Quality of a Pareto Front Approximation

Optimization of Complex Systems – April 2026

Andrea Brilli

Sapienza University of Rome

How good is our Pareto front approximation?

At the end of a DMS run, the algorithm returns a list $\mathcal{L}$ of mutually nondominated points.

Question: How do we know if this list is a *good* approximation of the Pareto front?

How good is our Pareto front approximation?

At the end of a DMS run, the algorithm returns a list $\mathcal{L}$ of mutually nondominated points.

Question: How do we know if this list is a *good* approximation of the Pareto front?

In single-objective optimization this is trivial: we compare function values. In multiobjective optimization, we need to evaluate a *set* of points.

A good approximation $\mathcal{L}$ should satisfy at least two properties:

- **Convergence**: the points in $\mathcal{L}$ should be *close* to the true Pareto front Ω_P.
- **Coverage**: the points in $\mathcal{L}$ should be *spread* across the entire front, leaving no large unexplored gaps.

How good is our Pareto front approximation?

At the end of a DMS run, the algorithm returns a list $\mathcal{L}$ of mutually nondominated points.

Question: How do we know if this list is a *good* approximation of the Pareto front?

In single-objective optimization this is trivial: we compare function values. In multiobjective optimization, we need to evaluate a *set* of points.

A good approximation $\mathcal{L}$ should satisfy at least two properties:

- **Convergence**: the points in $\mathcal{L}$ should be *close* to the true Pareto front Ω_P.
- **Coverage**: the points in $\mathcal{L}$ should be *spread* across the entire front, leaving no large unexplored gaps.

We focus today on **spread metrics**: quantitative measures of how well $\mathcal{L}$ covers the Pareto front.

These are also useful *inside* DMS itself, as we will see shortly.

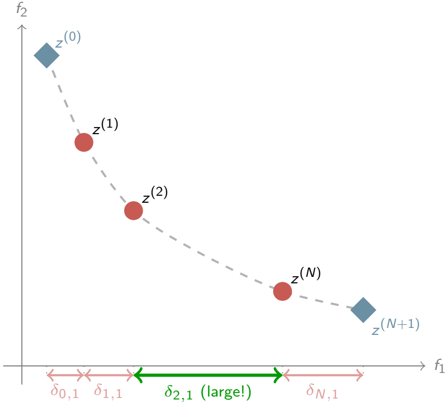
Setup: gaps between consecutive points

Let $\mathcal{L} = \{z^{(1)}, \ldots, z^{(N)}\} \subset \mathbb{R}^p$ be a Pareto front approximation with $p = 2$ objectives.

We augment $\mathcal{L}$ with two **extreme anchor points** $z^{(0)}$ and $z^{(N+1)}$: the endpoints of the (estimated) true Pareto front. For each objective $j \in \{1, \ldots, p\}$, sort all $N + 2$ points by increasing j-th component.

Define the **gap sequence** for objective j:

$$\delta_{i,j} = z_j^{(i+1)} - z_j^{(i)}, \quad i = 0, \ldots, N.$$



Setup: gaps between consecutive points

Let $\mathcal{L} = \{z^{(1)}, \dots, z^{(N)}\} \subset \mathbb{R}^p$ be a Pareto front approximation with $p = 2$ objectives.

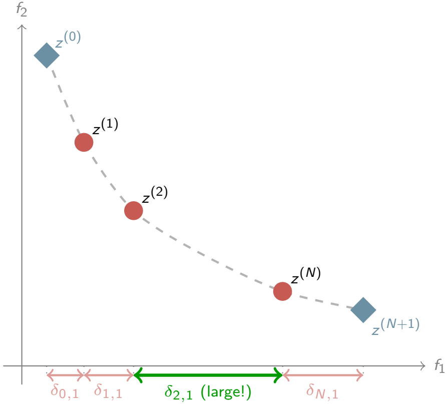
We augment $\mathcal{L}$ with two **extreme anchor points** $z^{(0)}$ and $z^{(N+1)}$: the endpoints of the (estimated) true Pareto front. For each objective $j \in \{1, \dots, p\}$, sort all $N + 2$ points by increasing j-th component.

Define the **gap sequence** for objective j:

$$\delta_{i,j} = z_j^{(i+1)} - z_j^{(i)}, \quad i = 0, \dots, N.$$

- $\delta_{0,j}$, $\delta_{N,j}$: **edge gaps** (distance from extreme anchors to the nearest list point)
- $\delta_{1,j}, \dots, \delta_{N-1,j}$: **interior gaps**

Ideal case: all gaps are equal and edge gaps are zero — the list perfectly tiles the front.



