CATOS: COMPUTE-ACCURACY TRADEOFF SCALARIZATION FOR MULTIOBJECTIVE OPTIMIZATION IN NEURAL ARCHITECTURE SEARCH

Anonymous authors
Paper under double-blind review

ABSTRACT

Neural architecture search (NAS) increasingly targets deployment on resource-constrained hardware such as mobile devices or edge accelerators, where accuracy must be weighed against efficiency metrics such as latency or FLOPs, motivating multiobjective optimization (MOO). Existing MOO methods for NAS rest on Pareto dominance, which induces only a partial ordering: whenever one architecture is more accurate and another more efficient, the two are incomparable, preventing a consistent ranking of tradeoffs across the search space. As a result, existing multi-objective NAS methods rely on heuristic ranking strategies built on Pareto-based indicators, which provide only indirect signals of tradeoffs and are sensitive to the sampled set of architectures. To address these limitations, we introduce CATOS (Compute-Accuracy Tradeoff Scalarization), a scalarization function built on novel mathematical constructs that ranks architectures under bi-objective tradeoffs and remains effective in low-sample regimes by leveraging a compute-accuracy scaling prior. Integrated into an evolutionary algorithm (CATOS-EA), our method outperforms state-of-the-art MOO algorithms as well as constrained approaches on NATS-Bench in nearly all settings. Notably, the constructs underlying CATOS transfer unchanged to the markedly different NLP benchmark HW-GPT-Bench, where the augmented CATOS outperforms the strongest baseline.

1 INTRODUCTION

The design of deep learning models is rarely a purely single-objective problem, but instead revolves around balancing several objectives, such as performance and computational cost. In practice, the most accurate model is seldom the preferred one; rather, practitioners often target models that trade off accuracy and efficiency, for example by maximizing accuracy under a latency constraint or minimizing cost subject to an accuracy requirement. This tradeoff becomes particularly critical in deployment on resource-constrained platforms such as mobile or edge devices, where strict limits on latency, memory, and compute must be respected. As a result, compute-accuracy tradeoffs are a central concern in neural architecture search (NAS) (Sukthanker et al., 2024b) as well as in model compression techniques such as pruning (Han et al., 2015) and quantization (Jacob et al., 2018).

To reason about such tradeoffs in NAS, the dominant conceptual tool is Pareto dominance (Deb, 2011), which multiobjective optimization (MOO) algorithms use to identify architectures that represent optimal tradeoffs – those where improving one objective necessarily worsens another. However, Pareto dominance induces only a *partial* ordering: whenever one architecture is more accurate and another more efficient, the two are mutually incomparable. Without additional preference information, it is therefore unclear how to rank or select among Pareto optimal architectures in practice.

In order to properly guide architecture search and balance the possibly contrary objectives, additional criteria are needed to select or rank architectures beyond pure Pareto dominance. Existing methods rely on hypervolume (Beume et al., 2007; Daulton et al., 2020) and related indicators, which often serve only as an indirect proxy for tradeoff, and depend solely on the sampled architectures, making them unstable in low-sample regimes. These limitations become particularly relevant in NAS, where architectures are typically evaluated within iterative search procedures under limited evaluation budgets (Elsken et al., 2019b). Rather than ranking models by their dominance relations, we use those

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

relations only to delimit a region of objective space, and score each model by its relative position within it – using a fitted compute-accuracy scaling prior (Rosenfeld et al., 2020) to interpolate across sparsely sampled regions.

Building on this idea, we introduce *CATOS* (Compute-Accuracy Tradeoff Scalarization), a tradeoff scalarization function for bi-objective compute-accuracy problems whose scaling prior assumes a power-law relationship between the objectives. At its core, CATOS is built on several novel geometric constructs: the *negative Pareto front*, the geometric mirror image of the classical Pareto front; *local ideal points*; and the *Pareto corridor*, the region in objective space within which CATOS is defined. A compute-accuracy scaling prior (Rosenfeld et al., 2020) is then fitted to the negative and positive fronts, yielding continuous boundaries even where samples are scarce.

We integrate CATOS into an evolutionary algorithm (EA) framework (Deb, 2011) for neural architecture search, which we call CATOS-EA. CATOS-EA outperforms constrained differential NAS approaches in nearly all settings on NATS-Bench (Dong et al., 2021), often by large margins on ImageNet16-120, while requiring only a single run per dataset rather than one per constraint. Against classical MOO baselines, it achieves the lowest hypervolume regret and epsilon-indicator (Zitzler et al., 2003) on the deployment-relevant region of NATS-Bench topology search space across all budgets and datasets, and almost always yields the lowest or second-lowest regret to the optimal value under constrained accuracy and latency evaluation. We further test the generalizability of CATOS's underlying constructs and its corridor-based scoring on the NLP benchmark HW-GPT-Bench (Sukthanker et al., 2024a), demonstrating flexibility across network types, tasks, and devices.

Our contributions can be summarized as follows:

- We introduce CATOS, a tradeoff scalarization function that applies dominance relations to delimit a region of objective space rather than to rank models directly, scoring each against a fitted compute-accuracy scaling prior to produce a continuous, sample-robust ranking.
- We introduce the formal constructs on which our corridor-based scoring rests: the negative Pareto front, local ideal points, and the Pareto corridor.
- We integrate CATOS into a multiobjective evolutionary NAS framework, CATOS-EA, improving over multiobjective and constrained baselines on NATS-Bench under deployment-oriented evaluation, and outperforming multiobjective baselines on HW-GPT.

2 Related Work

Scalarization functions for multiobjective optimization. Scalarization reduces multiple objectives to a scalar score, inducing a total preordering over candidate solutions (Miettinen, 1999). A prominent class of methods is the Achievement Scalarization Function (ASF) (Wierzbicki, 1980), which evaluates solutions based on their distance to a reference point and underpins decomposition-based and surrogate-assisted approaches such as MOEA/D (Zhang & Li, 2007) and ParEGO (Knowles, 2006). These methods rely on a preference geometry that is either specified by the user (Thiele et al., 2009; Said et al., 2010; Branke & Deb, 2005) – typically as reference points or weight vectors – or generated algorithmically (Das & Dennis, 1998; Qi et al., 2014), which works best when objectives share a comparable scale. Neither fits NAS well, where compute and accuracy span very different ranges and the achievable set of best tradeoffs is unknown a priori. CATOS instead derives its tradeoff geometry from a compute-accuracy scaling prior fitted on observed data during search, grounded in established scaling-law behavior of deep networks (Rosenfeld et al., 2020).

Evaluation of Multiobjective Evolutionary Algorithms in NAS. Multiobjective evolutionary algorithms (MOEAs) aim to approximate the set of Pareto optimal tradeoff solutions and are commonly evaluated using set-based indicators such as hypervolume (Deb, 2011), IGD, or the epsilon-indicator (Zitzler et al., 2003). These metrics primarily reward coverage of the explored solutions. However, in NAS, the primary interest is typically not full coverage of the architecture space, but identifying models that satisfy given efficiency requirements. As a result, many NAS studies employing MOEAs (Elsken et al., 2019a; Lu et al., 2019; Sinha et al., 2024; Zhang et al., 2020) ultimately report performance under a single constraint (e.g., latency or FLOPs). Next to this single-constraint evaluation, we apply loose performance filters (e.g., minimum accuracy and maximum latency) to the search space and compute Pareto-compliant set-based metrics on the resulting subset. This focuses evaluation on the deployment-relevant region of the search space rather than only isolated constraint settings.

Hardware-aware NAS. Hardware-aware NAS is generally divided into two paradigms: constrained search and multiobjective search. Constrained NAS spans several directions, two prominent ones being differentiable and evolutionary search. Differentiable approaches (e.g., ProxylessNAS (Cai et al., 2019), ConNAS (De Ruyter et al., 2026)) train a supernet under a fixed budget on latency, energy, or FLOPs and extract a matching architecture. Evolutionary approaches search architectures directly under a target constraint; OFA (Cai et al., 2020) is a notable variant that first trains a once-for-all supernet, then runs a constrained EA per deployment target to extract a matching sub-network. Both directions share the same limitation: the search is tied to a specific constraint, so changing the target requires rerunning the procedure. MOEAs – NSGA-II (Deb et al., 2002), MOEA/D (Zhang & Li, 2007), RVEA (Cheng et al., 2016), among others – take the opposite stance, approximating the full set of optimal tradeoffs in a single run that can serve any constraint post hoc. CATOS-EA is itself an MOEA, but we evaluate it against approaches from both paradigms: constrained differentiable methods (Section 4.2) and classical MOEAs (Section 4.3).

3 Multiobjective Optimization for Neural Architecture Search

3.1 Background

Multiobjective optimization is built on an objective function that maps a candidate solution to multiple, potentially conflicting objectives. In the context of NAS, this function assigns each architecture a vector of performance metrics, such as accuracy and latency, the architecture's *objective vector*. Without loss of generality, we assume all objectives are to be minimized. Objectives such as accuracy are expressed as error (i.e., $1 -$ accuracy). In the case of a single objective, the objective vector is in $\mathbb{R}$, whereas in the multiobjective setting it is in $\mathbb{R}^d$, with d being the number of objectives. The set of all achievable objective vectors induced by the search space defines the attainable set:

Definition 1 (Attainable Set)**.** Let $\mathcal{A}$ denote a set of architectures (e.g. a search space) and let $f\colon \mathcal{A} \to \mathbb{R}^d$ be an objective function mapping each architecture to a d-dimensional objective vector. The attainable set is given by

$$\mathcal{Y} \coloneqq f(\mathcal{A}) \subseteq \mathbb{R}^d.$$

When optimizing for multiple, potentially conflicting objectives, a total ordering over architectures cannot be naturally imposed. Instead, comparisons are based on the concept of *Pareto dominance* (Pareto, 1964). Intuitively, an architecture is preferred only if it is not worse in any objective and strictly better in at least one. In this case, one architecture is said to dominate the other.

Definition 2 (Pareto dominance)**.** Let $\alpha, \alpha' \in \mathcal{A}$ with $f(\alpha) = (f_1(\alpha), \ldots, f_d(\alpha))$. We say that α *dominates* α', written $\alpha \prec \alpha'$, iff

$$f_i(\alpha) \leq f_i(\alpha') \quad \forall i \in \{1, \ldots, d\}, \qquad \text{and} \qquad \exists j \in \{1, \ldots, d\} : f_j(\alpha) < f_j(\alpha').$$

We further introduce the concept of negative Pareto dominance, saying that α negatively dominates α', written $\alpha \prec^- \alpha'$, iff

$$f_i(\alpha) \geq f_i(\alpha') \quad \forall i \in \{1, \ldots, d\}, \qquad \text{and} \qquad \exists j \in \{1, \ldots, d\} : f_j(\alpha) > f_j(\alpha').$$

Negative dominance is the point-wise dual of standard dominance: $\alpha' \prec^- \alpha \Leftrightarrow \alpha' \succ \alpha \Leftrightarrow \alpha \prec \alpha'$. Negative dominance is a central element of CATOS and will be explicitly used in further definitions and proofs. We call it negative dominance because it is the standard dominance on the negated objectives: flipping the sign of every f_i turns $\prec$ into $\prec^-$. To distinguish the two notions, we henceforth refer to standard dominance as *positive Pareto dominance*, writing $\prec^+$ in place of $\prec$. All derived constructs are prefixed with *positive* or *negative* accordingly.

Architectures that are not dominated by any other element in $\mathcal{A}$ are of particular interest, as no alternative architecture is strictly better with respect to all objectives. Such architectures are referred to as *Pareto optimal*. The collection of all Pareto optimal architectures forms the Pareto optimal set, whose image in objective space is known as the *Pareto front*.

Definition 3 (Pareto optimal set and Pareto front)**.** A model $\alpha^\star \in \mathcal{A}$ is (positively) *Pareto optimal* iff there exists no $\alpha \in \mathcal{A}$ such that $\alpha \prec^+ \alpha^\star$. The *(positive) Pareto optimal set* is

$$\mathcal{P}^+_{\mathcal{A}} \coloneqq \{\alpha^\star \in \mathcal{A} \mid \nexists\, \alpha \in \mathcal{A} : \alpha \prec^+ \alpha^\star\},$$

and the *(positive) Pareto front* is its image in objective space, $\mathcal{F}_{\mathcal{A}}^{+} \coloneqq f(\mathcal{P}_{\mathcal{A}}^{+}) \subseteq \mathcal{Y}$. Replacing $\prec^{+}$ with $\prec^{-}$ analogously gives the *negative Pareto optimal set* $\mathcal{P}_{\mathcal{A}}^{-}$ and the *negative Pareto front* $\mathcal{F}_{\mathcal{A}}^{-} \coloneqq f(\mathcal{P}_{\mathcal{A}}^{-})$, the dual frontier, characterizing architectures that no other architecture is uniformly worse than.

The negative Pareto front complements the positive one as a reference bound, together delimiting the region in which CATOS operates. Each front is associated with a *Pareto boundary* – the dividing line between dominated and non-dominated regions of objective space, commonly visualized as a staircase (Mescoli et al., 2026); we formalize it in Section E.2 in the appendix. From here on we set $d = 2$. This reflects our deployment setting, where compute and accuracy are the two principal tradeoff axes (see Section 1), and avoids notational overhead with no corresponding gain in the constructions that follow.

To form a tight envelope around the positive and negative Pareto fronts, we introduce *local positive/negative ideal points*, a localized variant of the classical ideal point (Skulimowski, 1990) that restricts attention to pairs of non-dominated points.

