Les Cahiers du GERAD ISSN: 0711–2440

Using post-optimality analysis in blackbox optimization to improve design robustness

M. Fahed, S. Alarie, M. Gagnon, M. Kokkolaras

G–2026–79

October 2026

La collection *Les Cahiers du GERAD* est constituée des travaux de recherche menés par nos membres. La plupart de ces documents de travail a été soumis à des revues avec comité de révision. Lorsqu'un document est accepté et publié, le pdf original est retiré si c'est nécessaire et un lien vers l'article publié est ajouté.

Citation suggérée : M. Fahed, S. Alarie, M. Gagnon, M. Kokkolaras (Octobre 2026). Using post-optimality analysis in blackbox optimization to improve design robustness, Rapport technique, Les Cahiers du GERAD G– 2026–79, GERAD, HEC Montréal, Canada.

Avant de citer ce rapport technique, veuillez visiter notre site Web (https://www.gerad.ca/fr/papers/G-2026-79) afin de mettre à jour vos données de référence, s'il a été publié dans une revue scientifique.

The series *Les Cahiers du GERAD* consists of working papers carried out by our members. Most of these pre-prints have been submitted to peer-reviewed journals. When accepted and published, if necessary, the original pdf is removed and a link to the published article is added.

Suggested citation: M. Fahed, S. Alarie, M. Gagnon, M. Kokkolaras (October 2026). Using post-optimality analysis in blackbox optimization to improve design robustness, Technical report, Les Cahiers du GERAD G–2026–79, GERAD, HEC Montréal, Canada.

Before citing this technical report, please visit our website (https://www.gerad.ca/en/papers/G-2026-79) to update your reference data, if it has been published in a scientific journal.

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2026
– Bibliothèque et Archives Canada, 2026

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2026
– Library and Archives Canada, 2026

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Using post-optimality analysis in blackbox optimization to improve design robustness

Melad Fahed [a, d]

Stéphane Alarie [b, d]

Martin Gagnon [b]

Michael Kokkolaras [c, d]

[a] *Department of Mechanical Engineering, McGill University, Montréal (Qc), Canada, H3A 0C3*

[b] *Hydro-Québec Research Institute (IREQ), Varenne (Qc), Canada, J3X 1S1*

[c] *School of Mechanical and Automotive Engineering, Clemson University, South Carolina 29607, United States*

[d] *GERAD, Montréal (Qc), Canada, H3T 1J4*

melad.fahed@mail.mcgill.ca
alarie.stephane@hydroquebec.com
gagnon.martin11@hydroquebec.com
mkokkol@clemson.edu

October 2026
Les Cahiers du GERAD
G–2026–79

Copyright © 2026 Fahed, Alarie, Gagnon, Kokkolaras

Les textes publiés dans la série des rapports de recherche *Les Cahiers du GERAD* n'engagent que la responsabilité de leurs auteurs. Les auteurs conservent leur droit d'auteur et leurs droits moraux sur leurs publications et les utilisateurs s'engagent à reconnaître et respecter les exigences légales associées à ces droits. Ainsi, les utilisateurs:

- Peuvent télécharger et imprimer une copie de toute publication du portail public aux fins d'étude ou de recherche privée;
- Ne peuvent pas distribuer le matériel ou l'utiliser pour une activité à but lucratif ou pour un gain commercial;
- Peuvent distribuer gratuitement l'URL identifiant la publication.

Si vous pensez que ce document enfreint le droit d'auteur, contactez-nous en fournissant des détails. Nous supprimerons immédiatement l'accès au travail et enquêterons sur votre demande.

The authors are exclusively responsible for the content of their research papers published in the series *Les Cahiers du GERAD*. Copyright and moral rights for the publications are retained by the authors and the users must commit themselves to recognize and abide the legal requirements associated with these rights. Thus, users:

- May download and print one copy of any publication from the public portal for the purpose of private study or research;
- May not further distribute the material or use it for any profit-making activity or commercial gain;
- May freely distribute the URL identifying the publication.

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Abstract : In model calibration and other ill-posed inverse problems, many parameter sets fit the measurements almost equally well, yet an optimizer reports only one of them. The other evaluations are archived in an evaluation history and left unused. This paper proposes a post-optimality analysis that seeks, from that history alone, without new evaluations, a more robust calibration candidate within a tolerance of the best fit. A candidate is more robust when it stays within that tolerance under a larger change of its variables. The evaluations within the tolerance outline a region, and the candidate is its Chebyshev center, the point farthest from its walls. The move to the candidate is offered at two levels, the second using more of the history. The method is tested on 5 analytical problems and on a Kaplan turbine startup model calibration, where the histories give a point 2.09 times as robust as the solution.

Keywords : Blackbox optimization; post-optimality analysis; robustness; model calibration; derivative-free optimization

Acknowledgements: The authors thank Hydro-Québec for the data of the case study.

1 Introduction

Computational models of machines are calibrated by fitting their parameters to measurements. Such models are often built, calibrated and adapted to develop digital twins. A digital twin is a virtual representation of a physical system, linked to that system by a two-way stream of data (NASEM 2024). The calibrated values then stay in service until new measurements arrive, while the machine they describe keeps changing. An optimizer usually performs the calibration by minimizing the difference between the predictions and the measurements. It drives that difference as low as it can and ignores how far the parameters could change before the fit becomes unacceptable. It can therefore stop next to a steep rise of the objective, where a small change toward the rise makes the fit unacceptable.

Equifinality, in which distinct parameter sets fit the measurements almost equally well, is common in calibration (Beven and Binley 1992; Beven 2006). The optimizer returns one evaluation of lowest objective value, although other evaluations of the run often fit almost as well. The returned variable values may be the coefficients of a model that will be refitted, or a design that will be built with some error in its dimensions. The farther these values lie from the edge of the acceptable set, the larger the change they can take and stay acceptable.

Blackbox optimization treats the objective as a process that returns an output for a given input, with its inner workings unavailable for analysis (Audet and Hare 2017). When the objective is a simulation that is expensive to evaluate and may fail to return a value, direct-search optimization algorithms are a common choice. The mesh adaptive direct search algorithm (MADS) of Audet and Dennis (2006), implemented in NOMAD (Audet et al. 2022), is one such optimization algorithm. Each iteration of MADS has a search step and a poll step. The search may try points anywhere on a grid (the mesh), and the poll tries points on the mesh close to the best evaluation so far. When the poll finds no better point, the mesh is refined, until it reaches a minimum size. This refinement gives MADS its convergence analysis and a stopping criterion. As the mesh refines, NOMAD places its evaluations closer and closer around the solution. The variable values and the objective value of each evaluation can be written to a file called the evaluation history. In practice that file only confirms convergence, although it is all that remains when the simulation is too costly to rerun or no longer accessible.

Post-optimality analysis reuses the output of a finished run to establish how the solution responds to a change in the problem data or in the solution itself. Its classical form, the sensitivity theory of Fiacco (1983), needs derivatives at a confirmed optimum, whereas the history of a blackbox run records no derivative. Robust design instead changes the problem itself to include perturbations of the design. Robust optimization minimizes the worst objective value over those perturbations (Bertsimas, Nohadani, and Teo 2010), and design centering places the design farthest from the edge of the acceptable set (Vidigal and Director 1982). Each needs an optimization run of its own with new evaluations. To check how far the solution stays near-optimal when the problem is perturbed, the post-solve audit proposed by Hu (2026) also runs the solver again.

During a direct search, slopes (simplex gradients) estimated from past evaluations steer the next iterations (Custódio and Vicente 2007). After a single run, the evaluations have been reused for trade-off studies between the objective and a constraint (Audet, Dennis, and Le Digabel 2012) and for surrogate-based sensitivity analysis (Wang and Shoemaker 2014). Identifiability studies instead rerun the optimizer many times to check whether the measurements determine the parameter values (Raue et al. 2013). Methods from hydrology also draw new samples of the parameters and keep the parameter sets that fit the measurements within a tolerance (Spear and Hornberger 1980; Beven and Binley 1992). The evaluation history of a finished run is already such a sample, although its evaluations cluster along the path of the search, mostly near the solution.

