,p}(D)$.

Assumptions on f

Regularity of f

Let $p \geq 1$ and $\alpha > 1 + 1/p$. f is Lipschitz-continuous over $\overline{D}$ and $f \in W^{\alpha,p}(D)$.

Regularity of the level set

For all $A \subseteq \bar{D}$, there exist $\rho \in [0, \infty)$, $\delta \in (0, 1]$ s.t. for all $0 \leq a \leq \delta$ diam(A) we have

$$\left|\{x \in \overline{A} : |f(x)| \leq a\}\right| \leq \rho a.$$

Assumptions on f

Regularity of f

Let $p \geq 1$ and $\alpha > 1 + 1/p$. f is Lipschitz-continuous over $\overline{D}$ and $f \in W^{\alpha,p}(D)$.

Regularity of the level set

For all $A \subseteq \bar{D}$, there exist $\rho \in [0, \infty)$, $\delta \in (0, 1]$ s.t. for all $0 \leq a \leq \delta$ diam(A) we have

$$\left|\{x \in \overline{A} : |f(x)| \leq a\}\right| \leq \rho a.$$

Notation

We denote with $\gamma(\square) \geq 0$ a generic cell-dependent quantity that is p-summable over all cells of any (quasi-)uniform mesh of D, i.e., $\exists c_\gamma < \infty$ s.t. $\left(\sum_\square \gamma^p(\square)\right)^{1/p} \leq c_\gamma$.

Assumptions on the approximation

Noise-free accuracy of approximation operator

For a cell $\square$, let $h = \text{diam}(\square)$. Over all cells, the approximation operator $I^{\square}$ is solely based on point evaluations and satisfies

$$\|f - I^{\square} f\|_{L^p(\square)} \lesssim h^{\alpha} |f|_{W^{\alpha,p}(\square)} = \gamma(\square) h^{\alpha}.$$

Assumptions on the approximation

Noise-free accuracy of approximation operator

For a cell $\square$, let $h = \text{diam}(\square)$. Over all cells, the approximation operator $I^\square$ is solely based on point evaluations and satisfies

$$\|f - I^\square f\|_{L^p(\square)} \lesssim h^\alpha |f|_{W^{\alpha,p}(\square)} = \gamma(\square) h^\alpha.$$

Accuracy and cost of pointwise stochastic estimator

Let $p \geq 1$. For all $x \in \overline{D}$, the estimator $\tilde{f}(x,\omega)$ has bounded p-th moment and it can be evaluated so that a cost of $M(x) > 0$ corresponds to an error bound

$$\mathbb{E}\left[|f(x) - \tilde{f}(x,\cdot)|^p\right]^{1/p} \leq \sigma(x) M^{-\beta}(x),$$

for some $\sigma(x) > 0$ that is bounded everywhere over $\overline{D}$.

Error bound on local approximation

Theorem (approximation error)

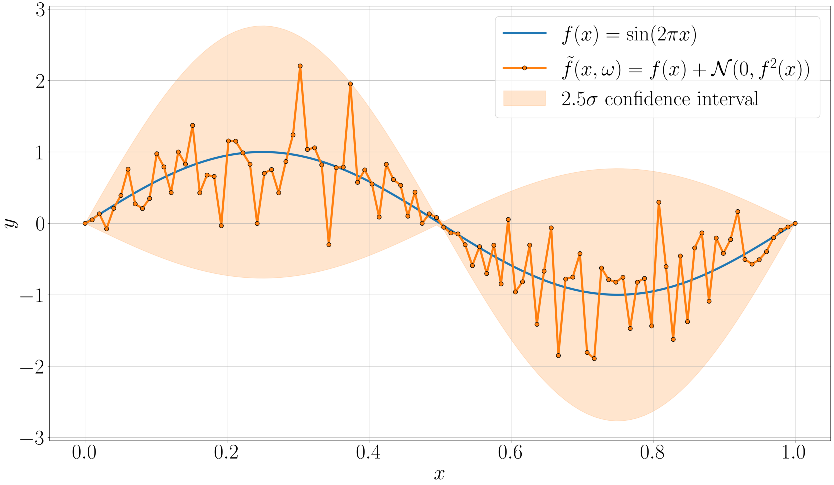
Let $\hat{f}^{\square} = I^{\square}\tilde{f}$ and choose $M(x_e) \sim (\sigma(x_e))^{1/\beta} h^{-\alpha/\beta}$ for all evaluation points $x_e \in \overline{\square}$. Then, $\hat{f}^{\square}$ is well defined and satisfies

$$\mathbb{E}\left[\|f - \hat{f}^{\square}\|^p_{L^p(\square)}\right]^{1/p} \lesssim \gamma(\square) h^{\alpha}.$$

Theoretical twist. $\tilde{f}$ not assumed continuous due to possibly independent estimations.

Local approximation with independent estimators

Key challenge. $\tilde{f}$ not assumed continuous due to possibly independent estimations.



Error measure and goal-oriented refinement criterion

Error measure and goal-oriented refinement criterion

Error measure: $E(D) = \sum_{\square} E(\square)$, where $E(\square) = \int_{\square} \mathbb{E}\left[\left|\mathbb{I}_{f \leq 0}(x) - \mathbb{I}_{\hat{f} \leq 0}(x)\right|\right] \mathrm{d}x.$

Error measure and goal-oriented refinement criterion

Error measure: $E(D) = \sum_{\square} E(\square)$, where $E(\square) = \int_{\square} \mathbb{E}\left[\left|\mathbb{I}_{f\leq 0}(x) - \mathbb{I}_{\hat{f}\leq 0}(x)\right|\right] \mathrm{d}x.$

Theorem (Goal-oriented refinement criterion)

Error measure and goal-oriented refinement criterion

Error measure: $E(D) = \sum_{\square} E(\square)$, where $E(\square) = \int_{\square} \mathbb{E}\left[\left|\mathbb{I}_{f\leq 0}(x) - \mathbb{I}_{\hat{f}\leq 0}(x)\right|\right] \mathrm{d}x.$

Theorem (Goal-oriented refinement criterion)

Let $0 \leq s = (p+1)^{-1/2} \leq 2^{-1/2} < 1$, and let $\hat{e}(\square)$ be any quantity (an a priori/posteriori error estimator) such that:

$$\mathbb{E}\left[\|f - \hat{f}^{\square}\|^p_{L^p(\square)}\right]^{1/p} \leq \hat{e}(\square) \leq \gamma(\square)h^{\alpha}.$$

Error measure and goal-oriented refinement criterion

Error measure: $E(D) = \sum_{\square} E(\square)$, where $E(\square) = \int_{\square} \mathbb{E}\left[\left| \mathbb{I}_{f \le 0}(x) - \mathbb{I}_{\hat{f} \le 0}(x) \right| \right] \mathrm{d}x$.

Theorem (Goal-oriented refinement criterion)

Let $0 \le s = (p+1)^{-1/2} \le 2^{-1/2} < 1$, and let $\hat{e}(\square)$ be any quantity (an a priori/posteriori error estimator) such that:

$$\mathbb{E}\left[\| f - \hat{f}^{\square} \|_{L^p(\square)}^p \right]^{1/p} \le \hat{e}(\square) \le \gamma(\square) h^{\alpha}.$$

For all $\varepsilon \in (0,1)$, let $h_L(\varepsilon) = \varepsilon^{\frac{p+1}{\alpha p}}$ and $h_1(\varepsilon) = (h_L(\varepsilon))^s$. Then, $0 < h_L < h_1$ and

Error measure and goal-oriented refinement criterion

Error measure: $E(D) = \sum_{\square} E(\square)$, where $E(\square) = \int_{\square} \mathbb{E}\left[\left|\mathbb{I}_{f \leq 0}(x) - \mathbb{I}_{\hat{f} \leq 0}(x)\right|\right] \mathrm{d}x$.

Theorem (Goal-oriented refinement criterion)

Let $0 \leq s = (p+1)^{-1/2} \leq 2^{-1/2} < 1$, and let $\hat{e}(\square)$ be any quantity (an a priori/posteriori error estimator) such that:

$$\mathbb{E}\left[\|f - \hat{f}^{\square}\|^{p}_{L^p(\square)}\right]^{1/p} \leq \hat{e}(\square) \leq \gamma(\square) h^{\alpha}.$$

For all $\varepsilon \in (0,1)$, let $h_L(\varepsilon) = \varepsilon^{\frac{p+1}{\alpha p}}$ and $h_1(\varepsilon) = (h_L(\varepsilon))^s$. Then, $0 < h_L < h_1$ and

$$\text{IF} \quad h \leq h_L \quad \text{OR} \quad \left(h \leq h_1(\varepsilon) \quad \text{AND} \quad \hat{e}(\square)^{1-s} \leq |\square|^{\frac{1-s}{p}} \inf_{x \in \square} |\hat{f}^{\square}(x,\omega)| \right),$$

THEN $E(\square) \leq \gamma_E^p(\square)\varepsilon$.