Definition 4 (Local Ideal Points and Ideal Fronts)**.** Let $\mathcal{F}^{+} = \{y^1, \dots, y^k\}$ with $k \geq 2$ and $y_1^1 < \cdots < y_1^k$ (implying $y_2^1 > \cdots > y_2^k$), and let $\mathcal{F}^{-} = \{w^1, \dots, w^m\}$ with $m \geq 2$ and $w_1^1 < \cdots < w_1^m$ (implying $w_2^1 > \cdots > w_2^m$). The *positive* and *negative ideal fronts* are the sets of *local positive* and *local negative ideal points*, respectively, given by

$$\mathcal{L}^{+} = \{\ell_{+}^{i}\}_{i=1}^{k-1}, \qquad \ell_{+}^{i} := (y_1^i,\ y_2^{i+1}),$$
$$\mathcal{L}^{-} = \{\ell_{-}^{i}\}_{i=1}^{m-1}, \qquad \ell_{-}^{i} := (w_1^{i+1},\ w_2^{i}).$$

Local ideal points have two structural properties that justify treating $\mathcal{L}^{+}$ and $\mathcal{L}^{-}$ as fronts in their own right. First, each ℓ_{+}^{i} dominates the two Pareto points it was built from and analogously for ℓ_{-}^{i} under negative dominance. Second, the local ideal points are themselves mutually non-dominated, and are non-dominated by the original fronts (see Section E.3 for the proof). Together, these properties imply that $\mathcal{L}^{+}$ is the positive Pareto front of the augmented set $\mathcal{F}^{+} \cup \mathcal{L}^{+}$, and analogously $\mathcal{L}^{-}$ is the negative Pareto front of $\mathcal{F}^{-} \cup \mathcal{L}^{-}$.

Definition 5 (Pareto Corridor)**.** Let $\mathcal{S} \subset \mathbb{R}^2$ be a finite set of objective vectors, and let $\mathcal{U}, \mathcal{V} \subseteq \mathcal{S}$. The *Pareto corridor* induced by $\mathcal{U}$ and $\mathcal{V}$ is

$$\mathcal{C}(\mathcal{U}, \mathcal{V}) \ := \ \big\{z \in \mathbb{R}^2 \mid \exists\, u \in \mathcal{U} : u \prec^{+} z\big\} \ \cap \ \big\{z \in \mathbb{R}^2 \mid \exists\, v \in \mathcal{V} : v \prec^{-} z\big\}.$$

Informally, the Pareto corridor $\mathcal{C}(\mathcal{U}, \mathcal{V})$ consists of all points that are (positively) dominated by $\mathcal{U}$ while simultaneously negatively dominated by $\mathcal{V}$. The following proposition formalizes key properties of this construction (see Section E.1 for the proof).

Proposition 1 (Properties of the Pareto Corridor)**.** Let $\mathcal{S} \subset \mathbb{R}^2$ be finite with (positive) Pareto front $\mathcal{F}^{+} = \{y^1, \dots, y^k\}$ and negative Pareto front $\mathcal{F}^{-} = \{w^1, \dots, w^m\}$, indexed as in Definition 4, with $|\mathcal{F}^{+}|, |\mathcal{F}^{-}| \geq 2$. Let $\mathcal{L}^{+}, \mathcal{L}^{-}$ be the corresponding ideal fronts.

(i) $\mathcal{C}(\mathcal{F}^{+}, \mathcal{F}^{-})$ may be empty, but $\mathcal{C}(\mathcal{F}^{+}, \mathcal{F}^{-}) \subseteq \mathcal{C}(\mathcal{L}^{+}, \mathcal{L}^{-})$.

(ii) $\mathcal{C}(\mathcal{L}^{+}, \mathcal{L}^{-}) \supseteq (y_1^1, w_1^2) \times (y_2^2, w_2^1)$, an open rectangle with strictly positive side lengths.

(iii) $\mathcal{C}(\mathcal{L}^{+}, \mathcal{L}^{-}) \subseteq [y_1^1, w_1^m] \times [y_2^k, w_2^1]$.

Consequently, the Pareto corridor $\mathcal{C}(\mathcal{L}^{+}, \mathcal{L}^{-})$ is bounded and contains a two-dimensional region. By (i), the corridor over the original Pareto fronts is therefore also bounded. The minimal configuration of $\mathcal{C}(\mathcal{L}^{+}, \mathcal{L}^{-})$ requires exactly two (positively) non-dominated points in $\mathcal{S}$, and in this minimal case $\mathcal{F}^{+} = \mathcal{F}^{-}$ is permitted. The degenerate case $|\mathcal{F}^{+}| = |\mathcal{F}^{-}| = 1$ is treated separately in Section E.1.1, where an analogous bounded two-dimensional corridor exists whenever y^1 is strictly smaller than w^1 in both coordinates.

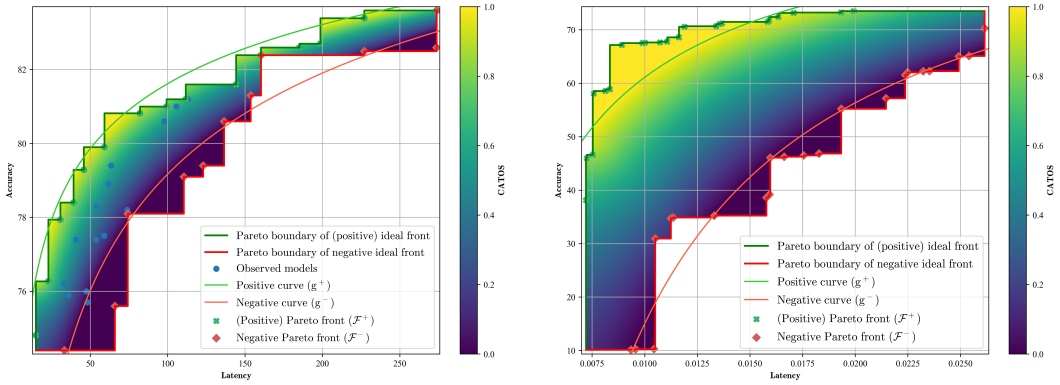
CATOS is therefore applicable as soon as $\mathcal{S}$ contains two non-dominated points, the minimal nontrivial configuration – making it applicable from the earliest stages of an architecture search. Moreover, the corridor over original Pareto fronts has an operational interpretation: a newly sampled architecture lies outside $\mathcal{C}(\mathcal{F}^{+}, \mathcal{F}^{-})$ if and only if it updates at least one of the two Pareto fronts.

Algorithm 1 CATOS-INIT

Require: Set of models $\mathcal{M} = \{m_i\}_{i=1}^N$, $m_i = (\ell_i, a_i) \in \mathbb{R}^2$
Ensure: Scores $\{s_i\}_{i=1}^N$, fitted curves g^+, g^-
$\mathcal{F}^+ \leftarrow$ POSITIVEPARETOFRONT($\mathcal{M}$)
$\mathcal{F}^- \leftarrow$ NEGATIVEPARETOFRONT($\mathcal{M}$)
$\mathcal{L}^+ \leftarrow$ POSITIVEIDEALPARETOFRONT($\mathcal{F}^+$)
$g^+ \leftarrow$ FITCURVE($\mathcal{L}^+$)
$g^- \leftarrow$ FITCURVE($\mathcal{F}^-$)
for $i = 1, \ldots, N$ **do**
  $s_i \leftarrow$ CATOS-SCORE(g^+, g^-, m_i)
end for
return $\left(\{s_i\}_{i=1}^N, g^+, g^-\right)$

Algorithm 2 CATOS-SCORE

Require: Curves g^+, g^-, model $m = (\ell, a) \in \mathbb{R}^2$
Ensure: Score $s \in [0, 1]$
if $a > g^+(\ell)$ **then**
  return 1
else if $a < g^-(\ell)$ **then**
  return 0
else
  $d_+ \leftarrow \min_{p \in g^+} \|m - p\|_2$
  $d_- \leftarrow \min_{p \in g^-} \|m - p\|_2$
  return $\dfrac{d_-}{d_+ + d_-}$
end if



(a) ImageNet, 38 trained models (Nottebaum et al., 2025).

(b) NATS-Bench topology on CIFAR-100.

Figure 1: **CATOS on two sets of models.** The shaded area represents the Pareto corridor $\mathcal{C}(\mathcal{L}^+, \mathcal{L}^-)$. Color encodes the CATOS score, ranging from 0 to 1.

3.2 CATOS FRAMEWORK

From now on, we depart from the pure minimization setting and instead aim to minimize cost (e.g., latency) while maximizing accuracy. The preceding definitions and results carry over unchanged under this sign convention, which aligns with visualization practice in the efficient-model literature (Nottebaum et al., 2025; Cai et al., 2023) (see Figure 1).

CATOS scores bi-objective vectors – pairs of (cost, accuracy) values obtained by evaluating architectures – by their position relative to fitted cost-accuracy frontiers. The framework consists of two sub-algorithms (see Algorithm 1, 2). CATOS-INIT takes a set of bi-objective vectors and fits two curves g^+ and g^- to the (positive) ideal front and the negative Pareto front respectively. CATOS-SCORE then assigns each vector a scalar score via its relative Euclidean position between g^+ and g^-. These distances are point-to-curve distances, computed by discretizing each curve into a polyline and projecting the vector onto its nearest segment via a k-d tree lookup. Consequently, scores grow with distance from g^- and shrink with distance from g^+. Vectors outside the area spanned by the two curves receive a fixed value of 0 (below g^-) or 1 (above g^+).

CATOS-SCORE is defined for all points in the Pareto corridor of the positive and negative ideal front: $\mathcal{C}(\mathcal{L}^+, \mathcal{L}^-)$. However, when a new model is added during search, refitting is required if the point lies outside the Pareto corridor $\mathcal{C}(\mathcal{F}^+, \mathcal{F}^-)$ – equivalently, if it extends one of the fronts; otherwise the existing curves remain valid and a single CATOS-SCORE call suffices.

A natural alternative would be to score against the Pareto boundaries themselves (see Section E.2, with a depiction of the Pareto boundary of an ideal front in Figure 1) rather than fitted curves, but the resulting distances would be dominated by the angular, staircase shape of the boundary rather than by

the underlying compute-accuracy tradeoff. The choice of curve family or interpolation method to yield g^+, g^- is therefore central to CATOS, and is addressed next.

Fitting a Scaling Prior. For FITCURVE we adopt the model-scaling form of Rosenfeld et al. (2020), $\varepsilon(m, n) \approx b(n)\, m^{-\beta(n)} + c_m(n)$, where m refers to the model size and n dataset size. We specialize it to our setting in three ways. First, since CATOS operates on a fixed dataset within a search, the data-dependent coefficients $b(n), \beta(n)$ become learnable scalars β, α. Second, we replace model size m with a generic cost axis (parameters, FLOPs, or latency). Finally, we aim for accuracy rather than error and fix the asymptotic ceiling to its natural upper bound of 100, removing a parameter that would be unreliable to fit on a 2–20 point Pareto front. This yields the two-parameter form

$$g(\text{cost}) \;=\; 100 - \beta \cdot \text{cost}^{-\alpha}. \tag{1}$$

Parameters α and β are estimated by least squares (FITCURVE, Algorithm 1), where each point's contribution is weighted by k-nearest neighbor distance (k = 3), such that the curve is not dominated by dense regions. Figure 1 visualizes CATOS scoring on two sets of models, together with all constructs underlying CATOS (Section D offers further notes on the visualization).

Properties. The scaling prior plays a dual role in CATOS: it keeps the method informative in the low-sample regime, and it ties CATOS to compute (e.g., parameters, FLOPs, latency) and accuracy as its objectives, assuming a power-law relationship between them. Beyond this, CATOS is architecture- and search-space-agnostic, relying only on observed objective values.

3.3 CATOS-EA: EMBEDDING CATOS IN EVOLUTIONARY SEARCH

To perform neural architecture search with CATOS, we embed it into a multiobjective evolutionary algorithm (EA) framework we call CATOS-EA. EAs maintain a population of candidate solutions (e.g., architectures) and iteratively replace weaker candidates with offspring derived from stronger ones, concentrating the population around promising regions of the search space. CATOS-EA is *selection-driven* (Deb et al., 2002): CATOS scoring specifies only how architectures are ranked for parent selection and survival, leaving offspring generation (mutation and crossover) to the search space. Each architecture must therefore map to a (cost, accuracy) pair, but is otherwise unconstrained.

Search Procedure. The population is initialized by sampling architectures uniformly from the search space (Deb et al., 2002). Each generation has three steps. (i) *Parent selection* via tournament: a random subset of the population is sampled, and the architecture with the highest CATOS score is chosen; this is repeated to select as many parents as needed. (ii) *Offspring generation* produces children from the parents using any search-space-appropriate operator. (iii) *Survival selection* removes architectures with the lowest CATOS scores until the population returns to its target size.

4 EXPERIMENTS AND RESULTS

4.1 SETTINGS

In the following we specify the setting for experiments on NATS-Bench (Dong et al., 2021) in Sections 4.2 and 4.3 and on HW-GPT-Bench (Sukthanker et al., 2024a) in Figure 2. For standard MOO baselines (NSGA-II, MOEA/D, etc.) we use the pymoo[1] (Blank & Deb, 2020) implementation.