To the authors' knowledge, these works stop short of using the history of one finished run to seek a more robust candidate and bound the robustness of the solution. The post-optimality analysis of this paper seeks that candidate and bounds that robustness after the optimization has finished, with no further search and no change to the problem. It builds the smallest convex region that contains

the evaluations within a tolerance of the best objective value (their convex hull). The candidate is the point of that hull farthest from its walls (its Chebyshev center). The move to the candidate is offered at two levels, and the stricter level moves only when the history shows the candidate to be more robust than the solution. This paper makes three contributions to blackbox optimization and model calibration.

1. It proposes a post-optimality analysis that seeks a more robust candidate from the evaluation history alone.
2. It offers the move to the Chebyshev center at two levels and tests both on 50 NOMAD runs of 5 analytical problems with known optima, in 1, 2, 5 and 10 variables.
3. It applies the analysis to the startup model calibration of a Kaplan turbine and tests the Chebyshev center on measured startups held out from the calibration.

2 Problem formulation

Consider the bound-constrained problem

$$\min_{\mathbf{x}\in[\mathbf{l},\mathbf{u}]} f(\mathbf{x}), \qquad \mathbf{l},\mathbf{u}\in\mathbb{R}^n, \tag{1}$$

with n variables, where f is a blackbox model, its derivatives are unavailable and an evaluation may fail. A run records its evaluations as the ordered set

$$\mathcal{H} = \{(\mathbf{x}_k, f_k)\}_{k=1}^{m}, \tag{2}$$

in which each pair is an evaluation, with variable values $\mathbf{x}_k$ and objective value f_k or a failure flag. The set $\mathcal{H}$ is called the evaluation history, and the run returns the successful evaluation of lowest objective value f^*, whose variable values $\mathbf{x}^*$ are called the solution. Variables can have very different ranges, so each variable is first mapped to $[-1, 1]$ between its bounds,

$$z_i = 2\,\frac{x_i - l_i}{u_i - l_i} - 1. \tag{3}$$

The variable values of an evaluation form a point, written $\mathbf{x}$ in the original units and $\mathbf{z}$ in these normalized coordinates, although the solution is written $\mathbf{x}^*$ in both. From here on, distances are measured in these normalized coordinates.

The tolerance ε is the share of the objective value that can be given up for robustness, typically a few percent. An evaluation is acceptable when its objective value exceeds the best value found by at most this share, and rejected otherwise. The same test applied to each $\mathbf{x}$ within the bounds gives the acceptable set

$$V(\varepsilon) = \{\mathbf{x}\in[\mathbf{l},\mathbf{u}] : f(\mathbf{x}) \le f^* + |f^*|\,\varepsilon\}. \tag{4}$$

A failed evaluation is also rejected, on the assumption that the model would fail there again.

A point is more robust when it stays acceptable under a larger change of its variables. The robustness radius $\rho(\mathbf{z})$ of a point $\mathbf{z}$ is the shortest straight-line distance from $\mathbf{z}$ to an unacceptable point $\mathbf{w}$ within the bounds,

$$\rho(\mathbf{z}) = \inf\{\|\mathbf{w}-\mathbf{z}\| : \mathbf{w}\in[-1,1]^n,\ \mathbf{w}\notin V(\varepsilon)\}. \tag{5}$$

Within the bounds a change of the point shorter than the radius keeps it acceptable, whatever its direction. A straight-line distance gives all directions of change the same weight. The normalization of Eq. (3) maps the range of each variable onto a length of 2. A radius r therefore lets a single variable change alone by up to $r\,(u_i - l_i)/2$ in the original units. When several variables change at once, their combined straight-line change must stay below r. Two variables that each change by r move the point

by $\sqrt{2}\,r$, beyond the radius. In the synthetic acceptable set of Fig. 1 the solution $\mathbf{x}^*$ lies near the edge, whereas the candidate ($\mathbf{z}_c$) lies farther inside and has the larger radius.

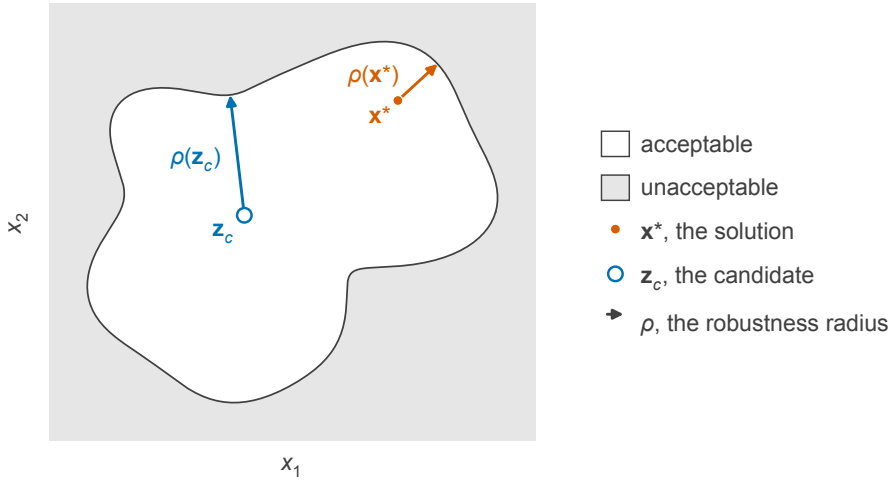

Figure 1: The robustness radius at 2 points of a synthetic acceptable set. Each arrow runs from a point to the nearest unacceptable point, and its length is the radius.

3 Proposed methodology

From the evaluation history, the method decides whether the solution $\mathbf{x}^*$ can be replaced by a more robust point. Each step makes one of several possible choices, and the other choices that could have been made in each step are discussed in Section 6. Fig. 2 shows the six steps on one example of a NOMAD run of the wedge in 2 variables. The acceptable valley of the wedge widens away from the narrow end that contains the solution, so a more robust point is known to exist away from the original $\mathbf{x}^*$ reported by the optimizer.

Step 1. Evaluation history. Eq. (3) maps the history $\mathcal{H}$ and its solution $\mathbf{x}^*$ to $[-1, 1]^n$, as in Fig. 2(a).

Step 2. Acceptable evaluations. Eq. (4) splits the history into acceptable and rejected evaluations, as in Fig. 2(b). Fig. 2 uses a tolerance of $\varepsilon = 5\%$. The threshold of this split is set from the best value in the run, so it depends on the run having converged.

Step 3. Hull construction. The method outlines the acceptable set with convex hulls, whose walls are flat faces. Such a region fits an acceptable set that is close to convex. Near a minimum where a smooth objective has a positive definite Hessian, the acceptable set approaches an ellipsoid as the tolerance shrinks. The flat walls are linear inequalities, so a linear program finds the point farthest from them. The acceptable set can be disconnected, made of separate parts. A hull over two parts would also cover the unacceptable gap between them, so the acceptable evaluations are grouped first. Single-linkage clustering (Sibson 1973) links them one pair at a time, starting with the two closest evaluations. Each new link joins the closest two evaluations that are not yet in the same group, until all evaluations are in one group. Single linkage does not need to be told how many groups to form, and it can follow a long, curved part. Inside one part the links are all short. Two separate parts are linked only across the gap between them, so that link is much longer. The method looks for this jump in the sorted link lengths. In the longer half of the lengths, it finds the largest ratio of one length to the next shorter one. The shorter half is left out, since large ratios also occur among the shortest links inside one part. When the largest ratio reaches the cut factor κ, the links longer than the jump are removed. Each set of evaluations still linked then forms a group. With a ratio below κ, all evaluations stay in one group, as in