Refinement criterion - some intuition

Take a deterministic problem so that $p = \infty$, $s = 0$, and there is no noise.

Refinement criterion - some intuition

Take a deterministic problem so that $p = \infty$, $s = 0$, and there is no noise.

Note that, over any cell $\square$, if either

$$|f - \hat{f}| \leq |\hat{f}| \quad \text{OR} \quad |f - \hat{f}| \leq |f|, \quad \forall x \in \overline{\square},$$

then the error is too small to cause a change in sign and $E(\square) = 0$.

Refinement criterion - some intuition

Take a deterministic problem so that $p = \infty$, $s = 0$, and there is no noise.

Note that, over any cell $\square$, if either

$$|f - \hat{f}| \leq |\hat{f}| \quad \text{OR} \quad |f - \hat{f}| \leq |f|, \quad \forall x \in \overline{\square},$$

then the error is too small to cause a change in sign and $E(\square) = 0$.

Reliable pointwise error estimation is tricky, but in $L^p(\square)$ it is not.

For any estimator $\hat{e}(\square)$ upper bounding the $L^\infty(\square)$ error, we can thus strengthen the above condition to, e.g.,

$$\|f - \hat{f}\|_{L^\infty(\square)} \leq \hat{e}(\square) \leq \inf_{x \in \overline{\square}} |\hat{f}|,$$

which is the condition in our refinement criterion.

The full algorithm

Adaptive level-set approximation algorithm

Input: Domain D, $\alpha, \beta > 0$, tolerance $\varepsilon > 0$, pointwise estimator $\tilde{f}(x, \omega)$ and error estimator $\hat{e}(\square)$. Initial quasi-uniform mesh of D.

Repeat until convergence:

- Evaluate $\tilde{f}$ at all cell nodes $\{x_e\}$ with budget $M(x_e) \sim (\sigma(x_e))^{1/\beta} h^{-\alpha/\beta}$.
- Update approximation $\hat{f}$ over the cell.
- **If** goal-oriented convergence criterion is not satisfied: Refine cell uniformly.

Proxy rootfinding: Find the zero level set of the resulting $\hat{f}$ going cell by cell.

Output: The approximant $\hat{f}$ and its zero level set.

Global convergence

Theorem (Global convergence)

Under the assumptions of the previous theorem, the adaptive algorithm converges in a finite number of steps to a $\hat{f}$ satisfying $E(D) \leq \theta\varepsilon$ for some $\theta > 0$.

Global convergence

Theorem (Global convergence)

Under the assumptions of the previous theorem, the adaptive algorithm converges in a finite number of steps to a $\hat{f}$ satisfying $E(D) \leq \theta\varepsilon$ for some $\theta > 0$.

Note: To obtain $E(D) \leq \varepsilon$ it is sufficient to rescale ε by θ.

Limitation: The constant θ depends on α, p and c_{γ_E}. The latter is difficult to estimate so in practice one can only make h_1 and h_L (or ε) small enough.

Cost and complexity analysis

Theorem (Total cost complexity)

The total expected cost of our algorithm satisfies

$$\mathbb{E}[W] \lesssim \varepsilon^{-\left(\frac{p+1}{p}\right)\left(\frac{1}{\beta}+\frac{d-1}{\alpha}\right)}.$$

This results in a cost-complexity reduction of $\mathcal{O}\left(\varepsilon^{\left(\frac{p+1}{\alpha p}\right)}\right)$ wrt a non-adaptive scheme.

Cost and complexity analysis

Theorem (Total cost complexity)

The total expected cost of our algorithm satisfies

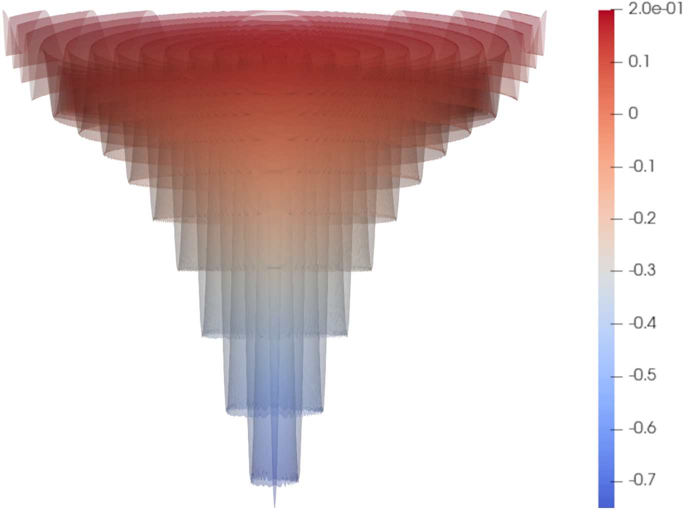
$$\mathbb{E}[W] \lesssim \varepsilon^{-\left(\frac{p+1}{p}\right)\left(\frac{1}{\beta}+\frac{d-1}{\alpha}\right)}.$$

This results in a cost-complexity reduction of $\mathcal{O}\left(\varepsilon^{\left(\frac{p+1}{\alpha p}\right)}\right)$ wrt a non-adaptive scheme.

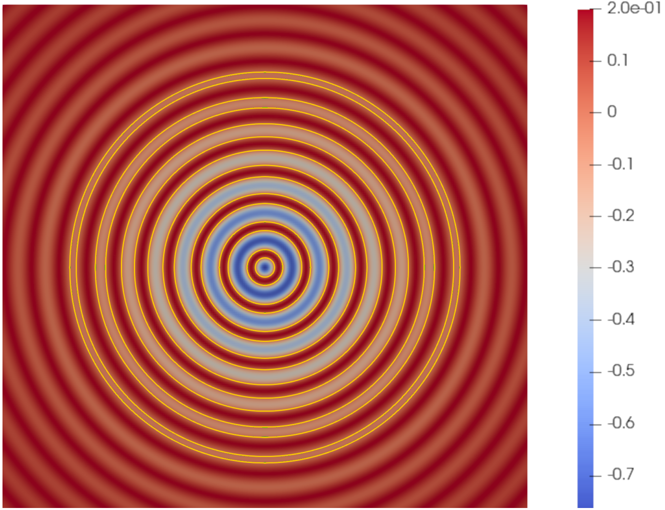
Same complexity as doing uniform refinement directly on the level set, but without knowing where it is beforehand!

4. Numerical results

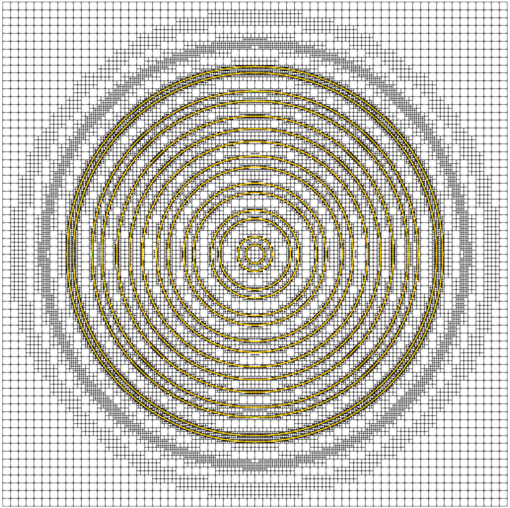
2D drop-wave function perturbed by noise



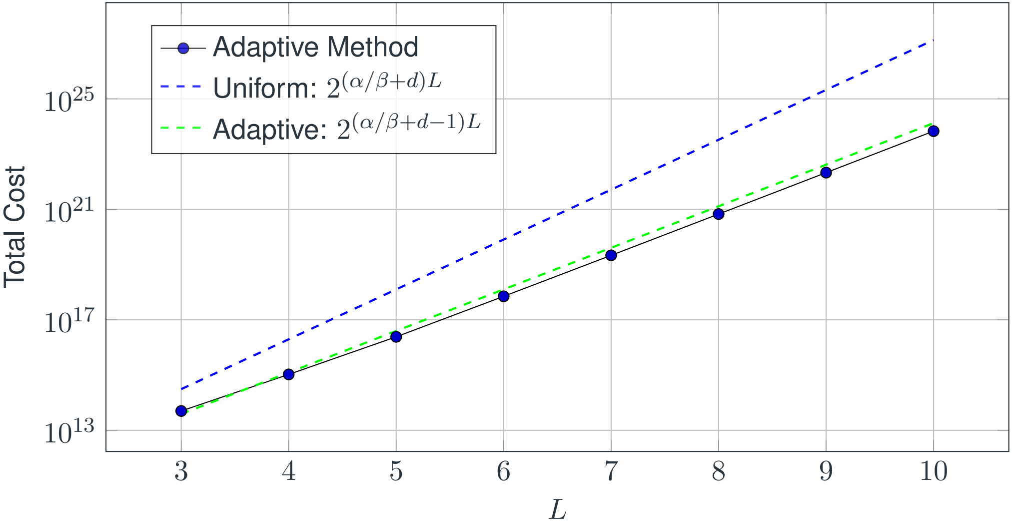
2D drop-wave function perturbed by noise