The Γ metric: the largest gap

Definition (Custódio et al., 2011)

Given a Pareto front approximation $\mathcal{L}$ *with gap sequence* $\{\delta_{i,j}\}$*, define*

$$\Gamma = \max_{j \in \{1,\ldots,p\}} \max_{i \in \{0,\ldots,N\}} \delta_{i,j}.$$

Γ is simply the size of the **largest gap** in the list, measured along any objective axis.

Question: What does a small Γ tell us? And what does it *not* tell us?

The Γ metric: the largest gap

Definition (Custódio et al., 2011)

Given a Pareto front approximation $\mathcal{L}$ with gap sequence $\{\delta_{i,j}\}$, define

$$\Gamma = \max_{j \in \{1,\ldots,p\}} \max_{i \in \{0,\ldots,N\}} \delta_{i,j}.$$

Γ is simply the size of the **largest gap** in the list, measured along any objective axis.

Question: What does a small Γ tell us? And what does it *not* tell us?

- **Small** Γ: no single region of the front is left poorly covered — the worst gap is small.
- **Large** Γ: at least one significant unexplored region exists.

What Γ does **not** tell us: it says nothing about the *rest* of the distribution. Two fronts can share the same Γ while one is perfectly uniform and the other has all points clustered in a corner.

Γ is a pessimist: it only reports the worst case.

Γ inside DMS: choosing the next poll center

Recall that in DMS, at each iteration we must **select a poll center** from the list $\mathcal{L}_k$.

Question: We have no preference information. Which point in $\mathcal{L}_k$ is the most "useful" to poll around next?

Γ inside DMS: choosing the next poll center

Recall that in DMS, at each iteration we must **select a poll center** from the list $\mathcal{L}_k$.

Question: We have no preference information. Which point in $\mathcal{L}_k$ is the most "useful" to poll around next?

Answer: Poll around the point that is adjacent to the **largest gap** in the current front approximation.

Formally: compute Γ on $\mathcal{L}_k$, identify the pair (i^*, j^*) achieving the maximum, and select as the next poll center the point $z^{(i^*)}$ or $z^{(i^*+1)}$ adjacent to that gap.

Intuition: We are deploying our limited function evaluations where coverage is most needed — a greedy gap-filling strategy.

Γ inside DMS: choosing the next poll center

Recall that in DMS, at each iteration we must **select a poll center** from the list $\mathcal{L}_k$.

Question: We have no preference information. Which point in $\mathcal{L}_k$ is the most "useful" to poll around next?

Answer: Poll around the point that is adjacent to the **largest gap** in the current front approximation.

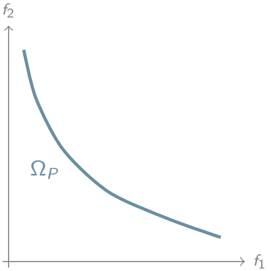
Formally: compute Γ on $\mathcal{L}_k$, identify the pair (i^*, j^*) achieving the maximum, and select as the next poll center the point $z^{(i^*)}$ or $z^{(i^*+1)}$ adjacent to that gap.

Intuition: We are deploying our limited function evaluations where coverage is most needed — a greedy gap-filling strategy.

Remark: This is a *heuristic*, not a convergence requirement. The theoretical results of DMS hold for any selection strategy. Γ-based ordering is used in practice because it tends to produce more uniform Pareto front approximations (Custódio et al., 2011, §6.3).

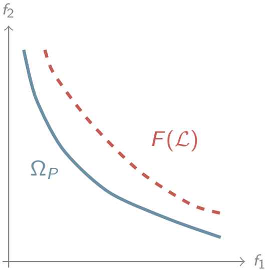
How good is our approximation of the Pareto front?

Question: In single-objective optimization, the quantity $f(x_k) - f(x^*)$ measures how far we are from the optimal value in the objective space. Given an approximation $\mathcal{L}$ of the Pareto front Ω_P, how do we define an analogous notion of distance?



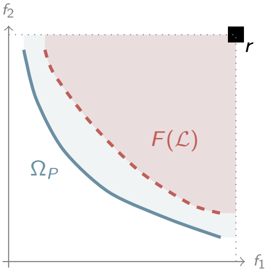
How good is our approximation of the Pareto front?