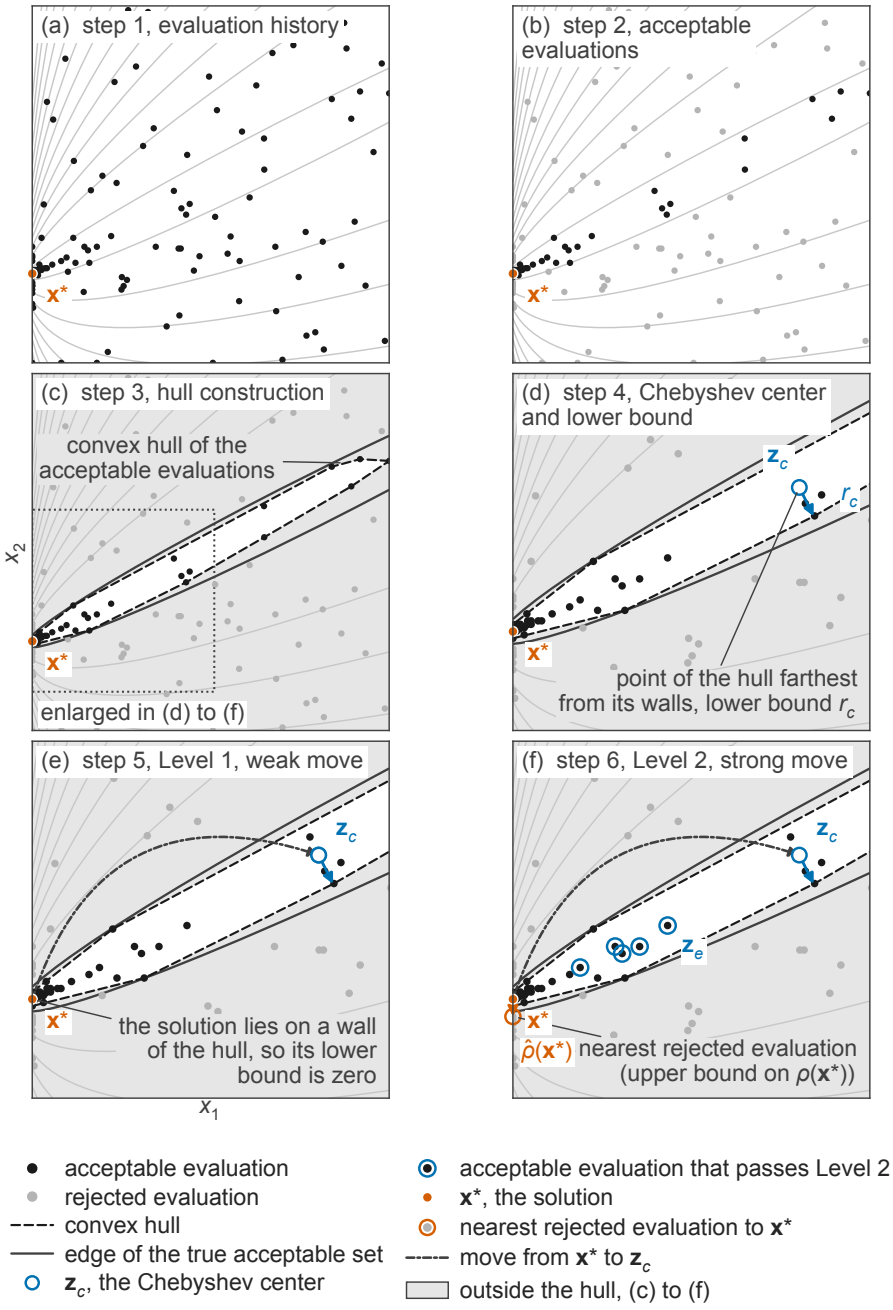

Figure 2: The six steps on one NOMAD run of the wedge in 2 variables, with 314 evaluations of which 236 are acceptable at $\varepsilon = 5\%$. The contours of the objective and the edge of the true acceptable set are drawn for reference only.

Fig. 2(*c*). This paper uses $\kappa = 5$. A lower κ can split a connected part, whereas a higher κ can merge separate parts.

The quickhull algorithm (Barber, Dobkin, and Huhdanpaa 1996) builds the hull of each group that has at least $n+1$ evaluations and spans all n dimensions. Each hull is written as

$$\mathcal{C} = \{\mathbf{z} : \mathbf{a}_i^\mathsf{T}\mathbf{z} \le b_i,\ i = 1, \ldots, q\}, \qquad \|\mathbf{a}_i\|_2 = 1, \tag{6}$$

where q is the number of walls, $\mathbf{a}_i$ is the outward unit vector perpendicular to wall i, and b_i sets the position of that wall. The hull is limited to about 10 variables, because the number of walls grows fast with the number of variables.

The method assumes from here on that all points of the hull are acceptable, $\mathcal{C} \subseteq V(\varepsilon)$. Each hull is tested against the rejected evaluations with Eq. (6) and discarded when one lies inside it. The test detects an unacceptable part of a hull only when a rejected evaluation lies in it.

Step 4. Candidate and lower bound. When the whole hull is acceptable, a point $\mathbf{z}$ inside it stays acceptable under all changes shorter than its distance $r(\mathbf{z})$ to the nearest wall. That distance is therefore a lower bound on the unknown radius of the point. Each vector $\mathbf{a}_i$ has unit length, so

$$r(\mathbf{z}) = \min_i \left(b_i - \mathbf{a}_i^\mathsf{T}\mathbf{z}\right), \tag{7}$$

which is zero on the walls. Any point of the hull could be the candidate. This paper takes the point with the largest lower bound, the Chebyshev center (Boyd and Vandenberghe 2004), which passes Level 2 whenever another point of the hull passes. Solving the linear program in the $n+1$ unknowns $\mathbf{z}$ and r gives the Chebyshev center of each hull,

$$\max_{\mathbf{z},r} \; r \quad \text{subject to} \quad \mathbf{a}_i^\mathsf{T}\mathbf{z} + r \le b_i, \; i = 1, \ldots, q. \tag{8}$$

The candidate $\mathbf{z}_c$ is the Chebyshev center whose bound $r_c = r(\mathbf{z}_c)$ is the largest among the hulls (Fig. 2(d)). Of the acceptable evaluations inside the hulls, the one with the largest bound is the recorded evaluation $\mathbf{z}_e$, and its objective value is known.

Step 5. Level 1, weak move. When all points of the hull are acceptable, the Chebyshev center $\mathbf{z}_c$ and all points within r_c of it lie within the tolerance. Level 1 therefore replaces $\mathbf{x}^*$ by $\mathbf{z}_c$ whenever a hull exists and otherwise refuses to decide (Tab. 1). This decision uses the history alone, whereas 1 evaluation of the center $\mathbf{z}_c$ measures the objective loss of the move (δ),

$$\delta = \frac{f(\mathbf{z}_c) - f^*}{|f^*|}, \tag{9}$$

and $\varepsilon - \delta$ is the remaining tolerance. The depth ratio $r_c/r(\mathbf{x}^*)$, which Level 1 also reports, measures how much deeper the center lies in the hull than the solution. The center maximizes the lower bound, so the ratio is at least 1 even when the solution has the larger radius.

Table 1: The two levels of decision. The claim of Level 2 assumes that all points of the hull are acceptable.

level	evaluations used	moves to $\mathbf{z}_c$ when	claim	reports
1, weak	acceptable, and rejected for the test of step 3	a hull exists	$\mathbf{z}_c$ lies deeper in the hull than $\mathbf{x}^*$	δ, $\varepsilon - \delta$ and $r_c/r(\mathbf{x}^*)$
2, strong	acceptable, and rejected for step 3 and the upper bound	$r_c > \hat{\rho}(\mathbf{x}^*)$	$\rho(\mathbf{z}_c) > \rho(\mathbf{x}^*)$	also $r_c/\hat{\rho}(\mathbf{x}^*)$ and $\mathbf{z}_e$

Step 6. Level 2, strong move. Since a comparison of lower bounds always favors the center, Level 2 needs an upper bound on the radius of the solution. Each rejected evaluation marks an unacceptable point at a known distance from a point $\mathbf{z}$, so Eq. (5) gives

$$\rho(\mathbf{z}) \le \hat{\rho}(\mathbf{z}) = \min_{\mathbf{z}_r \text{ rejected}} \|\mathbf{z}_r - \mathbf{z}\| \tag{10}$$

whatever the shape of the acceptable set. The outcome of Level 2 is then one of three,

$$\text{outcome} = \begin{cases} \text{move to } \mathbf{z}_c, & r_c > \hat{\rho}(\mathbf{x}^*), \\ \text{stay at } \mathbf{x}^*, & r_c \le \hat{\rho}(\mathbf{x}^*), \\ \text{refuse}, & \text{step 3 leaves no hull.} \end{cases} \tag{11}$$

A move is justified when all points of the hull are acceptable, since then

$$\rho(\mathbf{z}_c) \;\ge\; r_c \;>\; \hat{\rho}(\mathbf{x}^*) \;\ge\; \rho(\mathbf{x}^*). \tag{12}$$

The first inequality is the lower bound at the center, the third is the upper bound at the solution, and the middle one is the condition of Level 2. A move therefore shows that the center has the larger radius, although the history gives neither radius itself. A stay leaves the comparison unresolved, since with $r_c \leq \hat{\rho}(\mathbf{x}^*)$ the two bounds allow either point to have the larger radius. Fig. 3 draws the two bounds against the true radii on the synthetic acceptable set of Fig. 1, with $\mathbf{x}^*$ and $\mathbf{z}_c$ at other positions.