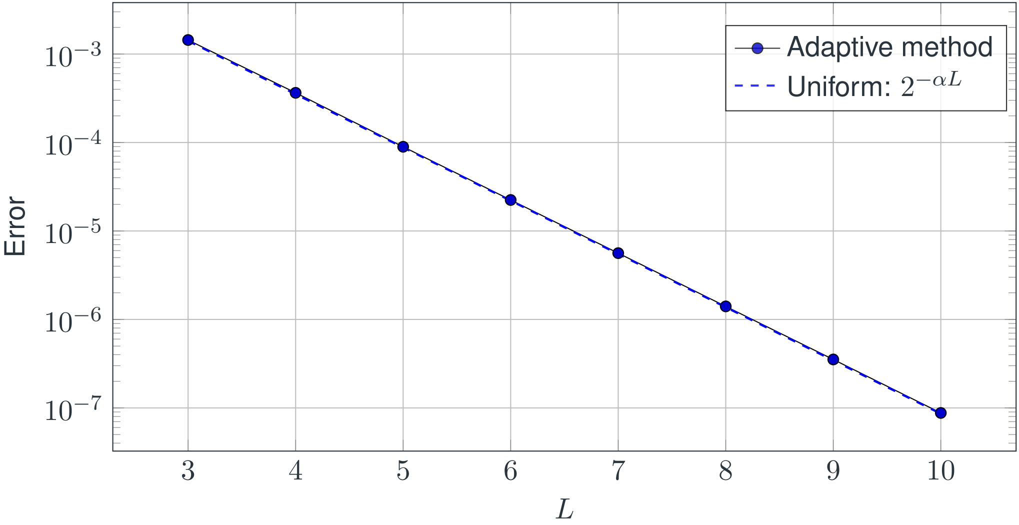
2D drop-wave function perturbed by noise



2D drop-wave function perturbed by noise

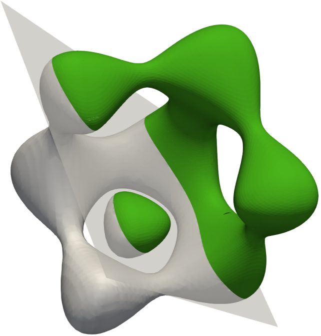


2D drop-wave function perturbed by noise

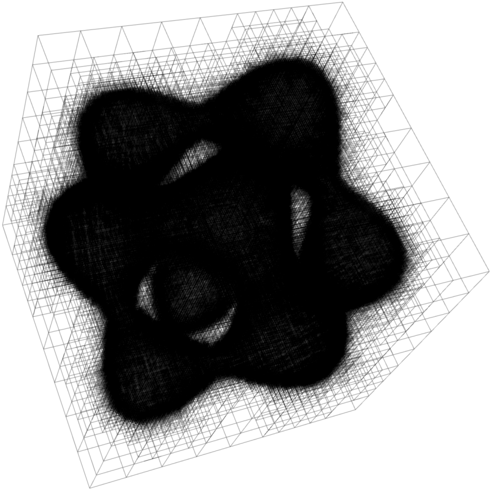


3D Styblinski-Tang function perturbed by noise

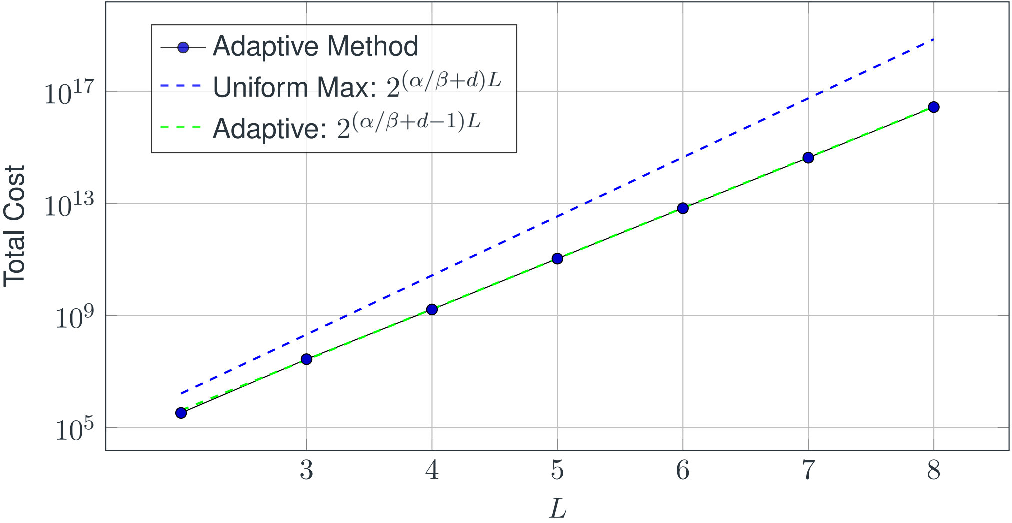
$$f(x) = \frac{1}{122}\left(\sum_{i=1}^{d} x_i^4 - 16x_i^2 + 5x_i\right) + 1, \quad x \in [-5,5]^d.$$



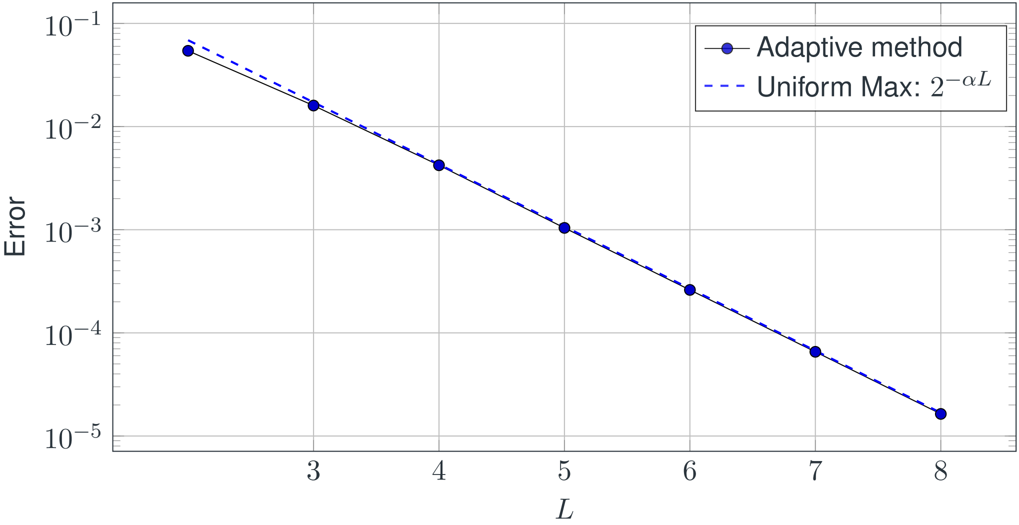
3D Styblinski-Tang function perturbed by noise



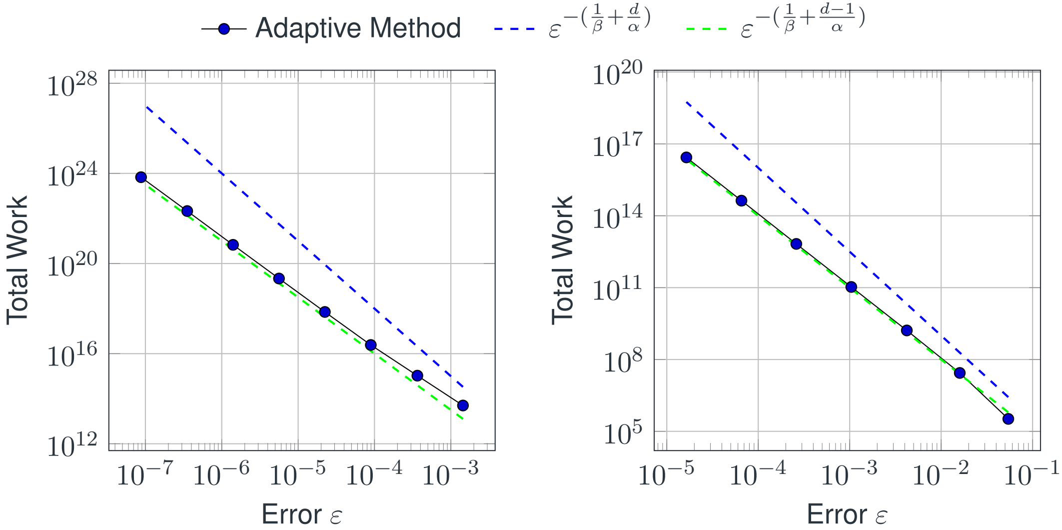
3D Styblinski-Tang function perturbed by noise