NATS-Bench protocol. NATS-Bench (Dong et al., 2021) comprises two sub-spaces: *topology* (varying cell operations, also called NAS-Bench-201 (Dong & Yang, 2020)) and *size* (varying per-layer channel widths on a fixed cell topology). We apply CATOS-EA to NATS-Bench, comparing against standard MOO baselines (NSGA-II, MOEA/D, etc.) via naslib (Mehta et al., 2022; Ruchte et al., 2020). Objectives are accuracy and batch-size-256 latency, both taken from (Dong et al., 2021). We consider three search budgets defined by the number of sampled models – *low* (60), *mid* (300), *high* (500) – and run each on CIFAR-10, CIFAR-100 (Krizhevsky, 2009), and ImageNet16-120 (Chrabaszcz et al., 2017). Crossover is disabled. Each generation samples one parent from a population of 10/30/50 architectures (*low*/*mid*/*high*) and produces one child by mutation, with operators following the standard procedure of the search space. When sampling or mutation yields a previously evaluated model, we resample up to 5 times before accepting a duplicate. Every

[1] `https://pymoo.org/`

configuration is repeated 100 times, and we report mean and standard deviation. For constrained evaluation (best accuracy/latency under constraint), we report regret to the optimal value (Carpentier & Valko, 2015) – the difference to the optimal feasible model – computed on the full search trace.

The constrained-differential comparison in Section 4.2 runs on both NATS-Bench sub-spaces at the mid budget (300 models). CATOS-EA does not optimize for any of the listed constraints in Section 4.2, and its trace is filtered post-hoc per constraint. The classical MOO comparison in Section 4.3 runs on the topology sub-space.

HW-GPT-Bench protocol. For Figure 2, we use perplexity and A100 latency as objectives, repeat each run five times, and use mutation for offspring generation with a population size of 300.

4.2 Comparison with constrained differential approaches

In Table 1 we compare CATOS-EA against constrained differential methods: ConNAS (De Ruyter et al., 2026), HDX (Hong et al., 2022), ProxylessNAS (Cai et al., 2019), and TF-NAS (Hu et al., 2020). These methods train one supernet per target constraint, while a single CATOS-EA run yields a Pareto front reusable across arbitrary constraints. Each reported accuracy from the constrained differential methods in Table 1 thus corresponds to one supernet training, while all CATOS-EA entries come from the same six configurations (one per dataset/search-space, 100 repeats) with traces filtered post-hoc. CATOS-EA does not optimize for any specific constraint, but almost consistently outperforms compared approaches. Although CATOS-EA's budget is 300 models in this setting, on the size search space it beats the next-best baseline after having seen 41 models on average.

Table 1: **NATS-Bench comparison with constrained differential methods.** CATOS-EA (mid setting, 300 models) is run once per dataset/search-space. Constraints: latency and energy on an NVIDIA Jetson TX2 Edge GPU (Li et al., 2021) for topology (left); parameters, peak memory, and FLOPs (De Ruyter et al., 2026) for size (right). Best method per (dataset, constraint) in bold.

	Topology			Size		
Method	**CIFAR-10**	**CIFAR-100**	**IN16-120**	**CIFAR-10**	**CIFAR-100**	**IN16-120**
Constraint:	*max accuracy s.t.* latency ≤ 5.31ms			*max accuracy s.t.* #params $\leq 261{,}650$		
ProxylessNAS	92.29 ± 0.00	67.48 ± 0.00	39.40 ± 0.00	85.87 ± 1.16	51.11 ± 2.35	25.27 ± 2.13
TF-NAS	93.27 ± 0.00	70.06 ± 0.00	39.75 ± 0.00	91.78 ± 0.00	63.64 ± 0.00	38.00 ± 0.00
HDX	75.83 ± 36.80	54.18 ± 29.73	31.69 ± 17.25	86.10 ± 1.28	51.71 ± 2.88	25.36 ± 2.42
ConNAS	93.18 ± 0.09	70.12 ± 0.16	40.59 ± 1.20	91.74 ± 0.29	66.22 ± 0.47	39.77 ± 0.77
CATOS-EA (ours)	$\mathbf{93.83 \pm 0.25}$	$\mathbf{71.13 \pm 0.62}$	$\mathbf{45.71 \pm 0.92}$	$\mathbf{92.05 \pm 0.17}$	$\mathbf{66.94 \pm 0.51}$	$\mathbf{40.29 \pm 0.78}$
Constraint:	*max accuracy s.t.* energy ≤ 23.95mJ			*max accuracy s.t.* peak mem ≤ 655kB		
ProxylessNAS	92.29 ± 0.00	67.48 ± 0.00	39.40 ± 0.00	91.33 ± 0.19	68.22 ± 0.28	38.49 ± 0.03
TF-NAS	93.01 ± 0.57	69.48 ± 1.29	39.62 ± 0.30	92.99 ± 0.10	$\mathbf{70.21 \pm 0.16}$	$\mathbf{45.25 \pm 0.27}$
HDX	92.29 ± 0.00	67.48 ± 0.00	39.40 ± 0.00	91.59 ± 0.19	68.03 ± 0.29	38.41 ± 0.07
ConNAS	92.97 ± 0.54	69.51 ± 1.48	40.32 ± 1.56	$\mathbf{93.28 \pm 0.15}$	69.70 ± 1.02	44.58 ± 0.43
CATOS-EA (ours)	$\mathbf{93.80 \pm 0.29}$	$\mathbf{70.99 \pm 0.62}$	$\mathbf{45.33 \pm 0.91}$	93.17 ± 0.23	68.74 ± 0.58	45.02 ± 0.74
Constraint:	*max accuracy s.t.* latency $\wedge$ energy			*max accuracy s.t.* #FLOPs ≤ 1196M		
ProxylessNAS	92.29 ± 0.00	67.48 ± 0.00	39.40 ± 0.00	89.55 ± 0.22	63.51 ± 0.67	32.43 ± 0.84
TF-NAS	93.27 ± 0.00	70.06 ± 0.00	39.75 ± 0.00	91.10 ± 0.25	66.56 ± 1.12	37.41 ± 0.94
HDX	92.29 ± 0.00	67.48 ± 0.00	39.40 ± 0.00	84.43 ± 0.99	49.02 ± 0.85	23.17 ± 1.51
ConNAS	93.14 ± 0.05	70.21 ± 0.17	41.45 ± 1.01	91.71 ± 0.28	65.44 ± 0.80	38.33 ± 0.36
CATOS-EA (ours)	$\mathbf{93.80 \pm 0.29}$	$\mathbf{70.98 \pm 0.61}$	$\mathbf{45.33 \pm 0.91}$	$\mathbf{92.55 \pm 0.18}$	$\mathbf{66.68 \pm 0.50}$	$\mathbf{42.25 \pm 0.72}$

4.3 Comparison with classical multiobjective optimization approaches

In Tables 2 and 3 we compare CATOS-EA against MOEA/D (Zhang & Li, 2007), NSGA-II (Deb et al., 2002), NSGA-III (Deb & Jain, 2013), R-NSGA-II (Deb & Sundar, 2006), and RVEA (Cheng et al., 2016). Both tables come from the same runs, reporting metrics on the final Pareto front (Table 2) and on the full trace (Table 3).

Reduced Pareto front evaluation. Not all parts of the Pareto front are equally interesting for deployment, but MOO algorithms enforce coverage across regions unlikely to be useful in practice. In Table 2 we therefore restrict the full NATS-Bench topology (Dong et al., 2021) benchmark to models above the 25th-quantile accuracy (CIFAR-10: $\geq$87.6%, CIFAR-100: $\geq$59.2%, ImageNet16-120: $\geq$29.4%) and below the 75th-quantile latency, retaining roughly half, and recompute the ground truth Pareto front on the restricted set. Restricting the benchmark sacrifices little accuracy: the best

Table 2: **Pareto-compliant evaluation against classical MOO algorithms.** We report Hypervolume Regret (HV-Reg., gap to the restricted ground-truth Pareto front) and the Epsilon-Indicator (ϵ-Ind.) on NATS-Bench topology. Lower is better for both. **Best** and second-best per (budget, metric, dataset).

Budget	Method	*Architectures filtered by* $\mathrm{lat.} \leq Q_{75}$ *and* $\mathrm{acc.} \geq Q_{25}$ *(∼50% of models)*					
		CIFAR-10		**CIFAR-100**		**ImageNet16-120**	
		HV-Reg.↓	ϵ-Ind.↓	HV-Reg.↓	ϵ-Ind.↓	HV-Reg.↓	ϵ-Ind.↓
Low	MOEA/D	0.495±.177	0.734±.213	0.614±.154	1.357±.501	0.644±.142	0.879±.226
	NSGA-III	0.163±.050	0.249±.075	0.294±.068	0.482±.122	0.301±.088	0.367±.093
	NSGA-II	0.169±.056	0.281±.082	0.299±.073	0.494±.129	0.297±.075	0.364±.076
	R-NSGA-II	0.148±.055	0.238±.094	0.294±.080	0.496±.149	0.305±.085	0.372±.094
	RVEA	0.309±.139	0.420±.164	0.426±.122	0.601±.159	0.433±.126	0.540±.170
	Ours	**0.120**±.059	**0.190**±.080	**0.265**±.077	**0.423**±.129	**0.263**±.088	**0.341**±.107
Mid	MOEA/D	0.457±.130	0.675±.183	0.599±.122	1.312±.585	0.630±.151	0.888±.241
	NSGA-III	0.055±.018	0.111±.031	0.145±.050	0.254±.053	0.118±.035	0.197±.055
	NSGA-II	0.054±.018	0.113±.036	0.145±.049	0.267±.060	0.134±.037	0.210±.047
	R-NSGA-II	0.055±.020	0.121±.042	0.139±.045	0.258±.058	0.116±.044	0.184±.049
	RVEA	0.100±.039	0.174±.057	0.203±.060	0.344±.086	0.179±.066	0.255±.084
	Ours	**0.042**±.020	**0.092**±.041	**0.110**±.051	**0.240**±.073	**0.097**±.041	**0.160**±.046
High	MOEA/D	0.545±.201	0.784±.229	0.587±.109	1.315±.465	0.591±.148	0.851±.228
	NSGA-III	0.040±.016	0.085±.026	0.114±.046	0.221±.052	0.089±.035	0.167±.040
	NSGA-II	0.042±.016	0.094±.028	0.117±.035	0.234±.044	0.092±.033	0.179±.050
	R-NSGA-II	0.037±.014	0.086±.031	0.101±.039	0.214±.049	0.084±.032	0.162±.042
	RVEA	0.084±.028	0.150±.045	0.163±.049	0.276±.062	0.133±.048	0.215±.067
	Ours	**0.027**±.014	**0.068**±.039	**0.083**±.045	**0.213**±.056	**0.063**±.032	**0.127**±.043

Table 3: **Multiobjective Optimization – Constrained Evaluation.** Regret to the optimum on NATS-Bench topology search space (Dong et al., 2021) for maximizing accuracy s.t. latency $\leq Q_{50}$ and minimizing latency s.t. accuracy $\geq Q_{75}$. Lower is better; **best** and second-best per column.

Budget	Method	**CIFAR-10**		**CIFAR-100**		**ImageNet16-120**	
		Δacc (pp)↓ s.t. lat $\leq Q_{50}$	Δlat (ms)↓ s.t. acc $\geq Q_{75}$	Δacc (pp)↓ s.t. lat $\leq Q_{50}$	Δlat (ms)↓ s.t. acc $\geq Q_{75}$	Δacc (pp)↓ s.t. lat $\leq Q_{50}$	Δlat (ms)↓ s.t. acc $\geq Q_{75}$
Low	MOEA/D	1.66±1.24	3.9±1.3	2.83±2.06	3.1±1.2	4.04±2.64	3.6±1.2
	NSGA-III	1.11±0.58	3.0±1.1	2.17±1.16	2.2±1.2	3.52±1.57	2.8±1.3
	NSGA-II	1.29±0.58	3.3±1.3	2.20±1.19	2.4±1.2	3.36±1.58	2.9±1.2
	R-NSGA-II	1.07±0.59	2.8±1.0	2.05±1.07	2.0±1.0	3.01±1.44	2.5±1.0
	RVEA	1.06±0.67	3.1±1.1	2.23±1.40	2.5±1.1	3.36±1.78	3.2±1.1
	Ours	**0.71**±0.57	**2.3**±1.0	**1.33**±1.00	**1.9**±1.0	**2.65**±1.31	**2.2**±1.0
Mid	MOEA/D	0.90±0.77	3.2±0.9	**0.35**±0.78	2.1±0.9	2.52±1.59	2.8±0.9
	NSGA-III	0.22±0.23	1.5±0.9	0.53±0.47	0.7±0.7	1.40±0.70	1.2±0.8
	NSGA-II	0.26±0.29	1.8±0.8	0.62±0.68	0.9±0.7	1.55±0.77	1.2±0.8
	R-NSGA-II	0.33±0.36	1.3±0.9	0.60±0.59	**0.5**±0.7	**1.10**±0.73	1.1±0.7
	RVEA	**0.19**±0.18	2.0±0.7	0.57±0.52	0.9±0.7	1.33±0.84	1.5±0.8
	Ours	0.22±0.35	**0.9**±0.9	0.39±0.55	0.6±0.4	1.17±0.63	**0.8**±0.6
High	MOEA/D	0.50±0.54	2.7±0.7	**0.11**±0.44	1.8±0.7	1.61±1.14	2.3±0.9
	NSGA-III	**0.11**±0.08	1.1±0.8	0.37±0.44	0.3±0.5	1.12±0.68	0.8±0.6
	NSGA-II	0.14±0.16	1.3±0.9	0.30±0.43	0.5±0.6	1.09±0.72	0.9±0.7
	R-NSGA-II	0.19±0.17	0.8±0.7	0.30±0.35	**0.0**±0.5	**0.80**±0.63	0.8±0.6
	RVEA	0.16±0.15	1.5±0.9	0.28±0.45	0.6±0.6	1.15±0.74	1.1±0.6
	Ours	0.15±0.25	**0.5**±0.6	0.22±0.22	0.4±0.3	1.04±0.50	**0.5**±0.5

model below the 75th-quantile latency is within 0.36% / 1.09% / 0.14% of the unrestricted maximum (CIFAR-10 / CIFAR-100 / ImageNet16-120). CATOS-EA outperforms all baselines in hypervolume and epsilon-indicator, with consistent gains across budgets and datasets.