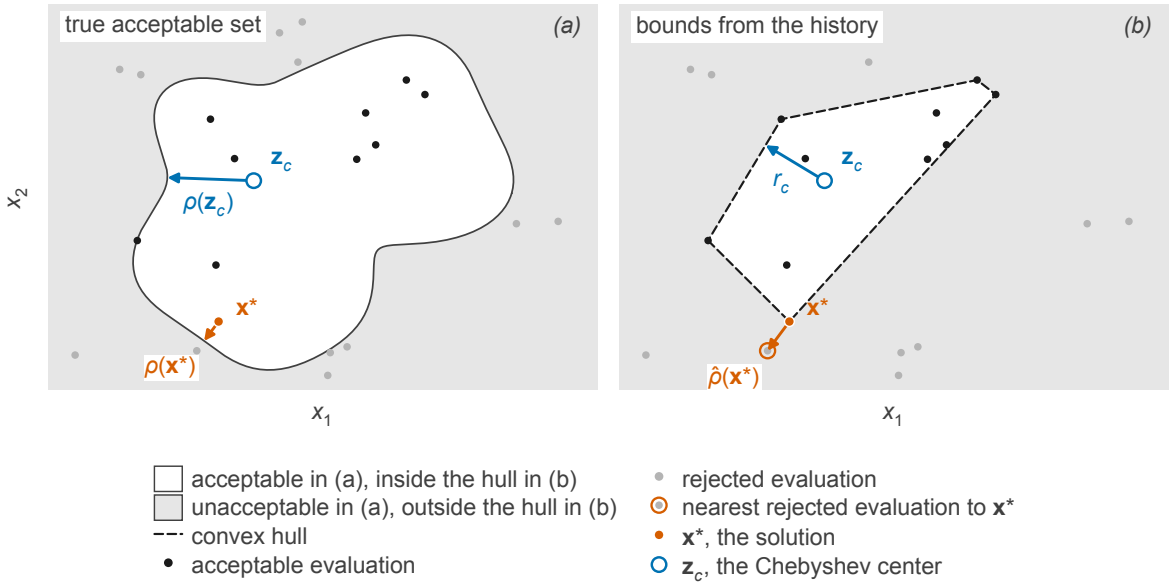

Figure 3: The two bounds against the truth on a synthetic acceptable set. (a) The true radii $\rho(\mathbf{x}^*)$ and $\rho(\mathbf{z}_c)$. (b) The lower bound r_c at $\mathbf{z}_c$ and the upper bound $\hat{\rho}(\mathbf{x}^*)$ to the nearest rejected evaluation (ringed).

An acceptable evaluation inside a hull passes Level 2 when its lower bound exceeds $\hat{\rho}(\mathbf{x}^*)$, since the chain of Eq. (12) then shows it to be more robust than the solution. The recorded evaluation $\mathbf{z}_e$ has the largest of these bounds and passes whenever another evaluation does.

In Fig. 2(e) the solution lies on a wall, where $r(\mathbf{x}^*) = 0$ and the depth ratio is infinite. The move to the center gives up 1.3% of the objective value and leaves 3.7% of the 5% tolerance. The lower bound $r_c = 0.0922$ exceeds the upper bound $\hat{\rho}(\mathbf{x}^*) = 0.0496$, so Level 2 moves (Fig. 2(f)). On this analytical problem the move raises the true radius from $\rho(\mathbf{x}^*) = 0.0353$ to $\rho(\mathbf{z}_c) = 0.1210$, a gain (the ratio of the two radii) of 3.4.

4 Benchmark on analytical test problems

4.1 Problem definition

Each benchmark problem has a closed-form objective, so the true radius of a point can be computed. This section uses the true radii only to check each decision. Each problem was optimized 10 times with NOMAD from different seeds, 50 runs in all, and analyzed at tolerances ε of 1, 2, 5, 10 and 20%. The runs of a problem share the Latin hypercube points of their initial search step. The 5 problems are defined in Appendix A and cover four unique shapes of the acceptable set.

The wedge of Fig. 2 is the elongated case, and the wedge in 5 variables is the same valley with a 4-dimensional cross-section. The two wells in 1 variable are the disconnected case, with a narrow well around the true optimum and a wide well that becomes acceptable at a tolerance between 2 and 5%. The pocket in 2 variables is a tilted valley with an unacceptable hole in its acceptable band, so the

hull can cover the hole. The quadratic in 10 variables has an interior optimum in a convex acceptable set, so the solution should be more robust than the center and Level 2 should stay.

4.2 Results

A move is considered harmful when it ends at a smaller true radius than the solution has. A stay is considered a missed move when the move would have ended at a larger true radius. Each move of Level 2 ended at a more robust point at all 5 tolerances (Tab. 2), whereas Level 1 made 13 to 30 harmful moves at each tolerance. The center lay within the tolerance in all runs with a hull, so evaluating it could not reveal a harmful move.

Table 2: Outcomes of Level 1 and Level 2 over the 10 runs of each problem at each tolerance. For the runs with a hull, a count of moves is followed by the harmful moves among them and a count of stays by the missed moves among them.

	tolerance ε				
problem (variables)	1%	2%	5%	10%	20%
Level 1			*moves (harmful)*		
two wells (1)	10 (10)	10 (10)	10 (0)	10 (0)	10 (0)
wedge (2)	10 (0)	10 (0)	10 (0)	10 (0)	10 (0)
pocket (2)	10 (10)	9 (9)	4 (4)	7 (7)	10 (10)
wedge (5)	10 (0)	10 (0)	10 (0)	10 (0)	10 (0)
quadratic (10)	10 (10)	10 (10)	9 (9)	10 (10)	10 (10)
total	50 (30)	49 (29)	43 (13)	47 (17)	50 (20)
Level 2			*moves (harmful)/stays (missed)*		
two wells (1)	0 (0)/10 (0)	0 (0)/10 (0)	10 (0)/0 (0)	10 (0)/0 (0)	10 (0)/0 (0)
wedge (2)	2 (0)/8 (8)	2 (0)/8 (8)	10 (0)/0 (0)	10 (0)/0 (0)	10 (0)/0 (0)
pocket (2)	0 (0)/10 (0)	0 (0)/9 (0)	0 (0)/4 (0)	0 (0)/7 (0)	0 (0)/10 (0)
wedge (5)	0 (0)/10 (10)	0 (0)/10 (10)	3 (0)/7 (7)	3 (0)/7 (7)	6 (0)/4 (4)
quadratic (10)	0 (0)/10 (0)	0 (0)/10 (0)	0 (0)/9 (0)	0 (0)/10 (0)	0 (0)/10 (0)
total	2 (0)/48 (18)	2 (0)/47 (18)	23 (0)/20 (7)	23 (0)/24 (7)	26 (0)/24 (4)

Fig. 4 shows the gain of each run with a hull at 5%. In all 10 runs of the two wells Level 2 moved from the narrow well to the same center of the wide one, a gain of 1.9. The 10 histories of the two wells are the same apart from the starting point. On the wedge in 2 variables Level 2 also moved in all runs to a center in the wider part of the valley, a gain of 3.2 to 3.9. The wedge in 5 variables had a similar gain, 2.5 to 4.1, but Level 2 moved in 3 of its 10 runs and had 7 missed moves.