3D Styblinski-Tang function perturbed by noise



Cost complexity (left: 2D, right: 3D)



Hyperelastic beam with uncertain Lamé parameters

$$\min_{u} \int_S \psi(u, \omega)\ \mathrm{d}s - \int_S B \cdot u\ \mathrm{d}s - \int_{\partial S} T \cdot u\ \mathrm{d}s$$

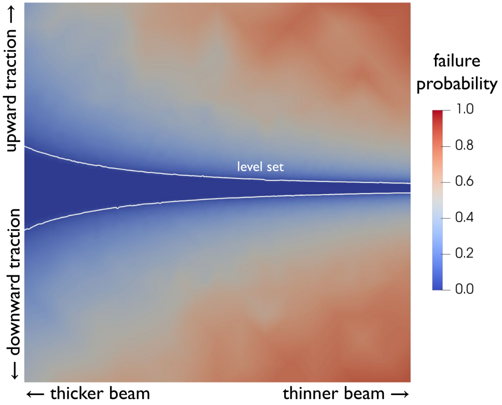
Uncertain elastic stored energy density model

Body force (gravity)

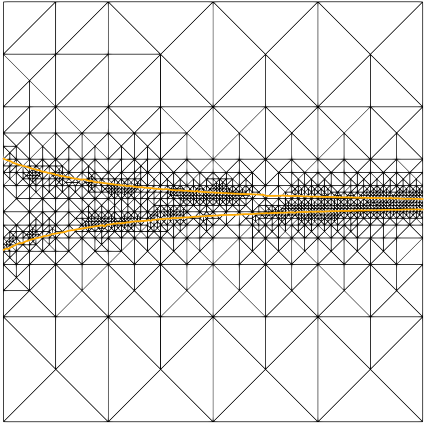
Traction

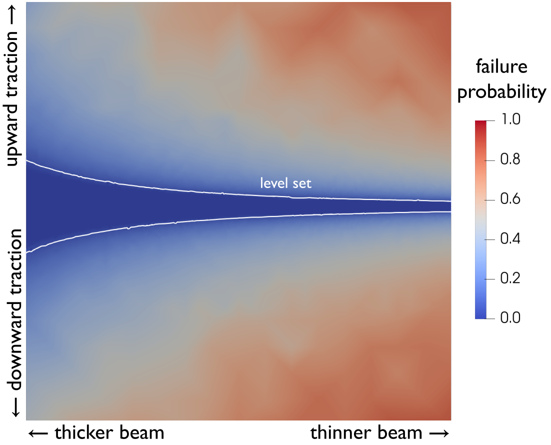
$$f(x) = \mathbb{P}(\text{failure}(x)) - p = \mathbb{P}(\|u\|_{L^\infty(S)} > u_\text{f}) - p$$

x = { beam thickness, traction }

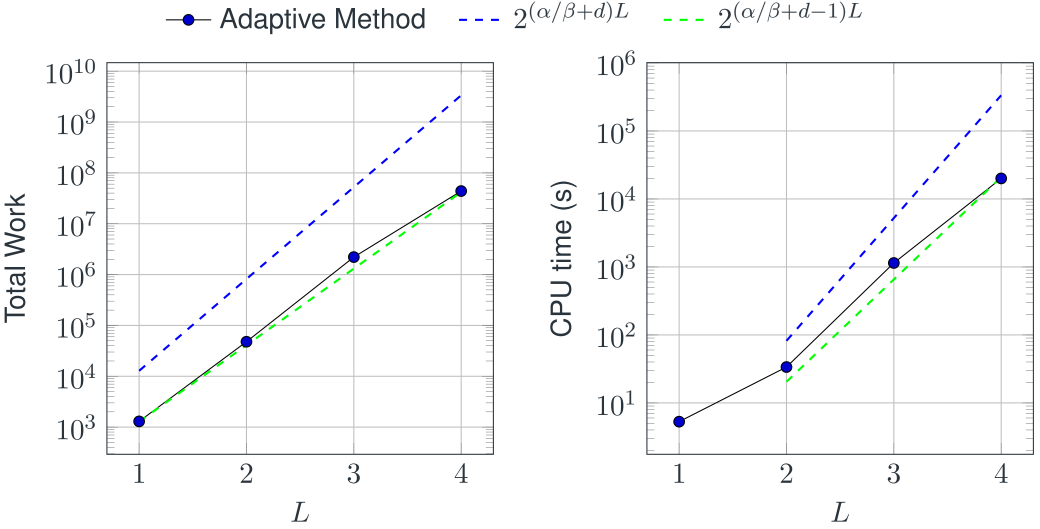


Hyperelastic beam with uncertain Lamé parameters

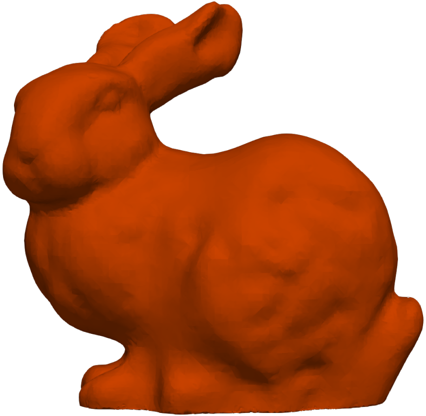




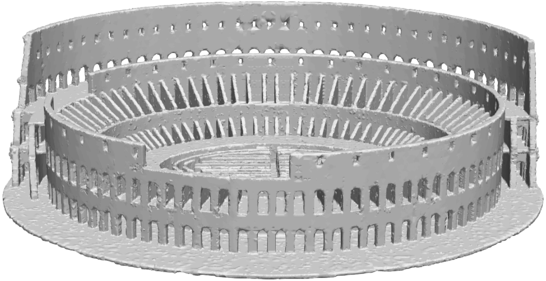
Hyperelastic beam with uncertain Lamé parameters



3D noisy signed distance functions



1 million evaluations



10 million evaluations

5. Conclusions

Conclusions

To sum up

- New adaptive algorithm for level set approximation of noisy functions.
- Cost complexity equivalent to performing uniform refinement directly on the level set, but without knowing where it is a priori!
- Suitable for problems with low-regularity. For smooth problems there is no benefit in space adaptivity and we recommend using a spectral method.

Conclusions

To sum up

- New adaptive algorithm for level set approximation of noisy functions.
- Cost complexity equivalent to performing uniform refinement directly on the level set, but without knowing where it is a priori!
- Suitable for problems with low-regularity. For smooth problems there is no benefit in space adaptivity and we recommend using a spectral method.

Next (current) steps

- Consider level-sets of Hausdorff dimension less that $d-1$; work analysis is exactly the same, the error metric is more tricky.
- Approximation of integrals over the level set.

Thank you for listening!

More info about me and my work at: `croci.github.io`

---

[1] M. Croci, A.-L. Haji-Ali, and I. C. Powell. An adaptive sampling algorithm for level-set approximation, 2025. URL `https://arxiv.org/abs/2509.14896`.

[2] A.-L. Haji-Ali, J. Spence, and A. L. Teckentrup. Adaptive multilevel Monte Carlo for probabilities. *SIAM Journal on Numerical Analysis*, 60(4):2125–2149, 2022.

[3] M. B. Giles and A.-L. Haji-Ali. Multilevel nested simulation for efficient risk estimation. *SIAM/ASA Journal on Uncertainty Quantification*, 7(2):497–525, 2019.

[4] M. Croci, M. B. Giles, M. E. Rognes, and P. E. Farrell. Efficient white noise sampling and coupling for Multilevel Monte Carlo with nonnested meshes. *SIAM/ASA Journal on Uncertainty Quantification*, 6(4): 1630–1655, 2018.

[5] J. P. Boyd. Finding the zeros of a univariate equation: proxy rootfinders, Chebyshev interpolation, and the companion matrix. *SIAM Review*, 55(2):375–396, 2013.

[6] A. Townsend and L. N. Trefethen. An extension of Chebfun to two dimensions. *SIAM Journal on Scientific Computing*, 35(6):C495–C518, 2013.