Constrained latency-accuracy evaluation. On the search trace, we evaluate the highest-accuracy model under a latency constraint and the lowest-latency model under an accuracy constraint. In Table 3, CATOS-EA consistently achieves the lowest or second-lowest regret across nearly all settings.

4.4 Multiobjective Comparison on HW-GPT-Bench

To test whether our corridor-based scoring generalizes beyond its design setting, we apply CATOS to the supernet surrogate benchmark HW-GPT-Bench (Sukthanker et al., 2024a), which differs in nearly

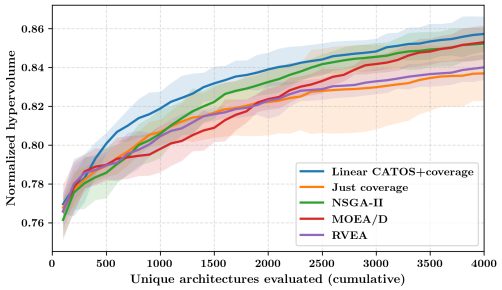


Figure 2: **Linear CATOS+coverage on HW-GPT.** Mean normalized hypervolume ± 2 SD. "Just coverage" uses only HV and NN-dist.

Table 4: **Curve ablation.** Mean±std over 100 runs on NATS-Bench topology. GT-PF is on the filtered set, Δacc/Δlat on the unfiltered.

Metric	Method	C-10	C-100	IN
Δacc (pp) ↓	RBF	1.21±.64	1.98±1.19	3.52±1.79
	Linear	0.88±.60	1.60±1.16	2.84±1.44
	CATOS	**0.71±.57**	**1.33±1.00**	**2.65±1.31**
Δlat (ms) ↓	RBF	2.9±1.2	2.1±1.1	2.8±1.2
	Linear	2.5±1.0	**1.7±1.1**	2.4±1.1
	CATOS	**2.3±1.0**	1.9±1.0	**2.2±1.0**
GT-PF (%) ↑	RBF	1.28±2.49	1.32±3.10	1.06±2.49
	Linear	1.95±3.77	1.85±4.88	1.43±2.82
	CATOS	**2.26±3.43**	**2.00±3.86**	**1.81±3.55**

every respect: transformer architectures, perplexity instead of accuracy, and surrogate A100 latencies. It also presents a different search regime: a partitioned objective space, a much larger search space (requiring a higher evaluation budget) and a denser Pareto front. The partitioned objective space and higher evaluation budget motivate a domain-specific change: we replace the scaling prior's power-law fit, which is most beneficial in low-sample regimes, with a piecewise-linear interpolant. The formulation permits this by construction, since CATOS requires only an interpolant of the positive and negative fronts, not a specific functional form. In this larger-space, denser-front regime, we also combine CATOS with coverage-oriented signals in a complementary split: CATOS carries the quality signal through its tradeoff score, while hypervolume (survival) and nearest-neighbor distance (parent selection) supply coverage and diversity, fused via rank-sum. Augmented CATOS surpasses every compared method in Figure 2, including HW-GPT-Bench's (Sukthanker et al., 2024a) strongest baseline, NSGA-II. The coverage-oriented signals alone perform considerably worse, confirming that CATOS's corridor-based scoring provides the decisive quality signal.

4.5 Ablation Study

We ablate the choice of g^+, g^- in CATOS by replacing the scaling prior with two alternatives: a thin-plate RBF kernel (RBF in Table 4) fit to all sampled points, and piecewise linear (linear in Table 4) interpolation of the positive ideal front and the negative Pareto front. We follow the NATS-Bench protocol of Section 4.1 and report results at the low budget, the regime in which a structural prior is most informative due to sparse sampling. Table 4 shows the constraint regrets Δacc and Δlat (Table 3 protocol) on the unfiltered benchmark and the percentage of ground-truth Pareto points found (GT-PF) during search on the filtered benchmark (filtered as in Table 2). The scaling prior achieves the lowest constraint regret on every dataset and recovers more Pareto points than either alternative, with piecewise linear interpolation in turn outperforming the thin-plate RBF.

5 Conclusion

We introduced *CATOS*, a tradeoff scalarization for bi-objective compute-accuracy problems, together with the formal constructs it rests on: the negative Pareto front, local ideal points, and the Pareto corridor. Our corridor-based architecture scoring yields a continuous ranking that remains stable in low-sample regimes due to the scaling prior. This makes CATOS especially suitable for neural architecture search with expensive model evaluations. Embedded in an evolutionary search, CATOS-EA delivers strong results against both constrained and multiobjective baselines, while its transfer to the markedly different HW-GPT-Bench demonstrates the generalizability of the corridor-based scoring. CATOS as presented targets the bi-objective compute-accuracy setting, but extending it to higher-dimensional tradeoffs or further substituting priors tailored to settings beyond compute-accuracy is a natural future direction (see Section A.1 for more details).

AI USE STATEMENT

In this work, we used generative AI tools to assist with writing proofs. For each proof, we specified the proof strategy and approach, an AI tool produced an initial write-up following our instructions, which we then revised into its final form. We also used AI tools to improve the readability of the proofs and to check them for errors. Beyond the proofs, we used AI tools to get feedback on our research methodology and on the positioning of our work, and to generate code helping to implement and evaluate our methods. We did not use generative AI tools to formulate full mathematical claims, provide critical ingredients for proofs, propose or refine hypotheses, develop theoretical models or conceptual frameworks, generate synthetic data, clean or reformat datasets, assist with translation, support qualitative or thematic data analysis, or interpret results. Additionally, we used AI tools to edit the paper for readability, identify relevant literature, brainstorm, review code for errors, format parts of the paper in LaTeX, and critique the manuscript, including spotting minor mistakes and assessing how clearly our proposed method is explained.

We reviewed all AI-assisted work and did not accept any output without independent verification. Several of us checked each proof by hand. We reviewed and tested all AI-generated code before using it in any reported experiment. We used AI tools as a search aid to find potentially relevant literature; we read every cited work in its original form, and all descriptions of related work reflect our own reading of the papers. We treated AI critique as one input among others and decided ourselves which suggestions to act on. We take full responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

REFERENCES

Jon Louis Bentley. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18(9):509–517, 1975.

Nicola Beume, Boris Naujoks, and Michael Emmerich. SMS-EMOA: Multiobjective selection based on dominated hypervolume. *European journal of operational research*, 181(3):1653–1669, 2007.

Julian Blank and Kalyanmoy Deb. Pymoo: Multi-objective optimization in python. *Ieee access*, 8: 89497–89509, 2020.

Jürgen Branke and Kalyanmoy Deb. Integrating user preferences into evolutionary multi-objective optimization. In *Knowledge incorporation in evolutionary computation*, pp. 461–477. Springer, 2005.

Han Cai, Ligeng Zhu, and Song Han. Proxylessnas: Direct neural architecture search on target task and hardware. In *International Conference on Learning Representations (ICLR)*, 2019.

Han Cai, Chuang Gan, Tianzhe Wang, Zhekai Zhang, and Song Han. Once-for-all: Train one network and specialize it for efficient deployment. In *International Conference on Learning Representations (ICLR)*, 2020.

Han Cai, Junyan Li, Muyan Hu, Chuang Gan, and Song Han. Efficientvit: Lightweight multi-scale attention for high-resolution dense prediction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 17302–17313, 2023.

Alexandra Carpentier and Michal Valko. Simple regret for infinitely many armed bandits. In *International Conference on Machine Learning*, pp. 1133–1141. PMLR, 2015.

Ran Cheng, Yaochu Jin, Markus Olhofer, and Bernhard Sendhoff. A reference vector guided evolutionary algorithm for many-objective optimization. *IEEE transactions on evolutionary computation*, 20(5):773–791, 2016.

Patryk Chrabaszcz, Ilya Loshchilov, and Frank Hutter. A downsampled variant of ImageNet as an alternative to the CIFAR datasets. *arXiv preprint arXiv:1707.08819*, 2017.

Indraneel Das and John E Dennis. Normal-boundary intersection: A new method for generating the pareto surface in nonlinear multicriteria optimization problems. *SIAM journal on optimization*, 8 (3):631–657, 1998.

Samuel Daulton, Maximilian Balandat, and Eytan Bakshy. Differentiable expected hypervolume improvement for parallel multi-objective bayesian optimization. *Advances in neural information processing systems*, 33:9851–9864, 2020.

Gregory De Ruyter, Mathias Verbeke, and Hans Hallez. Constraint-guided hardware-aware nas through gradient modification. In *The Fourteenth International Conference on Learning Representations*, 2026.

Kalyanmoy Deb. Multi-objective optimisation using evolutionary algorithms: an introduction. In *Multi-objective evolutionary optimisation for product design and manufacturing*, pp. 3–34. Springer, 2011.

Kalyanmoy Deb and Himanshu Jain. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: solving problems with box constraints. *IEEE transactions on evolutionary computation*, 18(4):577–601, 2013.

Kalyanmoy Deb and Jayavelmurugan Sundar. Reference point based multi-objective optimization using evolutionary algorithms. In *Proceedings of the 8th annual conference on Genetic and evolutionary computation*, pp. 635–642, 2006.

Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and TAMT Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation*, 6(2):182–197, 2002.

Xuanyi Dong and Yi Yang. NAS-Bench-201: Extending the scope of reproducible neural architecture search. In *International Conference on Learning Representations*, 2020.

Xuanyi Dong, Lu Liu, Katarzyna Musial, and Bogdan Gabrys. Nats-bench: Benchmarking nas algorithms for architecture topology and size. *IEEE transactions on pattern analysis and machine intelligence*, 44(7):3634–3646, 2021.

Thomas Elsken, Jan Hendrik Metzen, and Frank Hutter. Efficient multi-objective neural architecture search via lamarckian evolution. In *International Conference on Learning Representations (ICLR)*, 2019a.

Thomas Elsken, Jan Hendrik Metzen, and Frank Hutter. Neural architecture search: A survey. *Journal of Machine Learning Research*, 20(55):1–21, 2019b.

Song Han, Jeff Pool, John Tran, and William Dally. Learning both weights and connections for efficient neural network. *Advances in neural information processing systems*, 28, 2015.

Deokki Hong, Kanghyun Choi, Hye Yoon Lee, Joonsang Yu, Noseong Park, Youngsok Kim, and Jinho Lee. Enabling hard constraints in differentiable neural network and accelerator co-exploration. In *Proceedings of the 59th ACM/IEEE Design Automation Conference*, pp. 589–594, 2022.

Yibo Hu, Xiang Wu, and Ran He. TF-NAS: Rethinking three search freedoms of latency-constrained differentiable neural architecture search. In *European conference on computer vision*, pp. 123–139. Springer, 2020.

Benoit Jacob, Skirmantas Kligys, Bo Chen, Menglong Zhu, Matthew Tang, Andrew Howard, Hartwig Adam, and Dmitry Kalenichenko. Quantization and training of neural networks for efficient integer-arithmetic-only inference. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2704–2713, 2018.

Joshua Knowles. ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE transactions on evolutionary computation*, 10(1): 50–66, 2006.

Alex Krizhevsky. Learning multiple layers of features from tiny images. Technical report, University of Toronto, 2009.

Chaojian Li, Zhongzhi Yu, Yonggan Fu, Yongan Zhang, Yang Zhao, Haoran You, Qixuan Yu, Yue Wang, and Yingyan Lin. HW-NAS-Bench: Hardware-aware neural architecture search benchmark. In *The 9th International Conference on Learning Representations 2021 (ICLR 2021)*, 2021.

Zhichao Lu, Ian Whalen, Vishnu Boddeti, Yashesh Dhebar, Kalyanmoy Deb, Erik Goodman, and Wolfgang Banzhaf. Nsga-net: neural architecture search using multi-objective genetic algorithm. In *Proceedings of the genetic and evolutionary computation conference*, pp. 419–427, 2019.

Yash Mehta, Colin White, Arber Zela, Arjun Krishnakumar, Guri Zabergja, Shakiba Moradian, Mahmoud Safari, Kaicheng Yu, and Frank Hutter. NAS-Bench-Suite: Nas evaluation is (now) surprisingly easy. In *International Conference on Learning Representations*, 2022.

Riccardo Mescoli, Claudia Canali, and Riccardo Lancellotti. Investigating pareto front multi-objective optimization for edge computing service placement. *Computer Networks*, pp. 112213, 2026.

Kaisa Miettinen. *Nonlinear multiobjective optimization*, volume 12. Springer Science & Business Media, 1999.