Question: In single-objective optimization, the quantity $f(x_k) - f(x^*)$ measures how far we are from the optimal value in the objective space. Given an approximation $\mathcal{L}$ of the Pareto front Ω_P, how do we define an analogous notion of distance?



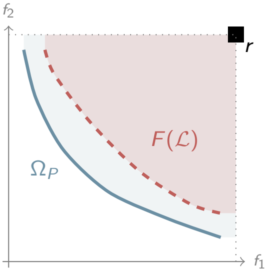
How good is our approximation of the Pareto front?

Question: In single-objective optimization, the quantity $f(x_k) - f(x^*)$ measures how far we are from the optimal value in the objective space. Given an approximation $\mathcal{L}$ of the Pareto front Ω_P, how do we define an analogous notion of distance?



How good is our approximation of the Pareto front?

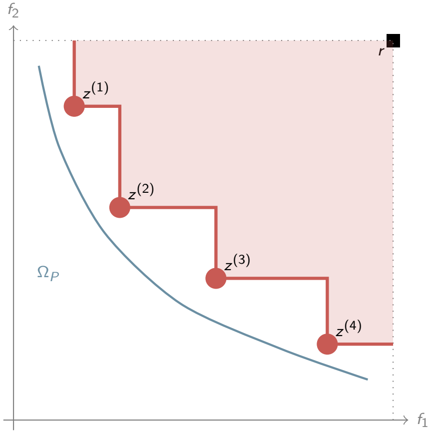
Question: In single-objective optimization, the quantity $f(x_k) - f(x^*)$ measures how far we are from the optimal value in the objective space. Given an approximation $\mathcal{L}$ of the Pareto front Ω_P, how do we define an analogous notion of distance?



The distance can be expressed in terms of how close the hypervolume of $F(\mathcal{L})$ is to the one of the true Pareto front Ω_P with respect to a reference point r.

Hypervolume of a finite approximation

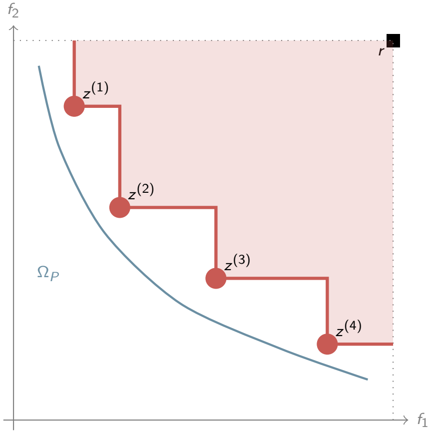
Question: What happens when our approximation $\mathcal{L}$ consists of a *finite* set of points?



The staircase connects the points of $\mathcal{L}$ in objective space. The **hypervolume** $\mathrm{HV}(\mathcal{L}, r)$ is the area of the region dominated by $\mathcal{L}$ and bounded above by the reference point r.

Hypervolume of a finite approximation

Question: What happens when our approximation $\mathcal{L}$ consists of a *finite* set of points?



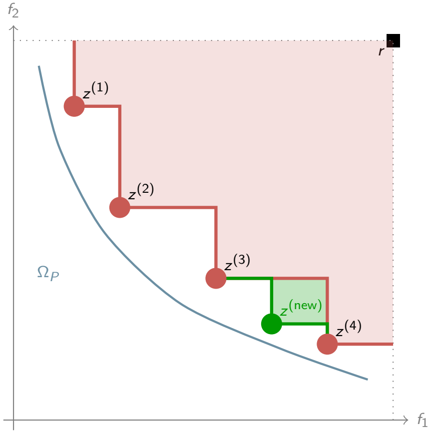
The staircase connects the points of $\mathcal{L}$ in objective space. The **hypervolume** $\mathrm{HV}(\mathcal{L}, r)$ is the area of the region dominated by $\mathcal{L}$ and bounded above by the reference point r.

The bigger the hypervolume of $\mathcal{L}$, the better the approximation of the true Pareto front.

Adding a new nondominated point to $\mathcal{L}$ can only **increase** the hypervolume.

Hypervolume of a finite approximation

Question: What happens when our approximation $\mathcal{L}$ consists of a *finite* set of points?



The staircase connects the points of $\mathcal{L}$ in objective space. The **hypervolume** $\text{HV}(\mathcal{L}, r)$ is the area of the region dominated by $\mathcal{L}$ and bounded above by the reference point r.

The bigger the hypervolume of $\mathcal{L}$, the better the approximation of the true Pareto front.