Acknowledgements

GOBIERNO DE ESPAÑA

MINISTERIO DE CIENCIA E INNOVACIÓN

Cofinanciado por la Unión Europea

AGENCIA ESTATAL DE INVESTIGACIÓN

ikerbasque

Basque Foundation for Science

This talk is supported by grant RYC2022-036312-I funded by MICIU / AEI / 10.13039/501100011033 and by ESF+. Additionally, it is also supported by the Basque Government through the BERC 2022-2025 program and by the Ministry of Science and Innovation through the BCAM Severo Ochoa accreditation CEX2021-001142-S / MICIN / AEI / 10.13039/501100011033.

Error bound on local approximation

Theorem (approximation error)

Let $\hat{f}^{\Box} = I^{\Box}\tilde{f}$ and choose $M(x_e) \sim (\sigma(x_e))^{1/\beta} h^{-\alpha/\beta}$ for all evaluation points $x_e \in \overline{\Box}$. Then, $\hat{f}^{\Box}$ is well defined and satisfies

$$\mathbb{E}\left[\|f - \hat{f}^{\Box}\|^p_{L^p(\Box)}\right]^{1/p} \lesssim \gamma(\Box) h^{\alpha}.$$

Key challenge. $\tilde{f}$ not assumed continuous due to possibly independent estimations.

Theorem proof (sketch)

Banach space notation:

Theorem proof (sketch)

Banach space notation:

- $\mathcal{B}(\overline{\square}, \mathbb{R}) :=$ space of everywhere-bounded functions over $\overline{\square}$. Norm: $\max_{x\in\overline{\square}} |\cdot|$.

Theorem proof (sketch)

Banach space notation:

- $\mathcal{B}(\overline{\square}, \mathbb{R}) :=$ space of everywhere-bounded functions over $\overline{\square}$. Norm: $\max_{x \in \overline{\square}} |\cdot|$.
- $\mathcal{B}(\overline{\square}, L^p(\Omega)) :=$ space of random fields with everywhere (in space) bounded p-th moment (in probability). Norm: $\max_{x \in \overline{\square}} \mathbb{E}[|\cdot|^p]^{1/p}$. **Note:** $\tilde{f} \in \mathcal{B}(\overline{\square}, L^p(\Omega))$.

Theorem proof (sketch)

Banach space notation:

- $\mathcal{B}(\overline{\square}, \mathbb{R}) :=$ space of everywhere-bounded functions over $\overline{\square}$. Norm: $\max_{x\in\overline{\square}} |\cdot|$.
- $\mathcal{B}(\overline{\square}, L^p(\Omega)) :=$ space of random fields with everywhere (in space) bounded p-th moment (in probability). Norm: $\max_{x\in\overline{\square}} \mathbb{E}[|\cdot|^p]^{1/p}$. **Note:** $\tilde{f} \in \mathcal{B}(\overline{\square}, L^p(\Omega))$.

Proof sketch: The first step is to show that any Lipschitz $I^{\square} : \mathcal{B}(\overline{\square}, \mathbb{R}) \to C^0(\overline{\square})$ can be extended to a Lipschitz operator $I^{\square} : \mathcal{B}(\overline{\square}, L^p(\Omega)) \to L^p(\Omega, C^0(\overline{\square}))$. Then,

Theorem proof (sketch)

Banach space notation:

- $\mathcal{B}(\overline{\square}, \mathbb{R}) :=$ space of everywhere-bounded functions over $\overline{\square}$. Norm: $\max_{x\in\overline{\square}} |\cdot|$.
- $\mathcal{B}(\overline{\square}, L^p(\Omega)) :=$ space of random fields with everywhere (in space) bounded p-th moment (in probability). Norm: $\max_{x\in\overline{\square}} \mathbb{E}[|\cdot|^p]^{1/p}$. **Note:** $\tilde{f} \in \mathcal{B}(\overline{\square}, L^p(\Omega))$.

Proof sketch: The first step is to show that any Lipschitz $I^{\square} : \mathcal{B}(\overline{\square}, \mathbb{R}) \to C^0(\overline{\square})$ can be extended to a Lipschitz operator $I^{\square} : \mathcal{B}(\overline{\square}, L^p(\Omega)) \to L^p(\Omega, C^0(\overline{\square}))$. Then,

$$\|f - \hat{f}^{\square}\|_{L^p(\Omega, L^p(\square))} \leq \|f - I^{\square} f\|_{L^p(\Omega, L^p(\square))} + \|I^{\square}(f - \tilde{f})\|_{L^p(\Omega, L^p(\square))}$$

Theorem proof (sketch)

Banach space notation:

- $\mathcal{B}(\overline{\Box},\mathbb{R}) :=$ space of everywhere-bounded functions over $\overline{\Box}$. Norm: $\max_{x\in\overline{\Box}}|\cdot|$.
- $\mathcal{B}(\overline{\Box},L^p(\Omega)) :=$ space of random fields with everywhere (in space) bounded p-th moment (in probability). Norm: $\max_{x\in\overline{\Box}}\mathbb{E}[|\cdot|^p]^{1/p}$. **Note:** $\tilde{f}\in\mathcal{B}(\overline{\Box},L^p(\Omega))$.

Proof sketch: The first step is to show that any Lipschitz $I^{\Box}:\mathcal{B}(\overline{\Box},\mathbb{R})\to C^0(\overline{\Box})$ can be extended to a Lipschitz operator $I^{\Box}:\mathcal{B}(\overline{\Box},L^p(\Omega))\to L^p(\Omega,C^0(\overline{\Box}))$. Then,

$$\begin{aligned}\|f-\hat{f}^{\Box}\|_{L^p(\Omega,L^p(\Box))} &\leq \|f-I^{\Box}f\|_{L^p(\Omega,L^p(\Box))}+\|I^{\Box}(f-\tilde{f})\|_{L^p(\Omega,L^p(\Box))}\\ &\leq \|f-I^{\Box}f\|_{L^p(\Omega,L^p(\Box))}+|\Box|^{1/p}\|I^{\Box}(f-\tilde{f})\|_{L^p(\Omega,C^0(\overline{\Box}))}\end{aligned}$$

Theorem proof (sketch)

Banach space notation:

- $\mathcal{B}(\overline{\square}, \mathbb{R}) :=$ space of everywhere-bounded functions over $\overline{\square}$. Norm: $\max_{x\in\overline{\square}} |\cdot|$.
- $\mathcal{B}(\overline{\square}, L^p(\Omega)) :=$ space of random fields with everywhere (in space) bounded p-th moment (in probability). Norm: $\max_{x\in\overline{\square}} \mathbb{E}[|\cdot|^p]^{1/p}$. **Note:** $\tilde{f} \in \mathcal{B}(\overline{\square}, L^p(\Omega))$.

Proof sketch: The first step is to show that any Lipschitz $I^{\square} : \mathcal{B}(\overline{\square}, \mathbb{R}) \to C^0(\overline{\square})$ can be extended to a Lipschitz operator $I^{\square} : \mathcal{B}(\overline{\square}, L^p(\Omega)) \to L^p(\Omega, C^0(\overline{\square}))$. Then,

$$\begin{aligned}
\|f - \hat{f}^{\square}\|_{L^p(\Omega, L^p(\square))} &\leq \|f - I^{\square} f\|_{L^p(\Omega, L^p(\square))} + \|I^{\square}(f - \tilde{f})\|_{L^p(\Omega, L^p(\square))} \\
&\leq \|f - I^{\square} f\|_{L^p(\Omega, L^p(\square))} + |\square|^{1/p} \|I^{\square}(f - \tilde{f})\|_{L^p(\Omega, C^0(\overline{\square}))} \\
&\leq \|f - I^{\square} f\|_{L^p(\Omega, L^p(\square))} + |\square|^{1/p} L_{I^{\square}} \max_{x_e \in \overline{\square}} \mathbb{E}[|f(x_e) - \tilde{f}(x_e, \cdot)|^p]^{1/p}
\end{aligned}$$

Theorem proof (sketch)

Banach space notation:

- $\mathcal{B}(\overline{\square}, \mathbb{R}) :=$ space of everywhere-bounded functions over $\overline{\square}$. Norm: $\max_{x \in \overline{\square}} |\cdot|$.
- $\mathcal{B}(\overline{\square}, L^p(\Omega)) :=$ space of random fields with everywhere (in space) bounded p-th moment (in probability). Norm: $\max_{x \in \overline{\square}} \mathbb{E}[|\cdot|^p]^{1/p}$. **Note:** $\tilde{f} \in \mathcal{B}(\overline{\square}, L^p(\Omega))$.