Moritz Nottebaum, Matteo Dunnhofer, and Christian Micheloni. LowFormer: Hardware efficient design for convolutional transformer backbones. In *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 7008–7018. IEEE, 2025.

Vilfredo Pareto. *Cours d'économie politique*, volume 1. Librairie Droz, 1964.

Yutao Qi, Xiaoliang Ma, Fang Liu, Licheng Jiao, Jianyong Sun, and Jianshe Wu. Moea/d with adaptive weight adjustment. *Evolutionary Computation*, 22(2):231–264, 2014.

Jonathan S. Rosenfeld, Amir Rosenfeld, Yonatan Belinkov, and Nir Shavit. A constructive prediction of the generalization error across scales. In *International Conference on Learning Representations (ICLR)*, 2020.

Michael Ruchte, Arber Zela, Julien Siems, Josif Grabocka, and Frank Hutter. NASLib: A modular and flexible neural architecture search library. `https://github.com/automl/NASLib`, 2020.

Lamjed Ben Said, Slim Bechikh, and Khaled Ghédira. The r-dominance: a new dominance relation for interactive evolutionary multicriteria decision making. *IEEE transactions on Evolutionary Computation*, 14(5):801–818, 2010.

Nilotpal Sinha, Abd El Rahman Shabayek, Anis Kacem, Peyman Rostami, Carl Shneider, and Djamila Aouada. Hardware aware evolutionary neural architecture search using representation similarity metric. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 2628–2637, 2024.

Andrzej Skulimowski. Applicability of ideal points in multicriteria decision-making. In *Proceedings of the 9th International Conference on Multiple Criteria Decision-Making, Fairfax, USA*, pp. 5–8, 1990.

Rhea S Sukthanker, Arber Zela, Benedikt Staffler, Aaron Klein, Lennart Purucker, Jörg K Franke, and Frank Hutter. Hw-gpt-bench: hardware-aware architecture benchmark for language models. *Advances in Neural Information Processing Systems*, 37:60776–60834, 2024a.

Rhea Sanjay Sukthanker, Arber Zela, Benedikt Staffler, Samuel Dooley, Josif Grabocka, and Frank Hutter. Multi-objective differentiable neural architecture search. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2024b.

Lothar Thiele, Kaisa Miettinen, Pekka J Korhonen, and Julian Molina. A preference-based evolutionary algorithm for multi-objective optimization. *Evolutionary computation*, 17(3):411–436, 2009.

Andrzej P Wierzbicki. The use of reference objectives in multiobjective optimization. In *Multiple criteria decision making theory and application: Proceedings of the third conference Hagen/Königswinter, West Germany, August 20–24, 1979*, pp. 468–486. Springer, 1980.

Li Lyna Zhang, Yuqing Yang, Yuhang Jiang, Wenwu Zhu, and Yunxin Liu. Fast hardware-aware neural architecture search. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 692–693, 2020.

Qingfu Zhang and Hui Li. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Transactions on evolutionary computation*, 11(6):712–731, 2007.

Eckart Zitzler, Lothar Thiele, Marco Laumanns, Carlos M Fonseca, and Viviane Grunert Da Fonseca. Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Transactions on evolutionary computation*, 7(2):117–132, 2003.

A Additional Considerations to CATOS

A.1 Discussion of Limitations and Future Work

Scope. CATOS currently handles two objectives; extending the proposed constructs (local ideal points, negative Pareto front and the Pareto corridor) to higher-dimensional tradeoffs is natural future work.

Modelling assumptions. While the constructs are indifferent to the choice of objectives, the scaling prior assumes a power-law relationship, which need not hold across all modalities, datasets or architectures. Further, if the search space is disconnected and no single curve fits, the prior can lead to deteriorated scoring. A general approximation of each Pareto front via piecewise linear (see Figure 2) or RBF interpolation would make it applicable to any two objectives, but Table 4 shows these interpolations perform worse in the low-sample regime.

Method behaviour. While CATOS induces a total order over tradeoffs, it does not inherently guarantee coverage of the full Pareto front. This can be mitigated by incorporating coverage oriented signals (crowding distance, hypervolume), into the ranking during search.

Practical and computational. The construction plus scaling prior admits edge cases such as singleton or degenerate Pareto fronts (see Section A.2), which we handle with well defined fallbacks that resolve automatically as more architectures are sampled. Finally, since a Pareto front update triggers recomputation and rescoring of all sampled models, the cost of scoring grows with the number of evaluated architectures.

A.2 Algorithmic Edge Case Handling

Creation of ideal fronts is not possible. If both Pareto fronts $\mathcal{F}^+$ and $\mathcal{F}^-$ have cardinality of one, we can not create ideal fronts. In this case CATOS score is defined in the Pareto corridor $\mathcal{C}(\mathcal{F}^+, \mathcal{F}^-)$ and accordingly both curves are fit to the respective Pareto point. Further, if one of the fronts has cardinality of two and the other has one, one ideal front can be constructed. Then the Pareto corridor in which CATOS is defined becomes $\mathcal{C}(\mathcal{L}^+, \mathcal{F}^-)$ or $\mathcal{C}(\mathcal{F}^+, \mathcal{L}^-)$.

Handling the singleton case for curve fitting. It can happen that a given set of models contains only a single (positively) non-dominated point, i.e. the (positive) Pareto front has size one. The same situation arises when the front contains exactly two points, since we then take their ideal point and are again left with a single point to which the positive curve g^+ must be fitted (see Algorithm 1). However, the scaling prior has two free parameters, so fitting it to a single point is underdetermined. In this case we simply fix $\alpha = 2$ and compute $\beta = (100 - \text{acc}) \cdot \text{cost}^\alpha$. This situation occurs rarely, and although the resulting score may be suboptimal, it is corrected quickly once new models are sampled and the positive ideal front grows.

Degenerate Pareto front. In rare cases, the ideal front consists of two points that are almost vertically aligned, in which case fitting the scaling prior fails, with α becoming extremely large. We handle this analogously to the singleton case above: since two points are available, we average their accuracy and cost, fix α, and compute $\beta = (100 - \text{acc}) \cdot \text{cost}^\alpha$. As before, this may lead to momentarily suboptimal scores, but the situation is an edge case that resolves automatically as the search progresses.

A.3 Extending the Scaling Prior to Error instead of Accuracy

The constructs that make up CATOS are invariant under coordinate-wise sign flips. Similarly the scaling prior can be reformulated from accuracy

$$g(cost) = 100 - \beta * cost^{-\alpha}$$

to error

$$g(cost) = \beta * cost^{-\alpha}$$

.

Moreover, the scaling prior can be replaced with a general approximation (for example linear or RBF) or a domain-specific curve form for bi-objective optimization, as the remaining constructs themselves are not specific to compute and accuracy.

A.4 Additional Details to CATOS Algorithm

Normalization is not directly mentioned in Algorithm 1 2, but we normalize all objectives to be between zero and one for the euclidean distance calculation. The normalization depends on all models seen. It is only changed, when either Pareto front is changed, then CATOS is recomputed as in Algorithm 1. If an added model does not change either Pareto front, it also doesn't require an update to the max/min-normalization, because it remains in the Pareto corridor $\mathcal{C}(\mathcal{F}^+, \mathcal{F}^-)$ and is inside the bounded region specified in Proposition 1 iii).

Finding the closest point on the scaling curves. In Algorithm 2 we compute the distance of an objective vector to the closest point on the positive and negative curve g^+ and g^-. The scaling prior formula however does not allow a closed form solution. We therefore sample each curve at 2000 points along the cost axis and treat it as a sequence of short line segments. For each query point we first locate the nearest sample using a k-d tree (Bentley, 1975), which is much faster than comparing against all samples, and then project the point onto the closer of the two adjacent segments, since the true closest point on the curve typically lies between two samples rather than exactly on one. This algorithm is vectorized in the implementation.

In the following we mention two edge cases for the curve fitting in Algorithm 1. The curve function is defined as:

$$g(\text{cost}) \;=\; 100 - \beta \cdot \text{cost}^{-\alpha}.$$

Why not fit the curve to the negative ideal front? Since we do not fit the negative curve g^- to the negative ideal front, there will be a higher amount of models which get an undiscriminating score of zero. However CATOS is not interested in distinguishing more precisely between models with bad tradeoff and the curve is then less prone to be dominated by negative ideal points given through sampled models with bad tradeoff.

B Wall-Clock Comparisons

We provide a direct wall-clock comparison against the efficient NSGA-II algorithm under the high-budget setting (500 model evaluations) on an Intel(R) Xeon(R) W-2125 CPU. Table 5 reports the mean runtime (in seconds) and standard deviation over ten runs.

Table 5: Wall-clock runtime (seconds) of CATOS-EA and NSGA-II under the high-budget setting (500 model evaluations). Values are reported as mean $\pm$ standard deviation.

Method	Runtime (s)
CATOS-EA	307.7 ± 44.1
NSGA-II	189.6 ± 7.2

Under this setting, CATOS-EA incurs roughly $1.6\times$ the runtime of NSGA-II (307.7 s versus 189.6 s). This gap should be interpreted with care: NSGA-II benefits from the highly optimized `pymoo`[2] implementation, whereas our current CATOS-EA implementation is not similarly tuned and leaves

[2] `https://pymoo.org/`

Table 6: **Multiobjective Optimization Algorithms - Filtering Analysis.** We report mean±std of percentage of ground truth Pareto points found with and without filtering NAS-Bench-201 models that are below 25th accuracy quantile and above 75th latency quantile (as in Table 2). Higher is better. **Best** and second-best per column within each budget.

Budget	Method	CIFAR-10		CIFAR-100		ImageNet16-120	
		No filt. (%) ↑	Filtered (%) ↑	No filt. (%) ↑	Filtered (%) ↑	No filt. (%) ↑	Filtered (%) ↑
Low	MOEA/D	0.63±1.18	0.00±0.00	0.45±1.36	0.18±1.17	0.00±0.00	0.00±0.00
	NSGA-III	0.91±1.80	0.89±1.98	0.41±1.30	0.31±1.51	0.25±0.91	0.29±1.13
	NSGA-II	0.80±1.52	0.58±1.95	0.59±1.53	0.38±1.68	0.29±0.97	0.33±1.21
	R-NSGA-II	1.11±1.93	1.05±2.58	0.27±1.08	0.23±1.31	0.25±0.91	0.33±1.21
	RVEA	0.40±1.14	0.47±1.84	0.41±1.30	0.31±1.51	0.11±0.61	0.14±0.81
	Ours	**1.69**±2.32	**2.26**±3.43	**2.14**±3.55	**2.00**±3.86	**1.50**±2.68	**1.81**±3.55
Mid	MOEA/D	4.49±2.92	0.00±0.00	3.77±1.71	0.00±0.00	0.00±0.00	0.00±0.00
	NSGA-III	9.06±5.77	6.79±5.87	2.09±2.83	2.46±4.05	1.71±2.45	1.86±2.77
	NSGA-II	12.49±6.91	7.68±5.34	3.09±3.14	2.77±4.15	1.79±2.45	1.81±2.76
	R-NSGA-II	10.11±5.81	8.37±6.62	2.05±3.04	1.85±3.79	2.07±2.43	2.67±3.18
	RVEA	3.57±3.41	2.63±3.68	1.64±2.84	1.15±2.95	0.93±1.72	1.00±2.05
	Ours	**14.26**±6.09	**16.89**±7.60	**10.64**±8.61	**12.00**±10.95	**9.39**±6.10	**11.05**±7.06
High	MOEA/D	5.60±3.10	1.18±2.19	4.23±1.33	0.31±1.51	0.00±0.00	0.00±0.00
	NSGA-III	16.49±7.64	14.00±7.67	4.50±4.19	4.77±6.04	2.86±2.86	3.24±3.62
	NSGA-II	19.80±8.23	13.53±8.75	5.09±3.60	4.62±5.44	2.68±2.87	2.86±3.50
	R-NSGA-II	17.54±7.27	14.95±7.73	3.14±3.84	3.00±4.47	3.11±2.93	3.86±3.85
	RVEA	5.49±3.78	3.00±3.43	2.27±3.05	1.31±3.09	1.39±2.20	1.71±2.82
	Ours	**20.97**±7.76	**25.21**±9.39	**18.18**±9.88	**20.38**±13.07	**16.79**±6.97	**19.67**±8.69

substantial room for optimization. We therefore regard the reported runtimes as an upper bound on CATOS-EA's overhead rather than a fundamental algorithmic limitation.

C Further Evaluation Analysis

C.1 Additional Analysis of Comparison with Multiobjective Optimization Algorithms

Effect of Filtering on Found Pareto Points In Table 6 we measure the percentage of ground truth Pareto points that were found during search. We compare how filtering out models as we did for Table 2 affects evaluation for MOO algorithms in this metric. CATOS-EA always performs best with the filtered and unfiltered Pareto front. More interestingly, CATOS-EA almost never finds a lower percentage of ground truth Pareto points after filtering, but rather the percentage stays comparably stable or improves, in contrast to the compared MOO algorithms.

C.2 Warm-Start Analysis

In order to analyze how CATOS-EA handles bad population initializations, we compare how runs with low HV-regret in the initial population improve compared to population initialization that have a higher HV-regret. Results are in Table 7 and Table 8. For this, we grouped 100 CATOS-EA runs by hypervolume regret after initial population sampling, then tracked how the mean HV-regret of the same group of runs developed over the search. Configurations are the same as in Table 2, using NATS-Bench topology search space, but only on CIFAR-10. From Table 7 and Table 8 we can see, that the runs with the best initial HV-regret, also have the best final HV-regret. However, the runs that start with the worst HV-regret, also have by far the greatest gains over the search. This is true for low and high setting.