Adding a new nondominated point to $\mathcal{L}$ can only **increase** the hypervolume.

A new point $z^{(\text{new})}$ is found: it expands the dominated region, strictly increasing the hypervolume by $\Delta\text{HV} > 0$.

The hypervolume indicator: notation

Definition (Hypervolume indicator)

Let $\mathcal{L} = \{z^{(1)}, \dots, z^{(N)}\} \subset \mathbb{R}^p$ *be a Pareto front approximation and let* $r \in \mathbb{R}^p$ *be a* **reference point** *such that* $z_j^{(i)} < r_j$ *for all* $i \in \{1, \dots, N\}$ *and all* $j \in \{1, \dots, p\}$*.*
The **hypervolume indicator** $\mathrm{HV}(\mathcal{L}, r)$ *is the p-dimensional volume of the region* **dominated by** $\mathcal{L}$ *and* **bounded above by** r*.*

The hypervolume indicator: notation

Definition (Hypervolume indicator)

Let $\mathcal{L} = \{z^{(1)}, \ldots, z^{(N)}\} \subset \mathbb{R}^p$ *be a Pareto front approximation and let* $r \in \mathbb{R}^p$ *be a* **reference point** *such that* $z_j^{(i)} < r_j$ *for all* $i \in \{1, \ldots, N\}$ *and all* $j \in \{1, \ldots, p\}$. *The* **hypervolume indicator** $\mathrm{HV}(\mathcal{L}, r)$ *is the p-dimensional volume of the region* **dominated by** $\mathcal{L}$ *and* **bounded above by** r.

- $p = 1$: reduces to $r_1 - \min_i z_1^{(i)}$ — exactly the single-objective gap to the optimum.
- $p = 2$: the area over the staircase (as in the previous slide).
- $p \geq 3$: a higher-dimensional volume bounded by a step surface and r.

The hypervolume indicator: notation

Definition (Hypervolume indicator)

Let $\mathcal{L} = \{z^{(1)}, \ldots, z^{(N)}\} \subset \mathbb{R}^p$ be a Pareto front approximation and let $r \in \mathbb{R}^p$ be a **reference point** *such that $z_j^{(i)} < r_j$ for all $i \in \{1, \ldots, N\}$ and all $j \in \{1, \ldots, p\}$.*
The **hypervolume indicator** $\mathrm{HV}(\mathcal{L}, r)$ *is the p-dimensional volume of the region* **dominated by** $\mathcal{L}$ *and* **bounded above by** r*.*

- $p = 1$: reduces to $r_1 - \min_i z_1^{(i)}$ — exactly the single-objective gap to the optimum.
- $p = 2$: the area over the staircase (as in the previous slide).
- $p \geq 3$: a higher-dimensional volume bounded by a step surface and r.

Key property (monotonicity): If $\mathcal{L}'$ dominates $\mathcal{L}$, then $\mathrm{HV}(\mathcal{L}', r) > \mathrm{HV}(\mathcal{L}, r)$ for *any* valid r.

The reference point: a hidden dependency

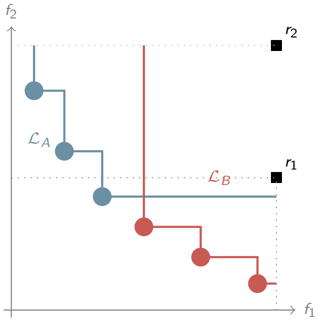
The definition of $\mathrm{HV}(\mathcal{L}, r)$ requires a reference point r that is **strictly dominated by all points in $\mathcal{L}$**.

Question: Does the *ranking* of two approximations by hypervolume depend on the choice of r?

The reference point: a hidden dependency

The definition of $\mathrm{HV}(\mathcal{L}, r)$ requires a reference point r that is **strictly dominated by all points in** $\mathcal{L}$.

Question: Does the *ranking* of two approximations by hypervolume depend on the choice of r?



Yes! Consider $\mathcal{L}_A$ (covering the high-f_2 region) and $\mathcal{L}_B$ (covering the low-f_2 region). These two fronts are incomparable under Pareto dominance.

With reference point r_1 (close to $\mathcal{L}_B$):

$$\mathrm{HV}(\mathcal{L}_B, r_1) > \mathrm{HV}(\mathcal{L}_A, r_1)$$

With reference point r_2 (far, balanced):

$$\mathrm{HV}(\mathcal{L}_A, r_2) > \mathrm{HV}(\mathcal{L}_B, r_2)$$

The choice of r implicitly encodes a preference over regions of the objective space.