Proof sketch: The first step is to show that any Lipschitz $I^{\square} : \mathcal{B}(\overline{\square}, \mathbb{R}) \to C^0(\overline{\square})$ can be extended to a Lipschitz operator $I^{\square} : \mathcal{B}(\overline{\square}, L^p(\Omega)) \to L^p(\Omega, C^0(\overline{\square}))$. Then,

$$\begin{aligned}
\|f - \hat{f}^{\square}\|_{L^p(\Omega, L^p(\square))} &\leq \|f - I^{\square} f\|_{L^p(\Omega, L^p(\square))} + \|I^{\square}(f - \tilde{f})\|_{L^p(\Omega, L^p(\square))} \\
&\leq \|f - I^{\square} f\|_{L^p(\Omega, L^p(\square))} + |\square|^{1/p} \|I^{\square}(f - \tilde{f})\|_{L^p(\Omega, C^0(\overline{\square}))} \\
&\leq \|f - I^{\square} f\|_{L^p(\Omega, L^p(\square))} + |\square|^{1/p} L_{I^{\square}} \max_{x_e \in \overline{\square}} \mathbb{E}[|f(x_e) - \tilde{f}(x_e, \cdot)|^p]^{1/p} \\
&\lesssim \gamma(\square) h^{\alpha} + |\square|^{1/p} \max_{x_e \in \overline{\square}} \sigma(x_e) M^{-\beta}(x_e) \lesssim \gamma(\square) h^{\alpha}.
\end{aligned}$$

Error measure and goal-oriented refinement criterion

Error measure: $E(D) = \sum_{\square} E(\square)$, where $E(\square) = \int_{\square} \mathbb{E}\left[\left|\mathbb{I}_{f\leq 0}(x) - \mathbb{I}_{\hat{f}\leq 0}(x)\right|\right] \mathrm{d}x.$

Theorem (Goal-oriented refinement criterion)

Let $0 \leq s = (p+1)^{-1/2} \leq 2^{-1/2} < 1$, and let $\hat{e}(\square)$ be any quantity (an a priori/posteriori error estimator) such that:

$$\mathbb{E}\left[\|f - \hat{f}^{\square}\|^{p}_{L^p(\square)}\right]^{1/p} \leq \hat{e}(\square) \leq \gamma(\square) h^{\alpha}.$$

For all $\varepsilon \in (0,1)$, let $h_L(\varepsilon) = \varepsilon^{\frac{p+1}{\alpha p}}$ and $h_1(\varepsilon) = (h_L(\varepsilon))^s$. Then, $0 < h_L < h_1$ and

$$\text{IF} \quad h \leq h_L \quad \text{OR} \quad \left(h \leq h_1(\varepsilon) \quad \text{AND} \quad \hat{e}(\square)^{1-s} \leq |\square|^{\frac{1-s}{p}} \inf_{x\in\square} |\hat{f}^{\square}(x,\omega)| \right),$$

THEN $E(\square) \leq \gamma_E^p(\square)\varepsilon$.

Proof (step 1)

First, we prove that $\exists h_L(\varepsilon)$ s.t. if $h \leq h_L$, then $E(\square) \leq \gamma^p(\square)\varepsilon$.

Proof (step 1)

First, we prove that $\exists h_L(\varepsilon)$ s.t. if $h \leq h_L$, then $E(\square) \leq \gamma^p(\square)\varepsilon$.

For any $\psi > 0$, it holds,

$$\mathbb{E}\left[\left|\mathbb{I}_{f\leq 0} - \mathbb{I}_{\hat{f}^{\square}\leq 0}\right|\right] \leq \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\geq|f|}\right]$$

Proof (step 1)

First, we prove that $\exists h_L(\varepsilon)$ s.t. if $h \leq h_L$, then $E(\Box) \leq \gamma^p(\Box)\varepsilon$.

For any $\psi > 0$, it holds,

$$\mathbb{E}\left[\left|\mathbb{I}_{f\leq 0} - \mathbb{I}_{\hat{f}^{\Box}\leq 0}\right|\right] \leq \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\Box}|\geq|f|}\right] = \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\Box}|\geq|f|}\mathbb{I}_{|f|<\psi}\right] + \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\Box}|\geq|f|}\mathbb{I}_{|f|\geq\psi}\right]$$

Proof (step 1)

First, we prove that $\exists h_L(\varepsilon)$ s.t. if $h \leq h_L$, then $E(\Box) \leq \gamma^p(\Box)\varepsilon$.

For any $\psi > 0$, it holds,

$$\begin{aligned}\mathbb{E}\left[\left|\mathbb{I}_{f\leq 0} - \mathbb{I}_{\hat{f}^{\Box}\leq 0}\right|\right] &\leq \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\Box}|\geq|f|}\right] = \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\Box}|\geq|f|}\mathbb{I}_{|f|<\psi}\right] + \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\Box}|\geq|f|}\mathbb{I}_{|f|\geq\psi}\right] \\ &\leq \mathbb{I}_{|f|<\psi} + \mathbb{P}(|f - \hat{f}^{\Box}| \geq \psi)\end{aligned}$$

Proof (step 1)

First, we prove that $\exists h_L(\varepsilon)$ s.t. if $h \le h_L$, then $E(\square) \le \gamma^p(\square)\varepsilon$.

For any $\psi > 0$, it holds,

$$\begin{aligned}\mathbb{E}\left[\left|\mathbb{I}_{f\le 0} - \mathbb{I}_{\hat{f}^{\square}\le 0}\right|\right] \le \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\ge|f|}\right] &= \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\ge|f|}\mathbb{I}_{|f|<\psi}\right] + \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\ge|f|}\mathbb{I}_{|f|\ge\psi}\right]\\ &\le \mathbb{I}_{|f|<\psi} + \mathbb{P}(|f-\hat{f}^{\square}| \ge \psi) \le \mathbb{I}_{|f|<\psi} + \psi^{-p}\mathbb{E}[|f-\hat{f}^{\square}|^p].\end{aligned}$$

Proof (step 1)

First, we prove that $\exists h_L(\varepsilon)$ s.t. if $h \leq h_L$, then $E(\square) \leq \gamma^p(\square)\varepsilon$.

For any $\psi > 0$, it holds,

$$\begin{aligned}\mathbb{E}\left[\left|\mathbb{I}_{f\leq 0} - \mathbb{I}_{\hat{f}^{\square}\leq 0}\right|\right] &\leq \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\geq|f|}\right] = \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\geq|f|}\mathbb{I}_{|f|<\psi}\right] + \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\geq|f|}\mathbb{I}_{|f|\geq\psi}\right]\\ &\leq \mathbb{I}_{|f|<\psi} + \mathbb{P}(|f-\hat{f}^{\square}|\geq\psi) \leq \mathbb{I}_{|f|<\psi} + \psi^{-p}\mathbb{E}[|f-\hat{f}^{\square}|^p].\end{aligned}$$

After integrating over $\square$ and invoking the approximation error bound from earlier we get

$$E(\square) \leq (\psi + \psi^{-p}h^{\alpha p})\left(\psi^{-1}\int_{\square}\mathbb{I}_{|f|<\psi}(x)\,\mathrm{d}x + \gamma^p(\square)\right)$$

Proof (step 1)

First, we prove that $\exists h_L(\varepsilon)$ s.t. if $h \leq h_L$, then $E(\square) \leq \gamma^p(\square)\varepsilon$.

For any $\psi > 0$, it holds,

$$\begin{aligned}\mathbb{E}\left[\left|\mathbb{I}_{f\leq 0} - \mathbb{I}_{\hat{f}^{\square}\leq 0}\right|\right] &\leq \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\geq|f|}\right] = \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\geq|f|}\mathbb{I}_{|f|<\psi}\right] + \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\geq|f|}\mathbb{I}_{|f|\geq\psi}\right]\\ &\leq \mathbb{I}_{|f|<\psi} + \mathbb{P}(|f-\hat{f}^{\square}|\geq\psi) \leq \mathbb{I}_{|f|<\psi} + \psi^{-p}\mathbb{E}[|f-\hat{f}^{\square}|^p].\end{aligned}$$

After integrating over $\square$ and invoking the approximation error bound from earlier we get

$$E(\square) \leq (\psi + \psi^{-p}h^{\alpha p})\left(\psi^{-1}\int_{\square}\mathbb{I}_{|f|<\psi}(x)\,\mathrm{d}x + \gamma^p(\square)\right) = \gamma^p(\square)(\psi + \psi^{-p}h^{\alpha p}).$$

Proof (step 1)

First, we prove that $\exists h_L(\varepsilon)$ s.t. if $h \le h_L$, then $E(\square) \le \gamma^p(\square)\varepsilon$.