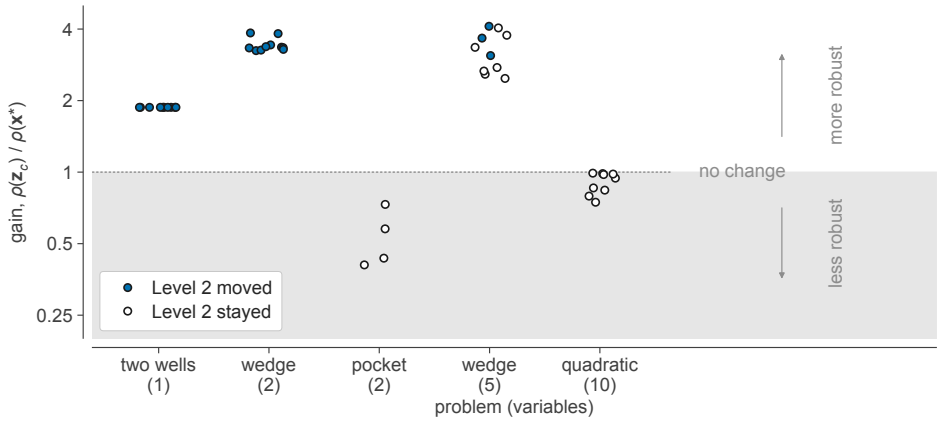

Figure 4: Each point is one benchmark run with a hull at $\varepsilon = 5\%$, at the height of its gain on a logarithmic axis.

In Fig. 5 the left stem of a run is the radius the lower bound gives away, and the right stem is the radius the upper bound adds. Level 2 therefore moves exactly when the true improvement $\rho(\mathbf{z}_c) - \rho(\mathbf{x}^*)$

exceeds the sum of the two stems. The left stems were longer in 5 variables, whereas the right stems stayed short. The lower bound reached 65 to 76% of the true radius at the center in 2 variables and 17 to 40% in 5. The upper bound was 1.1 to 1.4 times the true radius at the solution in 2 variables and 1.0 to 1.6 times in 5. The lower bound therefore caused most of the 7 missed moves.

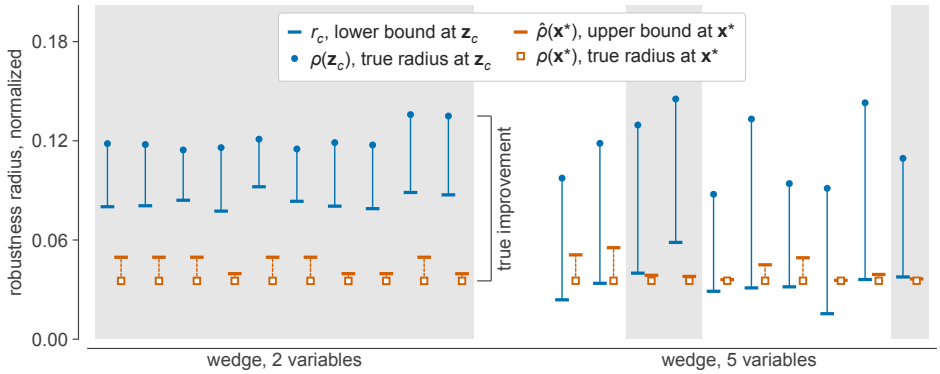

Figure 5: Each run of the wedge in 2 and in 5 variables is drawn as two stems, from r_c to $\rho(\mathbf{z}_c)$ and from $\rho(\mathbf{x}^*)$ to $\hat{\rho}(\mathbf{x}^*)$. The runs where Level 2 moved are shaded in gray.

On the quadratic Level 2 stayed in the 9 runs with a hull, and the solution was more robust than the center in each. Quickhull stopped on a numerical error in the other run while building the hull of 610 acceptable evaluations, so the method refused. The other 6 refusals came from the pocket, whose 10 runs all evaluated points inside the hole. The test of step 3 sees them only when the hull covers the hole. In those 6 runs the hull covered the hole and contained 1 to 5 rejected evaluations, as in Fig. 6(a). In the other 4 runs the hole lay outside the hull, as in Fig. 6(b). The centers beside the hole had 0.41 to 0.73 times the true radius at the solution. The 4 stays of Level 2 therefore kept the more robust point.

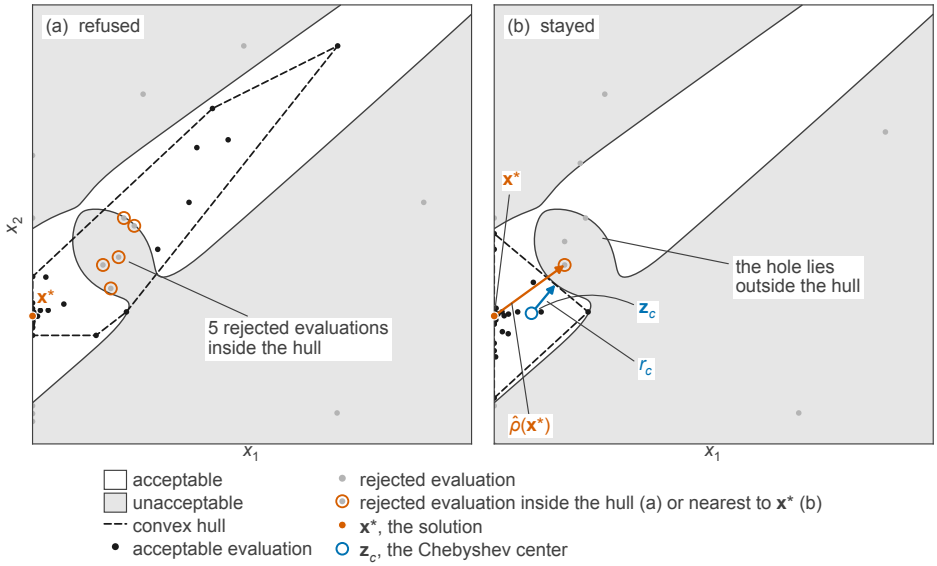

Figure 6: Two runs of the pocket show the test for a rejected evaluation inside the hull. (a) The hull covers the hole, and the rejected evaluations inside the hull are ringed. (b) The hole lies outside the hull, and the ringed rejected evaluation nearest to $\mathbf{x}^*$ sets the upper bound.

In 20 of the 23 runs that moved, 55 acceptable evaluations also passed Level 2, and all were more robust than the solution. The Chebyshev center was more robust than these evaluations in 19 of those

20 runs. Since the Chebyshev center maximizes the lower bound instead of the true radius, the volume centroid of its hull, the analytic center (Boyd and Vandenberghe 2004) and the mean of its group were compared with it. The analytic center is the point of the hull with the largest product of its distances to the walls. In most runs with 2 or more variables the volume centroid was more robust than the Chebyshev center, whereas the analytic center and the mean of the group were less robust. The lower bounds of the three points were at most r_c, so the Chebyshev center passed Level 2 most often, in 23 runs against 21, 16 and 10.

4.3 Sensitivity

The same histories were regrouped with $\kappa = 2$, 3, 4, 5, 7 and 10 in step 3. Of the 40 runs whose acceptable set is connected, $\kappa = 2$ split 16 of these sets into several groups and $\kappa = 3$ split 3. Values of κ from 4 to 10 gave the same groups as $\kappa = 5$, and all values separated the two wells. Each group of a split set has a smaller hull, and its smaller lower bound can only turn a move into a stay. On the pocket the smaller hulls also turned some refusals into stays.

At 1 and 2% the wide well was still unacceptable and the two wedges had smaller gains. Level 2 therefore moved in at most 2 runs, against 23 to 26 from 5% up. The tolerance sweep also tests the assumption that the whole hull is acceptable. At 10% the hole of the pocket is smaller than at 5%, and in 7 runs the hull covered all or part of it with no rejected evaluation inside. In those runs the lower bound exceeded the true radius at the center, whereas it was valid in all other runs with a hull. Level 2 stayed in all 7 runs, so the false bounds caused no harmful move.

The threshold of step 2 depends on the run having converged, so each history was also shortened to its first 25, 50 and 75% of evaluations. Level 2 then moved in 13, 21 and 23 runs, against 23 on the whole history, since a shorter run leaves fewer acceptable evaluations to build a hull. The two bounds stayed valid in all runs with a hull, so each of those moves ended at a point more robust than the solution of the shortened history.

5 Startup model calibration case study

The machine is a 50 MW Kaplan turbine of Hydro-Québec with adjustable guide vanes and runner blades. Each measured startup costs production time and fatigue life of the runner blades (Gagnon et al. 2010), so the startup model is calibrated on few startups. That calibration is an ill-posed inverse problem (Tarantola 2005), since distinct sets of coefficients reproduce the measured startups almost equally well.