How to choose the reference point in practice?

Question: If we do not know the true Pareto front Ω_P, how do we choose r?

How to choose the reference point in practice?

Question: If we do not know the true Pareto front Ω_P, how do we choose r?

An ideal choice would be the **nadir point** z^{nad}: the component-wise worst value over the true Pareto front,

$$z_j^{\text{nad}} = \max_{z \in \Omega_P} z_j, \qquad j = 1, \ldots, p.$$

Using $r = z^{\text{nad}}$ anchors the hypervolume to the natural extent of the problem.

Problem: z^{nad} requires knowledge of Ω_P, which we do not have.

How to choose the reference point in practice?

Question: If we do not know the true Pareto front Ω_P, how do we choose r?

An ideal choice would be the **nadir point** z^{nad}: the component-wise worst value over the true Pareto front,

$$z_j^{\text{nad}} = \max_{z \in \Omega_P} z_j, \qquad j = 1, \ldots, p.$$

Using $r = z^{\text{nad}}$ anchors the hypervolume to the natural extent of the problem.

Problem: z^{nad} requires knowledge of Ω_P, which we do not have.

Common practical strategies:

- **Estimated nadir from the current list**: $\hat{z}_j^{\text{nad}} = \max_{z \in \mathcal{L}} z_j + \varepsilon$, for a small offset $\varepsilon > 0$. The offset ensures r is strictly dominated by all points and adds robustness.
- **Fixed worst-case bound**: if domain knowledge provides upper bounds u_j on each objective, set $r_j = u_j$.

How to choose the reference point in practice?

Question: If we do not know the true Pareto front Ω_P, how do we choose r?

An ideal choice would be the **nadir point** z^{nad}: the component-wise worst value over the true Pareto front,

$$z_j^{\text{nad}} = \max_{z \in \Omega_P} z_j, \qquad j = 1, \dots, p.$$

Using $r = z^{\text{nad}}$ anchors the hypervolume to the natural extent of the problem.

Problem: z^{nad} requires knowledge of Ω_P, which we do not have.

Common practical strategies:

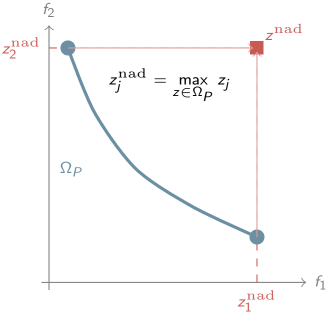
- **Estimated nadir from the current list**: $\hat{z}_j^{\text{nad}} = \max_{z \in \mathcal{L}} z_j + \varepsilon$, for a small offset $\varepsilon > 0$. The offset ensures r is strictly dominated by all points and adds robustness.
- **Fixed worst-case bound**: if domain knowledge provides upper bounds u_j on each objective, set $r_j = u_j$.

Practical takeaway: the reference point is a parameter of the indicator, not of the problem. When comparing two algorithms by hypervolume, both must use the **same** r, otherwise the comparison is meaningless.

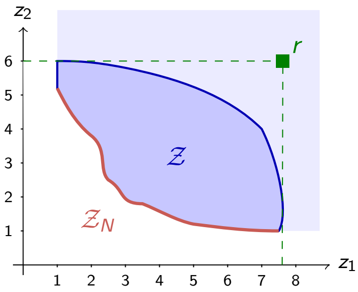
How to choose the reference point?

Question: If we do not know Ω_P, how do we choose r?

Nadir point z^{nad}



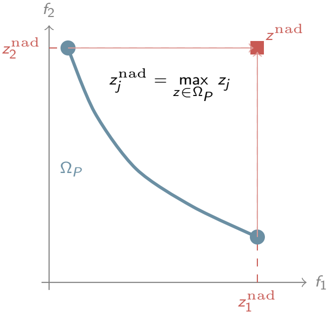
Worst case over $\mathcal{Z} = F(\mathcal{F})$



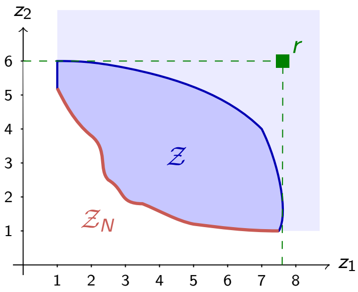
How to choose the reference point?

Question: If we do not know Ω_P, how do we choose r?