C.3 Budget Plots for Multiobjective Comparison

For the experiments reported in Table 3 and Table 2, we additionally plot in Figure 3 the evolution of mean hypervolume regret (averaged over the 100 runs) over the course of the search under reduced

Table 7: Warm-start analysis in low setting (population = 10). Measured mean HV-regret to ground truth, lower is better. NATS-Bench topology, CIFAR-10.

Step	best (0–20%)	20–40%	40–60%	60–80%	worst (80–100%)
10	0.1235	0.1766	0.2310	0.2558	0.3231
20	0.1116	0.1520	0.1781	0.1907	0.2578
30	0.1012	0.1359	0.1302	0.1647	0.2012
40	0.0920	0.1161	0.1149	0.1398	0.1698
50	0.0869	0.1035	0.0983	0.1313	0.1503
60	0.0821	0.0989	0.0900	0.1110	0.1315

Table 8: Warm-start analysis in high setting (population = 50). Measured mean HV-regret to ground truth, lower is better. NATS-Bench topology, CIFAR-10.

Step	best (0–20%)	20–40%	40–60%	60–80%	worst (80–100%)
50	0.0693	0.0968	0.1116	0.1292	0.1889
100	0.0565	0.0743	0.0920	0.0993	0.1058
200	0.0449	0.0578	0.0561	0.0596	0.0641
500	0.0314	0.0385	0.0364	0.0365	0.0494

Pareto-front evaluation (see Section 4.3). Most runs perform fewer unique architecture evaluations than their nominal budget. This results from our duplicate-handling procedure: when mutation produces a previously sampled architecture, it is resampled up to five times, and if a duplicate is still obtained, it is accepted. Because a single run terminating early would skew the aggregate curve, we plot results only up to the point at which the first run of any method exhausts its pool of unique architectures.

Results. After the initial uniform population sampling (low=10, mid=30, high=50), Figure 3 shows that CATOS-EA achieves lower hypervolume regret early in the search, while maintaining or widening its margin over the next-best competing method.

C.4 Extended Comparison with Pareto Baselines and Bayesian optimization Methods

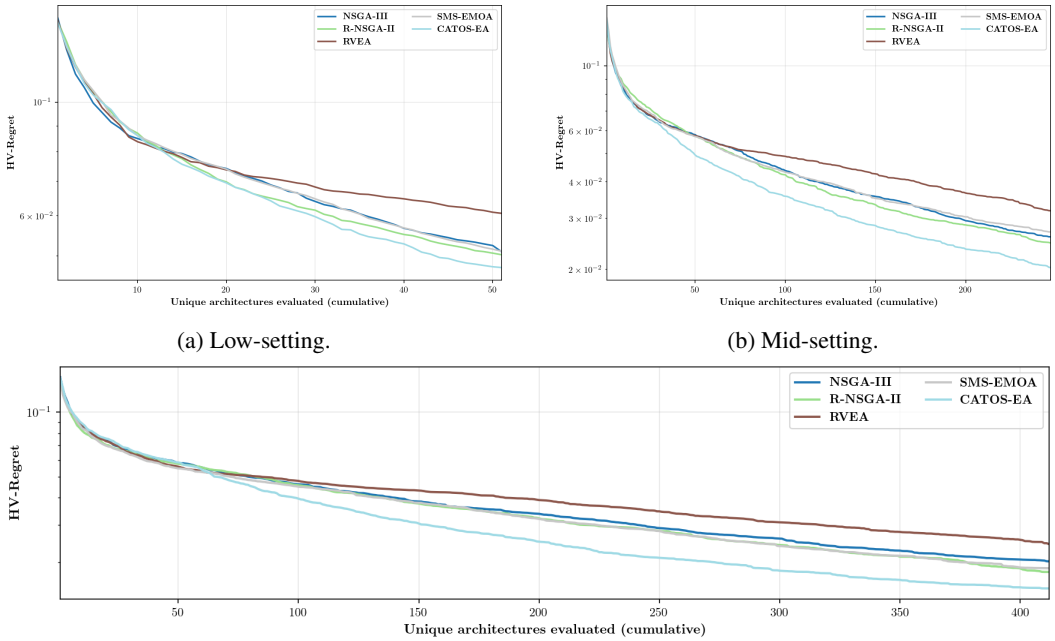
We further provide in Tables 9 and 10 an extended comparison with Pareto-based baselines (hypervolume, crowding distance), as well as the Bayesian multi-objective optimization approaches ParEGO (Knowles, 2006) and qEHVI (Daulton et al., 2020). CATOS-EA shows strong results throughout.

D Explanation of CATOS visualization

To clarify CATOS and its constructs, we walk through Figure 4 (a repetition of Figure 1) in detail. Both subfigures visualize CATOS-scoring together with its underlying constructs. The green points form the (positive) Pareto front and are connected by the (positive) Pareto boundary of the positive ideal front; analogously, the red points form the negative Pareto front, connected by the negative Pareto boundary of the negative ideal front.

Pareto boundary. Each boundary delineates the corresponding *ideal* front, not the Pareto front itself. Consequently, the positive Pareto points (excluding the endpoints) sit in the corner of the boundary rather than on its edge, indicating the directions of improvement (lower latency, higher accuracy). The negative Pareto points behave analogously: aside from the endpoints, they lie on the edge rather than in the corner, indicating the directions of worsening (higher latency, lower accuracy).

The Pareto boundary defined in Section E.2 extends with infinite tails; these are omitted from the figure, which shows only the portion contained in the Pareto corridor $\mathcal{C}(\mathcal{L}^+, \mathcal{L}^-)$.

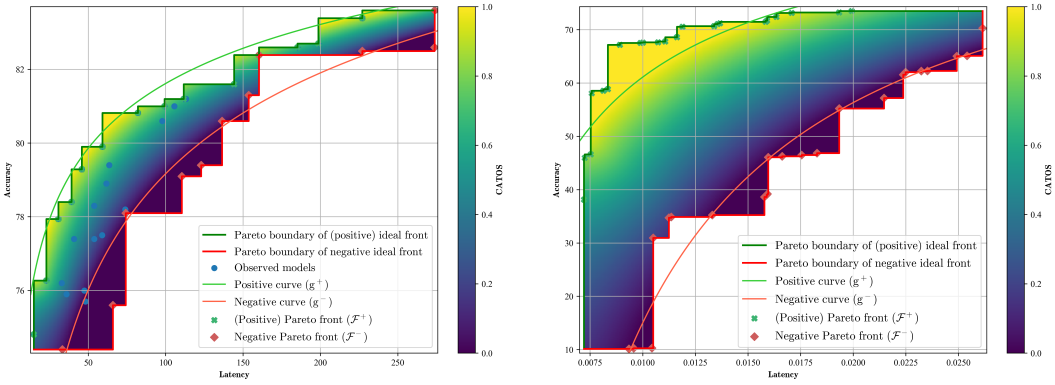


(a) Low-setting.

(b) Mid-setting.

(c) High-setting.

Figure 3: NATS-Bench topology search space, CIFAR-100 dataset. HV-regret to filtered ground truth Pareto front. Lower is better. Same runs as in Table 2, following the exact same evaluation protocol and the reduced Pareto front evaluation. Three different settings (see Section 4.3): low, mid, high. The y-axis is in logarithmic scale.



(a) ImageNet, 38 trained models.

(b) NATS-Bench topology on CIFAR-100.

Figure 4: **CATOS on two sets of models.** The shaded area represents the Pareto corridor $\mathcal{C}(\mathcal{L}^+, \mathcal{L}^-)$.

E Mathematical Proofs

Before proceeding with the proofs, we briefly introduce *strict dominance*, written $p \prec\!\!\prec y$, which requires $p_i < y_i$ in *every* coordinate; the negative analogue $\prec\!\!\prec^-$ flips all inequalities. Following the convention of Section 3.1, we will write $\prec\!\!\prec^+$ in place of $\prec\!\!\prec$ when symmetry with $\prec\!\!\prec^-$ aids clarity. Strict dominance is the natural strengthening of $\prec$ and is used in several of the proofs that follow.

E.1 Pareto Corridor

We first repeat the definition and proposition and then proceed with the proof of the proposition.

Table 9: **Pareto-compliant evaluation against classical MOO algorithms, Bayesian optimization baselines, and selection ablations.** We report Hypervolume Regret (HV-Reg., gap to the restricted ground-truth Pareto front) and the Epsilon-Indicator (ϵ-Ind.) on NATS-Bench topology. Lower is better for both. **Best** and second-best per (budget, metric, dataset).

Budget	Method	*Architectures filtered by* lat. $\leq Q_{75}$ *and* acc. $\geq Q_{25}$ *(∼50% of models)*					
		CIFAR-10		**CIFAR-100**		**ImageNet16-120**	
		HV-Reg. ↓	ϵ-Ind. ↓	HV-Reg. ↓	ϵ-Ind. ↓	HV-Reg. ↓	ϵ-Ind. ↓
Low	MOEA/D	0.495±.177	0.734±.213	0.614±.154	1.357±.501	0.644±.142	0.879±.226
	NSGA-III	0.163±.050	0.249±.075	0.294±.068	0.482±.122	0.301±.088	0.367±.093
	NSGA-II	0.169±.056	0.281±.082	0.299±.073	0.494±.129	0.297±.075	0.364±.076
	R-NSGA-II	0.148±.055	0.238±.094	0.294±.080	0.496±.149	0.305±.085	0.372±.094
	RVEA	0.309±.139	0.420±.164	0.426±.122	0.601±.159	0.433±.126	0.540±.170
	ParEGO	0.190±.060	0.292±.089	0.328±.060	0.510±.096	0.339±.066	0.401±.080
	qEHVI	0.196±.077	0.315±.110	0.321±.072	0.495±.104	0.332±.079	0.402±.096
	Crowding-only	0.204±.083	0.310±.103	0.335±.067	0.558±.145	0.364±.087	0.421±.090
	HV-only	0.137±.038	0.225±.071	0.274±.067	0.462±.122	0.286±.072	**0.338**±.077
	Ours	**0.120**±.059	**0.190**±.080	**0.265**±.077	**0.423**±.129	**0.263**±.088	0.341±.107
Mid	MOEA/D	0.457±.130	0.675±.183	0.599±.122	1.312±.585	0.630±.151	0.888±.241
	NSGA-III	0.055±.018	0.111±.031	0.145±.050	0.254±.053	0.118±.035	0.197±.055
	NSGA-II	0.054±.018	0.113±.036	0.145±.049	0.267±.060	0.134±.037	0.210±.047
	R-NSGA-II	0.055±.020	0.121±.042	0.139±.045	0.258±.058	0.116±.044	0.184±.049
	RVEA	0.100±.039	0.174±.057	0.203±.060	0.344±.086	0.179±.066	0.255±.084
	ParEGO	0.090±.029	0.157±.038	0.202±.052	0.331±.059	0.192±.054	0.270±.064
	qEHVI	0.106±.041	0.179±.055	0.207±.054	0.322±.071	0.232±.068	0.325±.083
	Crowding-only	0.102±.029	0.181±.047	0.223±.047	0.389±.081	0.217±.056	0.290±.061
	HV-only	0.054±.017	0.111±.038	0.148±.049	0.265±.050	0.127±.041	0.202±.054
	Ours	**0.042**±.020	**0.092**±.041	**0.110**±.051	**0.240**±.073	**0.097**±.041	**0.160**±.046
High	MOEA/D	0.545±.201	0.784±.229	0.587±.109	1.315±.465	0.591±.148	0.851±.228
	NSGA-III	0.040±.016	0.085±.026	0.114±.046	0.221±.052	0.089±.035	0.167±.040
	NSGA-II	0.042±.016	0.094±.028	0.117±.035	0.234±.044	0.092±.033	0.179±.050
	R-NSGA-II	0.037±.014	0.086±.031	0.101±.039	0.214±.049	0.084±.032	0.162±.042
	RVEA	0.084±.028	0.150±.045	0.163±.049	0.276±.062	0.133±.048	0.215±.067
	ParEGO	0.070±.021	0.136±.031	0.169±.042	0.284±.046	0.153±.041	0.242±.051
	qEHVI	0.074±.028	0.137±.044	0.173±.058	0.282±.072	0.196±.072	0.286±.092
	Crowding-only	0.079±.024	0.154±.041	0.165±.046	0.314±.069	0.163±.048	0.238±.059
	HV-only	0.046±.017	0.095±.027	0.116±.049	0.223±.041	0.103±.033	0.182±.041
	Ours	**0.027**±.014	**0.068**±.039	**0.083**±.045	**0.213**±.056	**0.063**±.032	**0.127**±.043

Definition 5 (Pareto Corridor)**.** Let $\mathcal{S} \subset \mathbb{R}^2$ be a finite set of objective vectors, and let $\mathcal{U}, \mathcal{V} \subseteq \mathcal{S}$. The *Pareto corridor* induced by $\mathcal{U}$ and $\mathcal{V}$ is

$$\mathcal{C}(\mathcal{U}, \mathcal{V}) \;:=\; \big\{z \in \mathbb{R}^2 \mid \exists\, u \in \mathcal{U} : u \prec^+ z\big\} \;\cap\; \big\{z \in \mathbb{R}^2 \mid \exists\, v \in \mathcal{V} : v \prec^- z\big\}.$$

Proposition 1 (Properties of the Pareto Corridor)**.** Let $\mathcal{S} \subset \mathbb{R}^2$ be finite with (positive) Pareto front $\mathcal{F}^+ = \{y^1, \ldots, y^k\}$ and negative Pareto front $\mathcal{F}^- = \{w^1, \ldots, w^m\}$, indexed as in Definition 4, with $|\mathcal{F}^+|, |\mathcal{F}^-| \geq 2$. Let $\mathcal{L}^+, \mathcal{L}^-$ be the corresponding ideal fronts.