5.1 Problem definition

Following Fahed et al. (2025), the startup model computes the rotational speed from the turbine torque characteristic. That characteristic gives the torque factor at each speed factor, guide vane opening and blade angle. The manufacturer's hill chart (the measured performance map) gives that characteristic only near the design point, whereas a typical startup could run far outside that design range. Fahed et al. (2025) therefore expressed the characteristic as a polynomial with 17 coefficients to extend it over the whole startup operating range. This paper calibrates these coefficients ($\mathbf{x}$) with NOMAD, as they did, by minimizing

$$e_{\text{tot}}(\mathbf{x}) = 0.05\,\frac{e_{\text{fit}}(\mathbf{x})}{e_{\text{fit}}(\mathbf{x}^0)} + 0.95\,\frac{e_{\text{sim}}(\mathbf{x})}{e_{\text{sim}}(\mathbf{x}^0)}, \tag{13}$$

in which the total error (e_{tot}) adds the hill-chart error (e_{fit}) and the startup error (e_{sim}), each normalized by its value at $\mathbf{x}^0$, the start point of the calibration of Fahed et al. (2025). The hill-chart error is the root-mean-square error (RMSE) between the polynomial and the hill chart. The startup error is the speed RMSE averaged over the 4 training startups.

Since the hull is limited to about 10 variables, the calibration is reduced to 4 of the 17 coefficients, and the other 13 stay at the values calibrated by Fahed et al. (2025). The 4 retained coefficients varied most among the acceptable evaluations of runs with all 17 coefficients free, since the fit constrains them least. The number 4 was chosen for this demonstration, because finding the best reduction is a separate question. NOMAD ran the reduced calibration 10 times, each run with its own seed and a budget of 1000 evaluations.

At the solution the hill-chart term makes up 5.2% of e_{tot}, so a tolerance of 5% on e_{tot} is nearly as large as the whole hill-chart term. Within that tolerance the evaluations of Fig. 7 reach a hill-chart error 97% above that of the solution. This paper therefore proposes the same tolerance on each error, with $\varepsilon = 5\%$ and the acceptable set

$$V(\varepsilon) = \{\mathbf{x} \in [\mathbf{l}, \mathbf{u}] : e_{\text{fit}}(\mathbf{x}) \le (1+\varepsilon)\, e_{\text{fit}}(\mathbf{x}^*), \quad e_{\text{sim}}(\mathbf{x}) \le (1+\varepsilon)\, e_{\text{sim}}(\mathbf{x}^*)\}. \tag{14}$$

NOMAD records e_{tot} alone, so the two errors were recomputed at the 7405 evaluations of the 10 runs. Steps 3, 4 and 6 apply unchanged, since they need only the resulting split into acceptable and rejected evaluations. Step 5 reports the change of each error in place of the objective loss (δ).

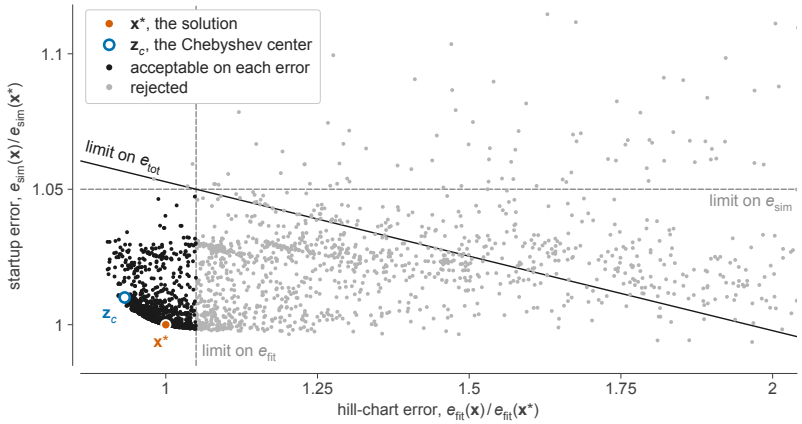

Figure 7: The evaluations of the 10 calibration runs in the plane of the two errors, each divided by its value at the solution $\mathbf{x}^*$. The gray points under the limit on e_{tot} are rejected by the tolerance on each error.

The calibration has no closed-form objective, and the true radius of a point is unknown, so each decision is checked against a reference radius estimated with new evaluations of the model. The point to be checked is perturbed by the same distance along 424 unit directions. These are the 16 directions to the corners of a 4-dimensional cube around the point, the 8 that change a single coefficient up or down, and 400 random directions. The reference radius is the largest distance at which all perturbed points stay acceptable, so a direction left out can only make it too large. The search for that distance starts from r_c at the center and from $\hat{\rho}(\mathbf{x}^*)$ at the solution. It doubles the distance until a perturbed point is unacceptable and then halves the interval 16 times.

5.2 Results

Level 1 moved in all 10 histories analyzed individually and gained a factor of 1.18 to 2.11 in reference radius. Level 2 stayed in each, since the lower bound of each run fell short of its upper bound, so all 10 stays were missed moves. Since the 10 histories record the same problem, they were pooled into one history. The solutions of 9 of the 10 runs lie within 2.2×10^{-6} of the pooled solution, under 1% of the pooled lower bound, so these runs share that solution. The largest lower bound of a single run (3.79×10^{-4}) exceeds the smallest upper bound (2.83×10^{-4}), but the two came from different

runs. The pooled history combines them, so its lower bound $r_c = 3.80 \times 10^{-4}$ exceeds its upper bound $\hat{\rho}(\mathbf{x}^*) = 2.83 \times 10^{-4}$ and Level 2 moved.

The upper bound is valid for any shape of the acceptable set, so only a false lower bound could make the move harmful. All 424 points at distance r_c from the center are acceptable, so the reference radius there is at least r_c. That reference radius is 3.977×10^{-4}, against 1.905×10^{-4} at the solution, a gain of 2.09. With 4000 random directions in place of 400, the two radii fell by at most 2.3%. Beyond the reference radius of the solution the center also keeps more perturbed points acceptable, 79% against 58% at distance 6×10^{-4} (Fig. 8).

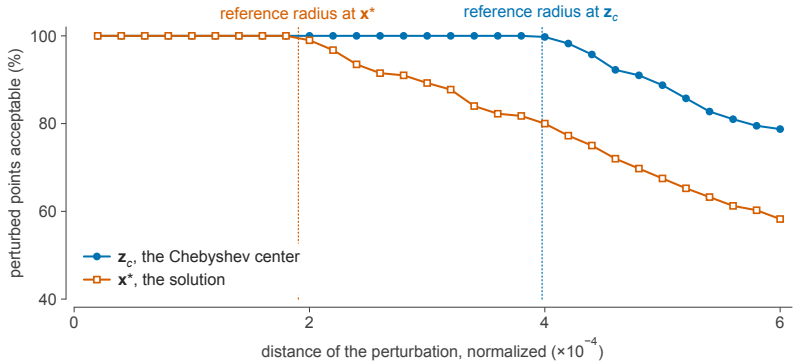

Figure 8: The share of perturbed points that stay acceptable around the solution $\mathbf{x}^*$ and the Chebyshev center $\mathbf{z}_c$, along 400 directions used for both points at all distances. The dotted lines are the reference radii of the two points.

When the pooled history was analyzed again at each tolerance from 1 to 20% in steps of 1%, the 424 points at distance r_c from the center were all acceptable. Level 2 moved at 10 of these 20 tolerances and stayed at the others, although the reference radius showed a gain of 1.35 to 2.04 at each stay.

One evaluation of the Chebyshev center in step 5 shows that the move lowers the hill-chart error by 6.8% and raises the startup error by 1.0%. The center and the solution were therefore compared on the 4 startups held out from the calibration (Fahed et al. 2025). The mean speed RMSE in Tab. 3 is 1.875 rpm at the center against 1.826 rpm at the solution, 2.7% higher and within the tolerance. The two points differ by at most 0.088 rpm on a single startup. These startups lie inside the range of the training startups, so the comparison measures interpolation only.