Nadir point z^{nad}



Worst case over $\mathcal{Z} = F(\mathcal{F})$



Practical takeaway: r is a *parameter* of the indicator, not of the problem. When comparing two algorithms, both must use the **same** r, otherwise the comparison is meaningless.

Recall: sufficient decrease in DMS

In the DMS convergence analysis, the **sufficient decrease** variant declares an iteration successful only if the new trial point $x_k + \alpha_k d$ *significantly improves* the list:

$$\exists\ i \in \{1, \dots, N\},\ j \in \{1, \dots, p\} \ \text{ such that } \ f_j(x_k + \alpha_k d) < f_j(z^{(i)}) - \rho(\alpha_k),$$

where $\rho : (0, +\infty) \to (0, +\infty)$ is a **forcing function**: continuous, nondecreasing, with $\rho(t)/t \to 0$ as $t \to 0$ (e.g., $\rho(t) = t^{1+a}$, $a > 0$).

Question: This condition tracks individual objective values of individual list points. Is there a more compact, *scalar* way to express the quality improvement at each successful iteration?

Recall: sufficient decrease in DMS

In the DMS convergence analysis, the **sufficient decrease** variant declares an iteration successful only if the new trial point $x_k + \alpha_k d$ *significantly improves* the list:

$$\exists\ i \in \{1, \ldots, N\},\ j \in \{1, \ldots, p\} \text{ such that } f_j(x_k + \alpha_k d) < f_j(z^{(i)}) - \rho(\alpha_k),$$

where $\rho : (0, +\infty) \to (0, +\infty)$ is a **forcing function**: continuous, nondecreasing, with $\rho(t)/t \to 0$ as $t \to 0$ (e.g., $\rho(t) = t^{1+a}$, $a > 0$).

Question: This condition tracks individual objective values of individual list points. Is there a more compact, *scalar* way to express the quality improvement at each successful iteration?

Observation: We already have a scalar measure of quality for $\mathcal{L}$: the **hypervolume** $\mathrm{HV}(\mathcal{L}, r)$. If a new nondominated point enters the list, the staircase strictly expands, and so does HV.

Sufficient increase of the hypervolume

The DMS sufficient decrease condition tracks individual objective values of individual list points. We now propose a **cleaner, scalar reformulation**: require that each successful iteration produces a minimum increase in the hypervolume of the list.

Definition (Sufficient hypervolume increase)

Given a reference point r, an iteration is declared **successful** *if the new trial point $x_k + \alpha_k d$ enters the list and:*

$$\mathrm{HV}(\mathcal{L}_{k+1}, r) \; > \; \mathrm{HV}(\mathcal{L}_k, r) + \rho(\alpha_k \|d\|),$$

where $\rho : (0, +\infty) \to (0, +\infty)$ is a **forcing function***: continuous, nondecreasing, with $\rho(t)/t \to 0$ as $t \downarrow 0$.*

Sufficient increase of the hypervolume

The DMS sufficient decrease condition tracks individual objective values of individual list points. We now propose a **cleaner, scalar reformulation**: require that each successful iteration produces a minimum increase in the hypervolume of the list.

Definition (Sufficient hypervolume increase)

Given a reference point r, an iteration is declared **successful** *if the new trial point $x_k + \alpha_k d$ enters the list and:*

$$\mathrm{HV}(\mathcal{L}_{k+1}, r) \ > \ \mathrm{HV}(\mathcal{L}_k, r) + \rho(\alpha_k \|d\|),$$

where $\rho : (0, +\infty) \to (0, +\infty)$ is a **forcing function***: continuous, nondecreasing, with $\rho(t)/t \to 0$ as $t \downarrow 0$.*

Question: Compare this with the DMS condition. What has changed structurally?

Sufficient increase of the hypervolume

The DMS sufficient decrease condition tracks individual objective values of individual list points. We now propose a **cleaner, scalar reformulation**: require that each successful iteration produces a minimum increase in the hypervolume of the list.

Definition (Sufficient hypervolume increase)

Given a reference point r, an iteration is declared **successful** *if the new trial point $x_k + \alpha_k d$ enters the list and:*

$$\mathrm{HV}(\mathcal{L}_{k+1}, r) \; > \; \mathrm{HV}(\mathcal{L}_k, r) + \rho(\alpha_k \|d\|),$$

where $\rho : (0, +\infty) \to (0, +\infty)$ is a **forcing function***: continuous, nondecreasing, with $\rho(t)/t \to 0$ as $t \downarrow 0$.*

Question: Compare this with the DMS condition. What has changed structurally?