(i) $\mathcal{C}(\mathcal{F}^+, \mathcal{F}^-)$ may be empty, but $\mathcal{C}(\mathcal{F}^+, \mathcal{F}^-) \subseteq \mathcal{C}(\mathcal{L}^+, \mathcal{L}^-)$.

(ii) $\mathcal{C}(\mathcal{L}^+, \mathcal{L}^-) \supseteq (y_1^1, w_1^2) \times (y_2^2, w_2^1)$, an open rectangle with strictly positive side lengths.

(iii) $\mathcal{C}(\mathcal{L}^+, \mathcal{L}^-) \subseteq [y_1^1, w_1^m] \times [y_2^k, w_2^1]$.

Proof of Proposition 1. We first establish two extremal observations used throughout. For an arbitrary point $s \in \mathcal{S} \subset \mathbb{R}^2$, write $s = (s_1, s_2)$, so that $\min_{s \in \mathcal{S}} s_1$ denotes the smallest first coordinate occurring anywhere in $\mathcal{S}$, and analogously for the other extrema.

(E1) $y_1^1 = \min_{s \in \mathcal{S}} s_1$ *and* $y_2^k = \min_{s \in \mathcal{S}} s_2$. Let $s^\star \in \mathcal{S}$ minimize s_1, breaking ties by smallest s_2. Suppose $s' \prec^+ s^\star$ for some $s' \in \mathcal{S}$. Then $s'_1 \leq s_1^\star$, but $s_1^\star$ is already the minimum, so $s'_1 = s_1^\star$. The tiebreaker forces $s'_2 \geq s_2^\star$, which combined with $s'_2 \leq s_2^\star$ yields $s' = s^\star$, contradicting strictness of $\prec^+$. Hence $s^\star \in \mathcal{F}^+$, and the indexing convention identifies $s^\star = y^1$. Symmetrically (swapping coordinates), $y_2^k = \min_{s \in \mathcal{S}} s_2$.

Table 10: **Multiobjective Optimization – Constrained Evaluation.** Regret to the optimum on NATS-Bench topology search space (Dong et al., 2021) for maximizing accuracy s.t. latency $\leq Q_{50}$ and minimizing latency s.t. accuracy $\geq Q_{75}$. Lower is better; **best** and second-best per column.

Budget	Method	CIFAR-10		CIFAR-100		ImageNet16-120	
		Δacc (pp)$\downarrow$ s.t. lat $\leq Q_{50}$	Δlat (ms)$\downarrow$ s.t. acc $\geq Q_{75}$	Δacc (pp)$\downarrow$ s.t. lat $\leq Q_{50}$	Δlat (ms)$\downarrow$ s.t. acc $\geq Q_{75}$	Δacc (pp)$\downarrow$ s.t. lat $\leq Q_{50}$	Δlat (ms)$\downarrow$ s.t. acc $\geq Q_{75}$
Low	MOEA/D	1.66±1.24	3.9±1.3	2.83±2.06	3.1±1.2	4.04±2.64	3.6±1.2
	NSGA-III	1.11±0.58	3.0±1.1	2.17±1.16	2.2±1.2	3.52±1.57	2.8±1.3
	NSGA-II	1.29±0.58	3.3±1.3	2.20±1.19	2.4±1.2	3.36±1.58	2.9±1.2
	R-NSGA-II	1.07±0.59	2.8±1.0	2.05±1.07	2.0±1.0	3.01±1.44	2.5±1.0
	RVEA	1.06±0.67	3.1±1.1	2.23±1.40	2.5±1.1	3.36±1.78	3.2±1.1
	ParEGO	1.34±0.56	3.4±1.1	2.45±0.95	2.6±0.9	3.50±1.68	3.0±1.2
	qEHVI	1.32±0.65	3.4±0.9	2.18±1.17	2.4±1.0	3.44±1.72	3.1±1.2
	Crowding-only	1.68±0.67	4.1±1.5	3.12±1.25	3.1±1.2	4.36±1.84	3.7±1.4
	HV-only	1.01±0.59	2.5±1.2	1.93±1.05	2.1±1.0	3.43±1.47	2.5±1.2
	Ours	**0.71**±0.57	**2.3**±1.0	**1.33**±1.00	**1.9**±1.0	**2.65**±1.31	**2.2**±1.0
Mid	MOEA/D	0.90±0.77	3.2±0.9	**0.35**±0.78	2.1±0.9	2.52±1.59	2.8±0.9
	NSGA-III	0.22±0.23	1.5±0.9	0.53±0.47	0.7±0.7	1.40±0.70	1.2±0.8
	NSGA-II	0.26±0.29	1.8±0.8	0.62±0.68	0.9±0.7	1.55±0.77	1.2±0.8
	R-NSGA-II	0.33±0.36	1.3±0.9	0.60±0.59	**0.5**±0.7	**1.10**±0.73	1.1±0.7
	RVEA	**0.19**±0.18	2.0±0.7	0.57±0.52	0.9±0.7	1.33±0.84	1.5±0.8
	ParEGO	0.50±0.37	2.0±0.9	0.92±0.64	1.3±0.7	1.89±0.79	1.8±0.7
	qEHVI	0.50±0.45	2.3±0.7	0.63±0.57	1.3±0.6	1.90±0.89	1.9±0.8
	Crowding-only	0.82±0.52	2.2±1.0	1.52±0.85	1.7±0.8	2.27±1.09	2.2±0.9
	HV-only	0.24±0.24	1.6±0.9	0.59±0.48	1.0±0.6	1.43±0.60	1.2±0.7
	Ours	0.22±0.35	**0.9**±0.9	0.39±0.55	0.6±0.4	1.17±0.63	**0.8**±0.6
High	MOEA/D	0.50±0.54	2.7±0.7	**0.11**±0.44	1.8±0.7	1.61±1.14	2.3±0.9
	NSGA-III	**0.11**±0.08	1.1±0.8	0.37±0.44	0.3±0.5	1.12±0.68	0.8±0.6
	NSGA-II	0.14±0.16	1.3±0.9	0.30±0.43	0.5±0.6	1.09±0.72	0.9±0.7
	R-NSGA-II	0.19±0.17	0.8±0.7	0.30±0.35	**0.0**±0.5	**0.80**±0.63	0.8±0.6
	RVEA	0.16±0.15	1.5±0.9	0.28±0.45	0.6±0.6	1.15±0.74	1.1±0.6
	ParEGO	0.36±0.28	1.7±0.9	0.65±0.51	0.9±0.6	1.42±0.51	1.6±0.7
	qEHVI	0.20±0.30	1.8±0.8	0.43±0.45	1.0±0.6	1.52±0.85	1.8±0.7
	Crowding-only	0.55±0.43	1.7±1.0	1.00±0.44	1.3±0.6	1.74±0.89	1.6±0.8
	HV-only	0.12±0.12	1.2±0.8	0.36±0.30	0.7±0.6	1.24±0.55	0.9±0.7
	Ours	0.15±0.25	**0.5**±0.6	0.22±0.22	0.4±0.3	1.04±0.50	**0.5**±0.5

(E2) $w_2^1 = \max_{s\in\mathcal{S}} s_2$ *and* $w_1^m = \max_{s\in\mathcal{S}} s_1$. The same argument, with $\prec^-$ in place of $\prec^+$ (every inequality flipped) and ties broken by largest s_1, identifies the maximizer of s_2 as w^1; symmetrically for s_1.

(i) $\mathcal{C}(\mathcal{F}^+, \mathcal{F}^-)$ may be empty, but $\mathcal{C}(\mathcal{F}^+, \mathcal{F}^-) \subseteq \mathcal{C}(\mathcal{L}^+, \mathcal{L}^-)$.

Possibly empty. We exhibit an instance with $|\mathcal{F}^+| = |\mathcal{F}^-| = 2$ for which the corridor is empty. Let $\mathcal{S} = \{y^1, y^2\}$ with $y^1 = (1, 3)$ and $y^2 = (2, 1)$. Since neither point dominates the other under $\prec^+$, $\mathcal{F}^+ = \mathcal{F}^- = \{y^1, y^2\}$. A point z lies in $\mathcal{C}(\mathcal{F}^+, \mathcal{F}^-)$ iff at least one of

$$(A)\ z_1 \geq 1,\ z_2 \geq 3 \quad \text{or} \quad (B)\ z_1 \geq 2,\ z_2 \geq 1$$

holds, and at least one of

$$(C)\ z_1 \leq 1,\ z_2 \leq 3 \quad \text{or} \quad (D)\ z_1 \leq 2,\ z_2 \leq 1$$

holds, each with at least one strict inequality (z cannot lie on either Pareto front). Checking all combinations: (A)∧(C) forces $z = y^1$, (B)∧(D) forces $z = y^2$ (both with all inequalities tight, contradiction); (A)∧(D) forces $z_2 \geq 3$ and $z_2 \leq 1$; (B)∧(C) forces $z_1 \geq 2$ and $z_1 \leq 1$. All four yield contradictions, so $\mathcal{C}(\mathcal{F}^+, \mathcal{F}^-) = \emptyset$.

Subset relation. Let $z \in \mathcal{C}(\mathcal{F}^+, \mathcal{F}^-)$, so there exist $y^i \in \mathcal{F}^+$ and $w^j \in \mathcal{F}^-$ with $y^i \prec^+ z$ and $w^j \prec^- z$. We show $z \in \mathcal{C}(\mathcal{L}^+, \mathcal{L}^-)$ by exhibiting $\ell_+^a \prec^+ z$ and $\ell_-^b \prec^- z$ for suitable indices.

For the positive side: if $i < k$, let $a = i$, so $\ell_+^a = (y_1^i, y_2^{i+1})$. From $y^i \prec^+ z$ we have $y_1^i \leq z_1$ and $y_2^i \leq z_2$; combined with $y_2^{i+1} < y_2^i$ (by the $\mathcal{F}^+$ ordering and $|\mathcal{F}^+| \geq 2$), this gives $\ell_+^a \prec^+ z$. If $i = k$, let $a = k - 1$, so $\ell_+^a = (y_1^{k-1}, y_2^k)$; then $y_1^{k-1} < y_1^k \leq z_1$ and $y_2^k \leq z_2$, again giving $\ell_+^a \prec^+ z$.

The negative side is symmetric: if $j > 1$, take $b = j - 1$, so $\ell^b_- = (w^j_1, w^{j-1}_2)$; if $j = 1$, take $b = 1$.

(ii) $\mathcal{C}(\mathcal{L}^+, \mathcal{L}^-) \supseteq (y^1_1, w^2_1) \times (y^2_2, w^1_2)$**.**

Recall $\ell^1_+ = (y^1_1, y^2_2)$ and $\ell^1_- = (w^2_1, w^1_2)$.

First coordinate. By (E1), $y^1_1 \leq w^1_1$. By the $\mathcal{F}^-$ ordering and $|\mathcal{F}^-| \geq 2$, $w^1_1 < w^2_1$. Chaining, $y^1_1 \leq w^1_1 < w^2_1$.

Second coordinate. By the $\mathcal{F}^+$ ordering and $|\mathcal{F}^+| \geq 2$, $y^2_2 < y^1_2$. By (E2), $y^1_2 \leq w^1_2$. Chaining, $y^2_2 < y^1_2 \leq w^1_2$.

Both strict inequalities give that the open rectangle $(y^1_1, w^2_1) \times (y^2_2, w^1_2)$ has strictly positive side lengths. For any z in this rectangle, $\ell^1_+ \prec^+ z$ (strict in both coordinates) and $z \prec^+ \ell^1_-$, equivalently $\ell^1_- \prec^- z$. Hence $z \in \mathcal{C}(\mathcal{L}^+, \mathcal{L}^-)$.

(iii) $\mathcal{C}(\mathcal{L}^+, \mathcal{L}^-) \subseteq [y^1_1, w^m_1] \times [y^k_2, w^1_2]$**.**

Let $z \in \mathcal{C}(\mathcal{L}^+, \mathcal{L}^-)$. Then $\ell^i_+ \prec^+ z$ for some $i \in \{1, \ldots, k-1\}$, giving $z_1 \geq y^i_1 \geq y^1_1$ and $z_2 \geq y^{i+1}_2 \geq y^k_2$. Similarly, $\ell^j_- \prec^- z$ for some j, equivalently $z \prec^+ \ell^j_-$, giving $z_1 \leq w^{j+1}_1 \leq w^m_1$ and $z_2 \leq w^j_2 \leq w^1_2$. Hence $z \in [y^1_1, w^m_1] \times [y^k_2, w^1_2]$. □

E.1.1 Singleton Edge Case

In the rare case that the negative or (positive) Pareto front just contains one point, the Pareto corridor can be still non-empty:

Proposition 2 (Pareto Corridor in the singleton case)**.** If $|\mathcal{F}^+| = |\mathcal{F}^-| = 1$, with $\mathcal{F}^+ = \{y^1\}$ and $\mathcal{F}^- = \{w^1\}$ satisfying $y^1 \prec\!\!\prec^+ w^1$, then

$$(y^1_1, w^1_1) \times (y^1_2, w^1_2) \;\subseteq\; \mathcal{C}(\mathcal{F}^+, \mathcal{F}^-) \;\subseteq\; [y^1_1, w^1_1] \times [y^1_2, w^1_2],$$

with strictly positive side lengths. In particular, $\mathcal{C}(\mathcal{F}^+, \mathcal{F}^-)$ is non-empty, bounded, and contains a two-dimensional region.