Table 3: The speed RMSE in rpm on each of the 4 startups held out from the calibration, in the order they were measured, and their mean.

	startup				
	1	2	3	4	mean RMSE
$\mathbf{x}^*$, the solution	1.503	1.214	1.280	3.309	1.826
$\mathbf{z}_c$, the Chebyshev center	1.496	1.302	1.327	3.375	1.875

6 Discussion

The lower bound assumes that the whole hull is acceptable. A convex hull can cover the holes of a nonconvex acceptable set, as on the pocket of Fig. 6, and step 3 detects a hole only when a rejected evaluation lies in it. The tolerance and the cut factor κ also decide whether a hole ends up inside a hull, and their best values depend on the problem (Section 4.3). Other shapes can leave such holes out. One is the alpha shape (Edelsbrunner, Kirkpatrick, and Seidel 1983), a hull whose boundary can bend inward around holes. With a suitable radius it can follow the acceptable set more closely, although finding the point farthest from its boundary costs more than the linear program of step 4.

When the budget allows a few more evaluations, the center $\mathbf{z}_c$ and points at distance r_c from it can be evaluated to test the lower bound, as in the case study. An evaluation that comes back rejected shows that part of the hull is unacceptable, and adding it to the history discards that hull. With the budget spent, the recorded evaluation $\mathbf{z}_e$ can replace the center, since its objective value is known.

A hull of a few hundred acceptable evaluations covers less of the acceptable set in more variables, so on the wedge the lower bound loosened from 2 to 5 variables (Fig. 5). The upper bound stayed tight, because NOMAD refines its mesh around the solution and leaves rejected evaluations near it. A larger budget did not help on the wedge in 5 variables, since the runs converged before spending it. A larger Latin hypercube in the initial search did not help either. The runs with the largest one spent 40,516 evaluations on average against 660, but their acceptable evaluations fell from 501 to 474. On the other hand, pooling independent runs that end at the same solution can combine the largest lower bound of one run with the smallest upper bound of another, as in the case study.

The number of walls limits the method to few variables (Fig. 9), since the hull of m points in n variables has at most of the order of $m^{\lfloor n/2 \rfloor}$ walls (McMullen 1970). A hull on a small subset of the acceptable evaluations has fewer walls but a smaller lower bound. New runs over fewer variables keep both bounds, as in the case study.

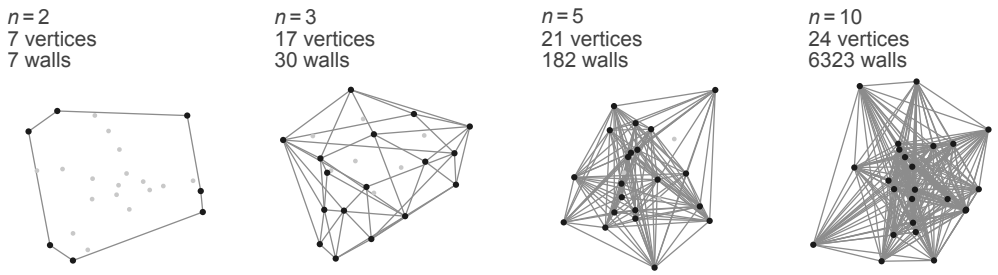

Figure 9: The hull of 24 random points of a ball in 2, 3, 5 and 10 variables, drawn by its vertices and edges projected onto a plane. Points inside the hull are gray, and the numbers of vertices and walls are given above each panel.

Tab. 4 lists the choice made in each step of the method and the other choices that could have been made. Of these, only points inside the hull were compared with the Chebyshev center (Section 4), and the rest remain untested. The bounds use only recorded variable and objective values, so they also apply to histories from other optimizers.

Table 4: The choice of this paper for each part of the analysis and other possible choices.

part	choice of this paper	other choices
evaluation history	NOMAD runs, alone or pooled into one history	other optimizers
acceptance	tolerance relative to the value at the solution, on the objective or on each error (Section 5)	absolute tolerance, threshold from the measurement noise
region	convex hull of each group	alpha shape (Edelsbrunner, Kirkpatrick, and Seidel 1983), union of balls, hull of a subset of the acceptable evaluations, region where a surrogate predicts acceptance
candidate	Chebyshev center	volume centroid, analytic center, mean of the group, recorded evaluation
decision	Level 1 and Level 2	upper bound from new evaluations near the solution, a required margin between the bounds
distance	straight line	largest change of a single variable, weighted distance

7 Conclusion

This paper proposed a post-optimality analysis that works from the evaluation history of a finished blackbox run and needs no new evaluation for its decision. On 5 analytical problems in 1, 2, 5 and 10 variables, each move of Level 2 increased the true radius at all 5 tolerances. Level 1 instead made a harmful move in 13 of the 43 runs with a hull at 5%, although each center lay within the tolerance. On the wedge the lower bound loosened from 2 to 5 variables, so 7 of the 10 runs in 5 variables ended in a missed move at that tolerance. On the startup model calibration of a Kaplan turbine, reduced to 4 coefficients, Level 2 moved only after the 10 histories were pooled. The move gained a factor of 2.09 in reference radius and changed the speed RMSE of each held-out startup by at most 0.088 rpm.

Future work will derive the tolerance from the spread of the evaluations and test the other choices of Tab. 4, such as histories from other optimizers. Open questions include how the bounds change as a history grows and how often a hull contains unacceptable points that the run did not evaluate.

Disclosure statement. The authors report there are no competing interests to declare.

Funding. This research was funded by Hydro-Québec and the Natural Sciences and Engineering Research Council of Canada (NSERC) [funding reference number 589463]. This work was supported by Mitacs through the Mitacs Accelerate Program.

Data availability statement. The evaluation histories of the benchmark and the analysis code are available from the corresponding author on request. The data of the case study are proprietary to Hydro-Québec and are not publicly available.

A The test problems

The lowest value of each objective is 1, or close to 1 on the pocket. All runs reached that value, so the threshold $f^* + |f^*|\,\varepsilon$ between acceptable and rejected evaluations is close to $1 + \varepsilon$. Tab. A.1 gives the starting point, the evaluation budget and the size of the initial Latin hypercube search of each problem. For each run the nominal starting point was moved in each variable by a random amount of up to 5% of its range. The seed of the run set these amounts, and the moved point was clipped to the bounds. The points of the initial Latin hypercube search were the same in all runs of a problem. The runs stopped before spending their budgets, when the mesh of NOMAD reached its minimum size. Fig. A.1 draws the 3 problems in 1 and 2 variables.

Table A.1: NOMAD 4.5.1 ran each test problem 10 times through its Python interface PyNomad. The column n gives the number of variables, and the start is the nominal starting point before the random move of each run. The budget is the largest number of evaluations allowed in a run, and the search is the number of points of the initial Latin hypercube search. The other parameters of NOMAD were left at their defaults.

problem	n	start	budget	search
two wells	1	0.62	400	60
wedge	2	$(0.7, 0.7)$	600	60
pocket	2	$(0.8, 0.6)$	400	30
wedge	5	$(0.7, 0.5, -0.5, 0.5, -0.5)$	1500	60
quadratic	10	$\mathbf{c} + 0.55$	1000	50

Two wells, 1 variable. Two Gaussian wells on $[-1, 1]$, the deeper one narrow and the shallower one wide,

$$f(x) = 1 + g(x) - \min g, \tag{A.1}$$
$$g(x) = -\exp\left(-\frac{(x+0.45)^2}{0.02}\right) - 0.985\exp\left(-\frac{(x-0.42)^2}{0.1568}\right).$$

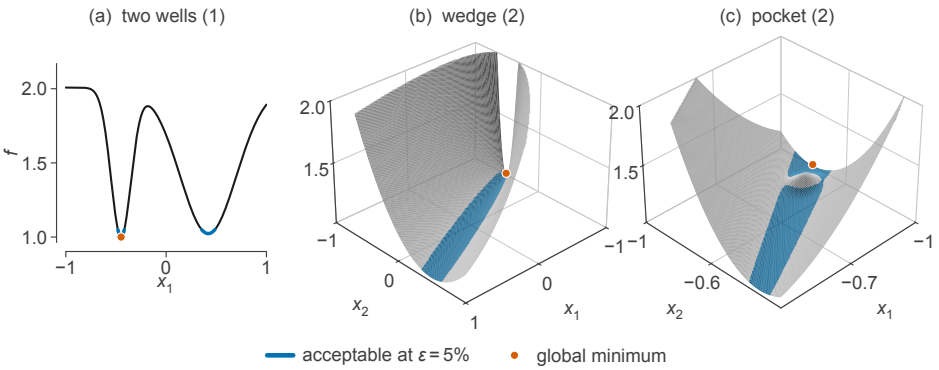

Figure A.1: The 3 test problems in 1 and 2 variables, (a) the two wells, (b) the wedge and (c) the pocket. The acceptable part of each function, within $\varepsilon = 5\%$ of its lowest value, is drawn in blue. Panel (c) shows the low corner of the domain of the pocket, where the bump rises above the threshold and opens a hole at the lower edge of the acceptable band.