The hypervolume condition is a **single scalar inequality**: the whole approximation quality must improve by at least $\rho(\alpha_k)$. It is *agnostic* to which point or which objective improves, only the net gain in dominated volume matters.

The hypervolume is bounded

Before proving convergence, we need to ensure that $\mathrm{HV}(\mathcal{L}_k, r)$ is **well-defined and finite** for all k.

Question: Under what conditions is $\mathrm{HV}(\mathcal{L}_k, r)$ guaranteed to be finite?

The hypervolume is bounded

Before proving convergence, we need to ensure that $\mathrm{HV}(\mathcal{L}_k, r)$ is **well-defined and finite** for all k.

Question: Under what conditions is $\mathrm{HV}(\mathcal{L}_k, r)$ guaranteed to be finite?

- The hypervolume is the volume of a region in $\mathbb{R}^p$. It is finite as long as that region is **bounded**.
- The dominated region is bounded between $\mathcal{L}_k$ and r: bounded above by r, bounded below by the ideal point z^{id}.
- So we need r to be finite — which is guaranteed if the **image set** $\mathcal{Z} = F(\mathcal{F})$ is bounded.

The hypervolume is bounded

Before proving convergence, we need to ensure that $\mathrm{HV}(\mathcal{L}_k, r)$ is **well-defined and finite** for all k.

Question: Under what conditions is $\mathrm{HV}(\mathcal{L}_k, r)$ guaranteed to be finite?

- The hypervolume is the volume of a region in $\mathbb{R}^p$. It is finite as long as that region is **bounded**.
- The dominated region is bounded between $\mathcal{L}_k$ and r: bounded above by r, bounded below by the ideal point z^{id}.
- So we need r to be finite — which is guaranteed if the **image set** $\mathcal{Z} = F(\mathcal{F})$ is bounded.

Assume all objective functions $f_1, \ldots, f_p$ are **bounded above and below**. Then the image set $\mathcal{Z}$ is bounded, and both the nadir-based and worst-case reference points are finite:

$$r_j = \max_{z \in \mathcal{Z}} z_j \; < \; +\infty, \qquad j = 1, \ldots, p.$$

Consequently, for any list $\mathcal{L}_k \subseteq \mathcal{Z}$:

$$0 \; \leq \; \mathrm{HV}(\mathcal{L}_k, \, r) \; \leq \; \prod_{j=1}^{p} \left(r_j - z_j^{\mathrm{id}} \right) \; < \; +\infty.$$

Step-size convergence: proof by contradiction

Theorem

Assume the objective functions are bounded above and below on $\mathcal{F}$. Let $\{(\mathcal{L}_k, \alpha_k)\}$ be the sequence generated by Algorithm 2.1 with the sufficient hypervolume increase condition. Then there exists a subsequence of iterates $\mathcal{K}$ such that $\alpha_k \to 0$ as $k \to \infty$, $k \in \mathcal{K}$.

Proof. Suppose by contradiction that **no** such subsequence exists.

Step-size convergence: proof by contradiction

Theorem

Assume the objective functions are bounded above and below on $\mathcal{F}$. Let $\{(\mathcal{L}_k, \alpha_k)\}$ be the sequence generated by Algorithm 2.1 with the sufficient hypervolume increase condition. Then there exists a subsequence of iterates $\mathcal{K}$ such that $\alpha_k \to 0$ as $k \to \infty$, $k \in \mathcal{K}$.

Proof. Suppose by contradiction that **no** such subsequence exists.

Then the step-size is eventually bounded away from zero: $\exists\, \bar{\alpha} > 0$ such that $\alpha_k \geq \bar{\alpha}$ for all k sufficiently large. Since ρ is nondecreasing:

$$\rho(\alpha_k) \;\geq\; \rho(\bar{\alpha}) \;=:\; \bar{\rho} \;>\; 0.$$

Step-size convergence: proof by contradiction

Theorem

Assume the objective functions are bounded above and below on $\mathcal{F}$. Let $\{(\mathcal{L}_k, \alpha_k)\}$ be the sequence generated by Algorithm 2.1 with the sufficient hypervolume increase condition. Then there exists a subsequence of iterates $\mathcal{K}$ such that $\alpha_k \to 0$ as $k \to \infty$, $k \in \mathcal{K}$.

Proof. Suppose by contradiction that **no** such subsequence exists.