For any $\psi > 0$, it holds,

$$\begin{aligned}\mathbb{E}\left[\left|\mathbb{I}_{f\le 0} - \mathbb{I}_{\hat{f}^\square \le 0}\right|\right] &\le \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^\square|\ge|f|}\right] = \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^\square|\ge|f|}\mathbb{I}_{|f|<\psi}\right] + \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^\square|\ge|f|}\mathbb{I}_{|f|\ge\psi}\right] \\ &\le \mathbb{I}_{|f|<\psi} + \mathbb{P}(|f-\hat{f}^\square| \ge \psi) \le \mathbb{I}_{|f|<\psi} + \psi^{-p}\mathbb{E}[|f-\hat{f}^\square|^p].\end{aligned}$$

After integrating over $\square$ and invoking the approximation error bound from earlier we get

$$E(\square) \le (\psi + \psi^{-p}h^{\alpha p})\left(\psi^{-1}\int_\square \mathbb{I}_{|f|<\psi}(x)\,\mathrm{d}x + \gamma^p(\square)\right) = \gamma^p(\square)(\psi + \psi^{-p}h^{\alpha p}).$$

Choosing $\psi \lesssim h^{\frac{\alpha p}{p+1}}$ to balance the two terms yields $E(\square) \le \gamma^p(\square)h^{\frac{\alpha p}{p+1}}$. Taking $h \le h_L(\varepsilon) = \varepsilon^{\frac{p+1}{\alpha p}}$ then yields the thesis.

Proof (step 2)

We now prove that $\exists h_1 > h_L$ s.t. if $h \leq h_1$ and $\hat{e}^{1-s}(\square) \leq |\square|^{\frac{1-s}{p}} \inf_{x\in\overline{\square}} |\hat{f}^{\square}|$, then $E(\square) \leq \gamma^p(\square)\varepsilon$.

Proof (step 2)

We now prove that $\exists h_1 > h_L$ s.t. if $h \leq h_1$ and $\hat{e}^{1-s}(\Box) \leq |\Box|^{\frac{1-s}{p}} \inf_{x \in \overline{\Box}} |\hat{f}^{\Box}|$, then $E(\Box) \leq \gamma^p(\Box)\varepsilon$. Under these assumptions, it holds

Proof (step 2)

We now prove that $\exists h_1 > h_L$ s.t. if $h \leq h_1$ and $\hat{e}^{1-s}(\square) \leq |\square|^{\frac{1-s}{p}} \inf_{x \in \overline{\square}} |\hat{f}^{\square}|$, then $E(\square) \leq \gamma^p(\square)\varepsilon$. Under these assumptions, it holds

$$\mathbb{E}\left[\left|\mathbb{I}_{f\leq 0} - \mathbb{I}_{\hat{f}^{\square}\leq 0}\right|\right] \leq \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\geq|\hat{f}^{\square}|}\right]$$

Proof (step 2)

We now prove that $\exists h_1 > h_L$ s.t. if $h \leq h_1$ and $\hat{e}^{1-s}(\square) \leq |\square|^{\frac{1-s}{p}} \inf_{x \in \overline{\square}} |\hat{f}^{\square}|$, then $E(\square) \leq \gamma^p(\square)\varepsilon$. Under these assumptions, it holds

$$\mathbb{E}\left[\left|\mathbb{I}_{f\leq 0} - \mathbb{I}_{\hat{f}^{\square}\leq 0}\right|\right] \leq \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\geq|\hat{f}^{\square}|}\right] \leq \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\geq|\square|^{-\frac{1-s}{p}}\hat{e}^{1-s}(\square)}\right]$$

Proof (step 2)

We now prove that $\exists h_1 > h_L$ s.t. if $h \leq h_1$ and $\hat{e}^{1-s}(\Box) \leq |\Box|^{\frac{1-s}{p}} \inf_{x \in \overline{\Box}} |\hat{f}^{\Box}|$, then $E(\Box) \leq \gamma^p(\Box)\varepsilon$. Under these assumptions, it holds

$$\begin{aligned}\mathbb{E}\left[\left|\mathbb{I}_{f\leq 0} - \mathbb{I}_{\hat{f}^{\Box}\leq 0}\right|\right] \leq \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\Box}|\geq|\hat{f}^{\Box}|}\right] &\leq \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\Box}|\geq|\Box|^{-\frac{1-s}{p}}\hat{e}^{1-s}(\Box)}\right] \\ &= \mathbb{P}\left(|f-\hat{f}^{\Box}| \geq |\Box|^{-\frac{1-s}{p}}\hat{e}^{1-s}(\Box)\right)\end{aligned}$$

Proof (step 2)

We now prove that $\exists h_1 > h_L$ s.t. if $h \leq h_1$ and $\hat{e}^{1-s}(\Box) \leq |\Box|^{\frac{1-s}{p}} \inf_{x\in\overline{\Box}} |\hat{f}^{\Box}|$, then $E(\Box) \leq \gamma^p(\Box)\varepsilon$. Under these assumptions, it holds

$$
\begin{aligned}
\mathbb{E}\left[\left|\mathbb{I}_{f\leq 0} - \mathbb{I}_{\hat{f}^{\Box}\leq 0}\right|\right] \leq \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\Box}|\geq|\hat{f}^{\Box}|}\right] &\leq \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\Box}|\geq|\Box|^{-\frac{1-s}{p}}\hat{e}^{1-s}(\Box)}\right] \\
&= \mathbb{P}\left(|f-\hat{f}^{\Box}| \geq |\Box|^{-\frac{1-s}{p}}\hat{e}^{1-s}(\Box)\right) \\
&\leq \left(|\Box|\hat{e}^{-p}(\Box)\right)^{1-s} \mathbb{E}[|f-\hat{f}^{\Box}|^p].
\end{aligned}
$$

Proof (step 2)

We now prove that $\exists h_1 > h_L$ s.t. if $h \le h_1$ and $\hat{e}^{1-s}(\square) \le |\square|^{\frac{1-s}{p}} \inf_{x\in\overline{\square}} |\hat{f}^{\square}|$, then $E(\square) \le \gamma^p(\square)\varepsilon$. Under these assumptions, it holds

$$\begin{aligned}
\mathbb{E}\left[\left|\mathbb{I}_{f\le 0} - \mathbb{I}_{\hat{f}^{\square}\le 0}\right|\right] \le \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\ge|\hat{f}^{\square}|}\right] &\le \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\ge|\square|^{-\frac{1-s}{p}}\hat{e}^{1-s}(\square)}\right] \\
&= \mathbb{P}\left(|f-\hat{f}^{\square}| \ge |\square|^{-\frac{1-s}{p}}\hat{e}^{1-s}(\square)\right) \\
&\le \left(|\square|\hat{e}^{-p}(\square)\right)^{1-s} \mathbb{E}[|f-\hat{f}^{\square}|^p].
\end{aligned}$$

Integrating over $\square$ and using that $\hat{e}(\square)$ upper bounds the error yields

$$E(\square) \le \left(|\square|\hat{e}^{-p}(\square)\right)^{1-s}\mathbb{E}\left[\|f-\hat{f}^{\square}\|^p_{L^p(\square)}\right]$$

Proof (step 2)

We now prove that $\exists h_1 > h_L$ s.t. if $h \le h_1$ and $\hat{e}^{1-s}(\Box) \le |\Box|^{\frac{1-s}{p}} \inf_{x \in \overline{\Box}} |\hat{f}^{\Box}|$, then $E(\Box) \le \gamma^p(\Box)\varepsilon$. Under these assumptions, it holds

$$\begin{aligned}
\mathbb{E}\left[\left|\mathbb{I}_{f \le 0} - \mathbb{I}_{\hat{f}^{\Box} \le 0}\right|\right] \le \mathbb{E}\left[\mathbb{I}_{|f - \hat{f}^{\Box}| \ge |\hat{f}^{\Box}|}\right] &\le \mathbb{E}\left[\mathbb{I}_{|f - \hat{f}^{\Box}| \ge |\Box|^{-\frac{1-s}{p}} \hat{e}^{1-s}(\Box)}\right] \\
&= \mathbb{P}\left(|f - \hat{f}^{\Box}| \ge |\Box|^{-\frac{1-s}{p}} \hat{e}^{1-s}(\Box)\right) \\
&\le \left(|\Box| \hat{e}^{-p}(\Box)\right)^{1-s} \mathbb{E}[|f - \hat{f}^{\Box}|^p].
\end{aligned}$$

Integrating over $\Box$ and using that $\hat{e}(\Box)$ upper bounds the error yields

$$E(\Box) \le \left(|\Box| \hat{e}^{-p}(\Box)\right)^{1-s} \mathbb{E}\left[\|f - \hat{f}^{\Box}\|_{L^p(\Box)}^p\right] \le |\Box|^{1-s} (\hat{e}(\Box))^{sp} \le \gamma^p(\Box) h^{\alpha s p},$$

since $|\Box| \sim \gamma^p(\Box)$ and $\hat{e}(\Box) \le \gamma(\Box) h^{\alpha}$.