Wedge, 2 variables. A valley along $x_2 = 0.5\,x_1$ on $[-1,1]^2$, narrow at $x_1 = -1$ and widening toward $x_1 = 1$, with its bottom lowest at the narrow end,

$$f(\mathbf{x}) = 1 + a(x_1)\,(x_2 - 0.5\,x_1)^2 + 0.016\,(x_1 + 1), \tag{A.2}$$
$$a(x_1) = \frac{40}{1 + 25\,(x_1 + 1)}.$$

Pocket, 2 variables. A tilted valley on $[-1,1]^2$ with a bump at $\mathbf{p} = (-0.9, -0.72)$,

$$f(\mathbf{x}) = 1 + 4\,(x_2 - 0.8\,x_1)^2 + 0.06\,(x_1 + 1) + 0.1\exp\left(-\frac{\|\mathbf{x} - \mathbf{p}\|_2^2}{0.05^2}\right). \tag{A.3}$$

At $\varepsilon = 5\%$ the bump opens an unacceptable hole at the lower edge of the acceptable band, so the hole joins the unacceptable region below the band. The acceptable strip above the hole is narrow, and the band closes around the hole only at larger tolerances.

Wedge, 5 variables. The wedge of Eq. (A.2) on $[-1,1]^5$, with its valley along the x_1 axis, a 4-dimensional cross-section and the same $a(x_1)$,

$$f(\mathbf{x}) = 1 + a(x_1)\sum_{i=2}^{5} x_i^2 + 0.016\,(x_1 + 1). \tag{A.4}$$

Its lowest value is at the point $(-1, 0, 0, 0, 0)$ of the narrow end.

Quadratic, 10 variables. A separable quadratic on $[-1,1]^{10}$,

$$f(\mathbf{x}) = 1 + \sum_{i=1}^{10} w_i\,(x_i - c_i)^2, \tag{A.5}$$

with weights $\mathbf{w}$ spanning 4 orders of magnitude and its lowest value at $\mathbf{c}$,

$$\mathbf{w} = (3,\ 1,\ 0.3,\ 0.1,\ 0.03,\ 0.01,\ 0.003,\ 0.001,\ 0.0005,\ 0.0002),$$
$$\mathbf{c} = (0.20,\ -0.35,\ 0.10,\ 0.45,\ -0.15,\ 0.30,\ -0.25,\ 0.05,\ 0.40,\ -0.10).$$

References

Audet, Charles, and J. E. Dennis, Jr. 2006. Mesh Adaptive Direct Search Algorithms for Constrained Optimization. SIAM J Optim 17 (1): 188–217. https://doi.org/10.1137/040603371.

Audet, Charles, J. E. Dennis, Jr., and Sébastien Le Digabel. 2012. Trade-off Studies in Blackbox Optimization. Optim Methods Softw 27 (4-5): 613–624. https://doi.org/10.1080/10556788.2011.571687.

Audet, Charles, and Warren Hare. 2017. Derivative-Free and Blackbox Optimization. Cham, Switzerland: Springer.

Audet, Charles, Sébastien Le Digabel, Viviane Rochon Montplaisir, and Christophe Tribes. 2022. Algorithm 1027: NOMAD Version 4: Nonlinear Optimization With the MADS Algorithm. ACM Trans Math Softw 48 (3): 35:1–35:22. https://doi.org/10.1145/3544489.

Barber, C. Bradford, David P. Dobkin, and Hannu Huhdanpaa. 1996. The Quickhull Algorithm for Convex Hulls. ACM Trans Math Softw 22 (4): 469–483. https://doi.org/10.1145/235815.235821.

Bertsimas, Dimitris, Omid Nohadani, and Kwong Meng Teo. 2010. Robust Optimization for Unconstrained Simulation-Based Problems. Oper Res 58 (1): 161–178. https://doi.org/10.1287/opre.1090.0715.

Beven, Keith. 2006. A Manifesto for the Equifinality Thesis. J Hydrol 320 (1–2): 18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007.

Beven, Keith, and Andrew Binley. 1992. The Future of Distributed Models: Model Calibration and Uncertainty Prediction. Hydrol Process 6 (3): 279–298. https://doi.org/10.1002/hyp.3360060305.

Boyd, Stephen, and Lieven Vandenberghe. 2004. Convex Optimization. Cambridge, UK: Cambridge University Press.

Custódio, A. L., and L. N. Vicente. 2007. Using Sampling and Simplex Derivatives in Pattern Search Methods. SIAM J Optim 18 (2): 537–555. https://doi.org/10.1137/050646706.

Edelsbrunner, Herbert, David G. Kirkpatrick, and Raimund Seidel. 1983. On the Shape of a Set of Points in the Plane. IEEE Trans Inf Theory 29 (4): 551–559. https://doi.org/10.1109/TIT.1983.1056714.

Fahed, Melad, Arthur Favrel, Martin Gagnon, Stéphane Alarie, and Michael Kokkolaras. 2025. Hydropower Unit Digital Twin Calibration Using Monitoring Data and Blackbox Optimization. IOP Conf Ser Earth Environ Sci 1561 (1): 012018. https://doi.org/10.1088/1755-1315/1561/1/012018.

Fiacco, Anthony V. 1983. Introduction to Sensitivity and Stability Analysis in Nonlinear Programming. Vol. 165 of Mathematics in Science and Engineering. New York: Academic Press.

Gagnon, M., S. A. Tahan, P. Bocher, and D. Thibault. 2010. Impact of Startup Scheme on Francis Runner Life Expectancy. IOP Conf Ser Earth Environ Sci 12 (1): 012107. https://doi.org/10.1088/1755-1315/12/1/012107.

Hu, Yi-Xiang. 2026. Position Paper: Post-Solve Robustness in Decision Engines: Feasible Regions and Smoothness Under Perturbations. arXiv:2606.00002.

McMullen, Peter. 1970. The Maximum Numbers of Faces of a Convex Polytope. Mathematika 17 (2): 179–184. https://doi.org/10.1112/S0025579300002850.

NASEM (National Academies of Sciences, Engineering, and Medicine). 2024. Foundational Research Gaps and Future Directions for Digital Twins. Washington, DC: The National Academies Press.

Raue, Andreas, Marcel Schilling, Julie Bachmann, Andrew Matteson, Max Schelker, Daniel Kaschek, Sabine Hug, et al. 2013. Lessons Learned from Quantitative Dynamical Modeling in Systems Biology. PLoS ONE 8 (9): e74335. https://doi.org/10.1371/journal.pone.0074335.

Sibson, Robin. 1973. SLINK: An Optimally Efficient Algorithm for the Single-Link Cluster Method. Comput J 16 (1): 30–34. https://doi.org/10.1093/comjnl/16.1.30.

Spear, Robert C., and George M. Hornberger. 1980. Eutrophication in Peel Inlet—II. Identification of Critical Uncertainties via Generalized Sensitivity Analysis. Water Res 14 (1): 43–49. https://doi.org/10.1016/0043-1354(80)90040-8.

Tarantola, Albert. 2005. Inverse Problem Theory and Methods for Model Parameter Estimation. Philadelphia, PA: SIAM.

Vidigal, Luis M., and Stephen W. Director. 1982. A Design Centering Algorithm for Nonconvex Regions of Acceptability. IEEE Trans Comput-Aided Des Integr Circuits Syst 1 (1): 13–24. https://doi.org/10.1109/TCAD.1982.1269992.

Wang, Yilun, and Christine A. Shoemaker. 2014. Sensitivity Analysis for Computationally Expensive Models Using Optimization and Objective-Oriented Surrogate Approximations. arXiv:1410.7291.