Then the step-size is eventually bounded away from zero: $\exists\, \bar{\alpha} > 0$ such that $\alpha_k \geq \bar{\alpha}$ for all k sufficiently large. Since ρ is nondecreasing:

$$\rho(\alpha_k) \;\geq\; \rho(\bar{\alpha}) \;=:\; \bar{\rho} \;>\; 0.$$

Question: Can the number of successful iterations be finite?

Step-size convergence: proof by contradiction

Theorem

Assume the objective functions are bounded above and below on $\mathcal{F}$. Let $\{(\mathcal{L}_k, \alpha_k)\}$ be the sequence generated by Algorithm 2.1 with the sufficient hypervolume increase condition. Then there exists a subsequence of iterates $\mathcal{K}$ such that $\alpha_k \to 0$ as $k \to \infty$, $k \in \mathcal{K}$.

Proof. Suppose by contradiction that **no** such subsequence exists.

Then the step-size is eventually bounded away from zero: $\exists\, \bar{\alpha} > 0$ such that $\alpha_k \geq \bar{\alpha}$ for all k sufficiently large. Since ρ is nondecreasing:

$$\rho(\alpha_k) \;\geq\; \rho(\bar{\alpha}) \;=:\; \bar{\rho} \;>\; 0.$$

Question: Can the number of successful iterations be finite?

No. If successful iterations were finite, then from some index k_0 onwards *all* iterations are unsuccessful. But at each unsuccessful iteration $\alpha_k \leftarrow \beta\, \alpha_k$ with $\beta \in (0, 1)$, so $\alpha_k \to 0$, contradicting $\alpha_k \geq \bar{\alpha}$.

Step-size convergence: proof by contradiction (cont.)

We have shown the number of successful iterations must be **infinite**.

Question: If there are infinitely many successful iterations, each increasing $\mathrm{HV}(\mathcal{L}_k, r)$ by at least $\bar{\rho} > 0$, what can we conclude?

Step-size convergence: proof by contradiction (cont.)

We have shown the number of successful iterations must be **infinite**.

Question: If there are infinitely many successful iterations, each increasing $\mathrm{HV}(\mathcal{L}_k, r)$ by at least $\bar{\rho} > 0$, what can we conclude?

At every successful iteration $k \in \mathcal{S}$:

$$\mathrm{HV}(\mathcal{L}_{k+1},\, r) \;\geq\; \mathrm{HV}(\mathcal{L}_k,\, r) + \bar{\rho}.$$

Since $|\mathcal{S}| = +\infty$, summing over all successful iterations:

$$\mathrm{HV}(\mathcal{L}_k,\, r) \;\xrightarrow{k\to\infty}\; +\infty,$$

Step-size convergence: proof by contradiction (cont.)

We have shown the number of successful iterations must be **infinite**.

Question: If there are infinitely many successful iterations, each increasing $\mathrm{HV}(\mathcal{L}_k, r)$ by at least $\bar{\rho} > 0$, what can we conclude?

At every successful iteration $k \in \mathcal{S}$:

$$\mathrm{HV}(\mathcal{L}_{k+1},\, r) \;\geq\; \mathrm{HV}(\mathcal{L}_k,\, r) + \bar{\rho}.$$

Since $|\mathcal{S}| = +\infty$, summing over all successful iterations:

$$\mathrm{HV}(\mathcal{L}_k,\, r) \xrightarrow{k \to \infty} +\infty,$$

which contradicts the fact that the hypervolume of the Pareto Front is finite. □

Step-size convergence: proof by contradiction (cont.)

We have shown the number of successful iterations must be **infinite**.

Question: If there are infinitely many successful iterations, each increasing $\mathrm{HV}(\mathcal{L}_k, r)$ by at least $\bar{\rho} > 0$, what can we conclude?

At every successful iteration $k \in \mathcal{S}$:

$$\mathrm{HV}(\mathcal{L}_{k+1},\, r) \;\geq\; \mathrm{HV}(\mathcal{L}_k,\, r) + \bar{\rho}.$$

Since $|\mathcal{S}| = +\infty$, summing over all successful iterations:

$$\mathrm{HV}(\mathcal{L}_k,\, r) \;\xrightarrow{k\to\infty}\; +\infty,$$

which contradicts the fact that the hypervolume of the Pareto Front is finite. □

Remark: this proof avoids reasoning about individual objective components and reduces everything to the monotone convergence of a single scalar $\mathrm{HV}(\mathcal{L}_k, r)$.