Proof (step 2)

We now prove that $\exists h_1 > h_L$ s.t. if $h \le h_1$ and $\hat{e}^{1-s}(\square) \le |\square|^{\frac{1-s}{p}} \inf_{x\in\overline{\square}} |\hat{f}^{\square}|$, then $E(\square) \le \gamma^p(\square)\varepsilon$. Under these assumptions, it holds

$$\begin{aligned}
\mathbb{E}\left[\left|\mathbb{I}_{f\le 0} - \mathbb{I}_{\hat{f}^{\square}\le 0}\right|\right] \le \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\ge|\hat{f}^{\square}|}\right] &\le \mathbb{E}\left[\mathbb{I}_{|f-\hat{f}^{\square}|\ge|\square|^{-\frac{1-s}{p}}\hat{e}^{1-s}(\square)}\right] \\
&= \mathbb{P}\left(|f-\hat{f}^{\square}| \ge |\square|^{-\frac{1-s}{p}}\hat{e}^{1-s}(\square)\right) \\
&\le (|\square|\hat{e}^{-p}(\square))^{1-s}\, \mathbb{E}[|f-\hat{f}^{\square}|^p].
\end{aligned}$$

Integrating over $\square$ and using that $\hat{e}(\square)$ upper bounds the error yields

$$E(\square) \le (|\square|\hat{e}^{-p}(\square))^{1-s}\mathbb{E}\left[\|f-\hat{f}^{\square}\|^p_{L^p(\square)}\right] \le |\square|^{1-s}(\hat{e}(\square))^{sp} \le \gamma^p(\square)h^{\alpha s p},$$

since $|\square| \sim \gamma^p(\square)$ and $\hat{e}(\square) \le \gamma(\square)h^{\alpha}$.Taking $h \le h_1(\varepsilon) = \varepsilon^{\frac{1}{\alpha s p}} = h_L^s$ yields $E(\square) \le \gamma^p(\square)\varepsilon$ and the thesis since $h_1 > h_L$ for all $\varepsilon < 1$.

Algorithm cost

We now indicate with $\ell \in \{1, \dots, L\}$ a cell refinement level. The (random) work of our method on a cell $\square_\ell$ is given by the recursive formula

$$W_\ell^{\square_\ell} \lesssim N h_\ell^{-\alpha/\beta} \mathbb{I}_{e_\ell \leq \delta_\ell} + \mathbb{I}_{e_\ell > \delta_\ell} \sum_{\square_{\ell+1} \in \mathcal{R}(\square_\ell)} W_{\ell+1}^{\square_{\ell+1}}$$

where N is the number of evaluation nodes per cell, $e_\ell = \hat{e}(\square_\ell)^{1-s}$, $\delta_\ell(\square) = |\square_\ell|^{(1-s)/p} \inf_{x \in \overline{\square}_\ell} |\hat{f}^{\square_\ell}|$, and $\mathcal{R}(\square_\ell)$ is a uniform refinement of $\square_\ell$.

Algorithm cost

We now indicate with $\ell \in \{1, \dots, L\}$ a cell refinement level. The (random) work of our method on a cell $\Box_\ell$ is given by the recursive formula

$$W_\ell^{\Box_\ell} \lesssim N h_\ell^{-\alpha/\beta} \mathbb{I}_{e_\ell \le \delta_\ell} + \mathbb{I}_{e_\ell > \delta_\ell} \sum_{\Box_{\ell+1} \in \mathcal{R}(\Box_\ell)} W_{\ell+1}^{\Box_{\ell+1}}$$

where N is the number of evaluation nodes per cell, $e_\ell = \hat{e}(\Box_\ell)^{1-s}$,
$\delta_\ell(\Box) = |\Box_\ell|^{(1-s)/p} \inf_{x \in \overline{\Box}_\ell} |\hat{f}^{\Box_\ell}|$, and $\mathcal{R}(\Box_\ell)$ is a uniform refinement of $\Box_\ell$.

The total (random) work of our method is then bounded by

$$\sum_{\Box \in D_{h_1}} W_1^{\Box} \lesssim N h_1^{-(\frac{\alpha}{\beta}+d)} + N \sum_{\ell=1}^{L} h_\ell^{-\frac{\alpha}{\beta}} m_\ell, \quad \text{where} \quad m_\ell = \sum_{\Box_\ell \in D_{h_\ell}} \mathbb{I}_{e_\ell > \delta_\ell}$$

is the (random) number of cells on level ℓ that have been refined.

Bounding the expected work

The expected work is then bounded by

$$\sum_{\square\in D_{h_1}} \mathbb{E}[W_1^{\square}] \lesssim N h_1^{-\frac{\alpha}{\beta}-d} + N \sum_{\ell=1}^{L} h_\ell^{-\frac{\alpha}{\beta}} \mathbb{E}[m_\ell].$$

Bounding the expected work

The expected work is then bounded by

$$\sum_{\square\in D_{h_1}} \mathbb{E}[W_1^{\square}] \lesssim N h_1^{-\frac{\alpha}{\beta}-d} + N \sum_{\ell=1}^{L} h_\ell^{-\frac{\alpha}{\beta}} \mathbb{E}[m_\ell].$$

By using our assumptions on the regularity of the level set, we can prove the bound

$$\mathbb{E}[m_\ell] \lesssim h_\ell^{1-d},$$

Bounding the expected work

The expected work is then bounded by

$$\sum_{\square \in D_{h_1}} \mathbb{E}[W_1^{\square}] \lesssim N h_1^{-\frac{\alpha}{\beta}-d} + N \sum_{\ell=1}^{L} h_\ell^{-\frac{\alpha}{\beta}} \mathbb{E}[m_\ell].$$

By using our assumptions on the regularity of the level set, we can prove the bound

$$\mathbb{E}[m_\ell] \lesssim h_\ell^{1-d},$$

which in turn yields a total expected work bounded by

$$\sum_{\square \in D_{h_1}} \mathbb{E}[W_1^{\square}] \lesssim N \left(h_1^{-\frac{\alpha}{\beta}-d} + h_L^{-\frac{\alpha}{\beta}-d+1} \right) \lesssim N h_L^{-\frac{\alpha}{\beta}-d+1}.$$

Total cost complexity

Recall that ensuring an error bound of $\theta\varepsilon$ requires $h_L = \varepsilon^{(p+1)/(\alpha p)}$. The asymptotic cost complexity of our algorithm is then given by

$$\sum_{\square \in D_{h_1}} \mathbb{E}[W_1^{\square}] \lesssim N h_L^{-\frac{\alpha}{\beta}-d+1} \lesssim \varepsilon^{-\left(\frac{p+1}{p}\right)\left(\frac{1}{\beta}+\frac{d-1}{\alpha}\right)}.$$

This results in a cost-complexity reduction of $\mathcal{O}\left(\varepsilon^{\left(\frac{p+1}{\alpha p}\right)}\right)$ wrt a non-adaptive scheme.

Total cost complexity

Recall that ensuring an error bound of $\theta\varepsilon$ requires $h_L = \varepsilon^{(p+1)/(\alpha p)}$. The asymptotic cost complexity of our algorithm is then given by

$$\sum_{\square \in D_{h_1}} \mathbb{E}[W_1^{\square}] \lesssim N h_L^{-\frac{\alpha}{\beta}-d+1} \lesssim \varepsilon^{-\left(\frac{p+1}{p}\right)\left(\frac{1}{\beta}+\frac{d-1}{\alpha}\right)}.$$

This results in a cost-complexity reduction of $\mathcal{O}\left(\varepsilon^{\left(\frac{p+1}{\alpha p}\right)}\right)$ wrt a non-adaptive scheme.

Same complexity as doing uniform refinement directly on the level set, but without knowing where it is beforehand!

Total cost complexity

Recall that ensuring an error bound of $\theta\varepsilon$ requires $h_L = \varepsilon^{(p+1)/(\alpha p)}$. The asymptotic cost complexity of our algorithm is then given by

$$\sum_{\square \in D_{h_1}} \mathbb{E}[W_1^{\square}] \lesssim N h_L^{-\frac{\alpha}{\beta}-d+1} \lesssim \varepsilon^{-\left(\frac{p+1}{p}\right)\left(\frac{1}{\beta}+\frac{d-1}{\alpha}\right)}.$$

This results in a cost-complexity reduction of $\mathcal{O}\left(\varepsilon^{\left(\frac{p+1}{\alpha p}\right)}\right)$ wrt a non-adaptive scheme.

Same complexity as doing uniform refinement directly on the level set, but without knowing where it is beforehand!

Conjecture: f lacking smoothness away from the level set does not affect these results.