Proof of Proposition 2. **Lower inclusion.** By the hypothesis $y^1 \prec\!\!\prec^+ w^1$, both $y^1_1 < w^1_1$ and $y^1_2 < w^1_2$ hold strictly, so the open rectangle $(y^1_1, w^1_1) \times (y^1_2, w^1_2)$ has strictly positive side lengths. Any z in this rectangle satisfies $y^1_1 < z_1 < w^1_1$ and $y^1_2 < z_2 < w^1_2$, giving $y^1 \prec^+ z$ (strict in both coordinates) and $z \prec^+ w^1$, i.e. $w^1 \prec^- z$. Hence $z \in \mathcal{C}(\mathcal{F}^+, \mathcal{F}^-)$.

Upper inclusion. Let $z \in \mathcal{C}(\mathcal{F}^+, \mathcal{F}^-)$. Since $\mathcal{F}^+ = \{y^1\}$ and $\mathcal{F}^- = \{w^1\}$, the corridor condition reduces to $y^1 \prec^+ z$ and $w^1 \prec^- z$. The first gives $z_1 \geq y^1_1$ and $z_2 \geq y^1_2$; the second, equivalent to $z \prec^+ w^1$, gives $z_1 \leq w^1_1$ and $z_2 \leq w^1_2$. Hence $z \in [y^1_1, w^1_1] \times [y^1_2, w^1_2]$. □

A similar argument handles the asymmetric setting in which one front has cardinality two and the other has cardinality one. Constructing the one available ideal front reduces this case to a singleton-vs-singleton configuration, and the strict-dominance hypothesis required by Proposition 2 holds automatically:

Proposition 3 (Pareto Corridor in the Asymmetric Case)**.** Suppose $|\mathcal{F}^+| = 2$ and $|\mathcal{F}^-| = 1$, with $\mathcal{F}^+ = \{y^1, y^2\}$ (indexed as in Definition 4) and $\mathcal{F}^- = \{w^1\}$. Then $\ell^1_+ \prec\!\!\prec^+ w^1$, and

$$(y^1_1,\, w^1_1) \times (y^2_2,\, w^1_2) \;\subseteq\; \mathcal{C}(\mathcal{L}^+, \mathcal{F}^-) \;\subseteq\; [y^1_1,\, w^1_1] \times [y^2_2,\, w^1_2],$$

with strictly positive side lengths; in particular $\mathcal{C}(\mathcal{L}^+, \mathcal{F}^-)$ is non-empty, bounded, and contains a two-dimensional region. The symmetric statement holds for $|\mathcal{F}^+| = 1$ and $|\mathcal{F}^-| = 2$, with $\mathcal{C}(\mathcal{F}^+, \mathcal{L}^-)$ in place of $\mathcal{C}(\mathcal{L}^+, \mathcal{F}^-)$.

Proof. Since $|\mathcal{F}^-| = 1$ and $\mathcal{S}$ is finite, every $s \in \mathcal{S} \setminus \{w^1\}$ is negatively dominated by some element of $\mathcal{F}^-$, and w^1 is the only candidate. In particular w^1 negatively dominates both y^1 and y^2, giving $w^1_t \geq y^j_t$ for all $t, j \in \{1, 2\}$. Combined with the indexing $y^1_1 < y^2_1$ and $y^2_2 < y^1_2$, this chains to

$$\ell^1_{+,1} \;=\; y^1_1 \;<\; y^2_1 \;\leq\; w^1_1, \qquad \ell^1_{+,2} \;=\; y^2_2 \;<\; y^1_2 \;\leq\; w^1_2,$$

so $\ell_+^1 \prec\!\!\prec^+ w^1$. Applying Proposition 2 to the singleton sets $\{\ell_+^1\}$ and $\{w^1\}$ in place of $\mathcal{F}^+, \mathcal{F}^-$ yields the displayed inclusions. The case $|\mathcal{F}^+| = 1, |\mathcal{F}^-| = 2$ follows by applying the order-reversing involution $z \mapsto -z$. □

E.2 Pareto Boundary

When plotting a Pareto front, a step-wise function is used (Mescoli et al., 2026), which we call *Pareto boundary* and now formally introduce. Particularly, we will define a (positive) Pareto boundary and a negative Pareto boundary. We refer to it as a boundary since it induces a partition of the objective space into dominated and non-dominated regions.

Definition 6 (Positive and negative Pareto boundary)**.** Given a positive Pareto front $\mathcal{F}$ the *positive Pareto boundary* is

$$\mathcal{B}(\mathcal{F}) \;\coloneqq\; \big\{\, y \in \mathbb{R}^d \setminus \mathcal{F} \;\mid\; \big(\exists p \in \mathcal{F} : p \prec y\big) \;\wedge\; \big(\forall p' \in \mathcal{F} : \neg(p' \prec\!\!\prec y)\big) \,\big\}.$$

The *negative Pareto boundary* is defined analogously with respect to the negative Pareto front $\mathcal{F}^-$:

$$\mathcal{B}^-(\mathcal{F}^-) \;\coloneqq\; \big\{\, y \in \mathbb{R}^d \setminus \mathcal{F}^- \;\mid\; \big(\exists p \in \mathcal{F}^- : p \prec^- y\big) \;\wedge\; \big(\forall p' \in \mathcal{F}^- : \neg(p' \prec\!\!\prec^- y)\big) \,\big\}.$$

Colloquially, the positive Pareto boundary are the points in objective space, which are positively Pareto dominated but not strictly dominated (analogously for the negative boundary). Next we will prove that this actually represents the staircase by giving an explicit geometric description of $\mathcal{B}(\mathcal{F})$ and $\mathcal{B}^-(\mathcal{F}^-)$ in $\mathbb{R}^2$. We restrict to $d = 2$, as $d > 2$ only adds notational complexity without giving more structural insight for the following characterization.

Proposition 4 (Characterization of Positive and Negative Pareto Boundaries)**.** Let $\mathcal{F} = \{p^1, \dots, p^k\}$ be a finite positive Pareto front with $p^i = (p_1^i, p_2^i)$ indexed such that $p_1^1 < p_1^2 < \cdots < p_1^k$, which by Pareto optimality implies $p_2^1 > p_2^2 > \cdots > p_2^k$. Analogously, let $\mathcal{F}^- = \{q^1, \dots, q^m\}$ be a finite negative Pareto front with $q^i = (q_1^i, q_2^i)$ indexed such that $q_1^1 < \cdots < q_1^m$, which by negative Pareto optimality implies $q_2^1 > \cdots > q_2^m$. Then the positive Pareto boundary $\mathcal{B}(\mathcal{F})$ is given by

$$\begin{aligned}\mathcal{B}(\mathcal{F}) \;\coloneqq\; &\underbrace{\big(\{p_1^1\} \times (p_2^1, +\infty)\big)}_{\mathcal{B}^0} \cup \bigcup_{i=1}^{k-1} \underbrace{\big((p_1^i,\, p_1^{i+1}) \times \{p_2^i\}\big)}_{\mathcal{B}_i^1} \cup \bigcup_{i=1}^{k-1} \underbrace{\big(\{p_1^{i+1}\} \times (p_2^{i+1},\, p_2^i]\big)}_{\mathcal{B}_i^2} \\ &\cup \underbrace{\big((p_1^k,\, +\infty) \times \{p_2^k\}\big)}_{\mathcal{B}^3},\end{aligned}$$

and the negative Pareto boundary $\mathcal{B}^-(\mathcal{F}^-)$ is given by

$$\begin{aligned}\mathcal{B}^-(\mathcal{F}^-) \;\coloneqq\; &\underbrace{\big((-\infty,\, q_1^1) \times \{q_2^1\}\big)}_{\mathcal{B}^{-,0}} \cup \bigcup_{i=1}^{m-1} \underbrace{\big(\{q_1^i\} \times [q_2^{i+1},\, q_2^i)\big)}_{\mathcal{B}_i^{-,1}} \cup \bigcup_{i=1}^{m-1} \underbrace{\big((q_1^i,\, q_1^{i+1}) \times \{q_2^{i+1}\}\big)}_{\mathcal{B}_i^{-,2}} \\ &\cup \underbrace{\big(\{q_1^m\} \times (-\infty,\, q_2^m)\big)}_{\mathcal{B}^{-,3}}.\end{aligned}$$

By convention, the corner (p_1^{i+1}, p_2^i) belongs to $\mathcal{B}_i^2$ and the corner (q_1^i, q_2^{i+1}) belongs to $\mathcal{B}_i^{-,1}$; with these choices each union is a partition.

Lemma 1 (Pointwise criterion)**.** Given Pareto dominance and strict dominance are defined as:

$$p \prec y \iff p \le y \text{ componentwise and } p \neq y, \qquad p \prec\!\!\prec y \iff p_i < y_i \text{ for both } i,$$

and similarly are $\prec^-, \prec\!\!\prec^-$ given by reversing every inequality. For $p, y \in \mathbb{R}^2$ with $p \neq y$ then holds:

(i) $p \prec y \wedge \neg(p \prec\!\!\prec y)$ holds if and only if (A) or (B) holds, where

$$\text{(A)} \quad p_1 = y_1 \text{ and } p_2 < y_2, \qquad \text{(B)} \quad p_2 = y_2 \text{ and } p_1 < y_1.$$

(ii) $p \prec^- y \wedge \neg(p \prec\!\!\prec^- y)$ holds if and only if (A^-) or (B^-) holds, where

$$(\text{A}^-) \quad p_1 = y_1 \text{ and } p_2 > y_2, \qquad (\text{B}^-) \quad p_2 = y_2 \text{ and } p_1 > y_1.$$

In each part, (A) and (B) (resp. (A^-) and (B^-)) are mutually exclusive under the hypothesis $p \neq y$.

Proof. Part (i). Expanding the two conjuncts:

$$p \prec y \iff p_1 \le y_1,\ p_2 \le y_2,\ p \neq y; \qquad \neg(p \prec\!\!\prec y) \iff p_1 \ge y_1 \text{ or } p_2 \ge y_2.$$

Their conjunction forces $p_i = y_i$ for at least one i (combine $p_i \le y_i$ with $p_i \ge y_i$ on whichever coordinate witnesses the second clause). Both coordinates equal would give $p = y$, excluded by hypothesis, so exactly one coordinate i satisfies $p_i = y_i$. The other coordinate satisfies $p_j \le y_j$ with $p_j \neq y_j$, i.e. $p_j < y_j$. The two cases are (A) and (B). The converse is immediate: if (A) or (B) holds, the componentwise inequality and at least one tight coordinate are visible by inspection.

Part (ii). The map $\sigma : z \mapsto -z$ is an order-reversing involution: $p \prec^- y \iff \sigma(p) \prec \sigma(y)$, and likewise $p \prec\!\!\prec^- y \iff \sigma(p) \prec\!\!\prec \sigma(y)$. Applying part (i) to $\sigma(p), \sigma(y)$ yields conditions (A) or (B) on $(-p, -y)$, which translate back to (A^-) or (B^-) on (p, y).

Mutual exclusivity. If both (A) and (B) held, then $p_1 = y_1$ and $p_2 = y_2$, so $p = y$, contradicting the hypothesis. Same for (A^-), (B^-). □

Proof of the positive boundary characterization. Throughout we use Lemma 1(i). For a fixed $y \in \mathbb{R}^2$, define the binding indices

$$i_y^{\le} := \max\{\, j : p_1^j \le y_1 \,\}, \qquad i_y^{<} := \max\{\, j : p_1^j < y_1 \,\},$$

with $\max \emptyset := 0$. Because p_2^j is strictly decreasing in j, the conditions for y to be weakly but not strictly dominated take the compact form

$$\text{(W)} \quad i_y^{\le} \ge 1 \text{ and } y_2 \ge p_2^{\,i_y^{\le}}, \qquad \text{(N)} \quad i_y^{<} = 0 \text{ or } y_2 \le p_2^{\,i_y^{<}}.$$

(W) is weak dominance by the smallest-p_2 candidate among indices with $p_1^j \le y_1$; (N) is the no-strict-dominance constraint applied to all j with $p_1^j < y_1$ at once, again binding at the largest such j.

($\supseteq$). Each piece witnesses (W) and (N), and the displayed pair $(p^\star, y)$ satisfies (A) or (B):

- $\mathcal{B}^0$: $y = (p_1^1, y_2)$, $y_2 > p_2^1$. Then $i_y^{\le} = 1$, $i_y^{<} = 0$, so (W),(N) hold; (p^1, y) satisfies (A).
- $\mathcal{B}_i^1$: $y = (y_1, p_2^i)$, $y_1 \in (p_1^i, p_1^{i+1})$. Then $i_y^{\le} = i_y^{<} = i$, and $y_2 = p_2^i$ saturates both bounds; (p^i, y) satisfies (B).
- $\mathcal{B}_i^2$: $y = (p_1^{i+1}, y_2)$, $y_2 \in (p_2^{i+1}, p_2^i]$. Then $i_y^{\le} = i + 1$, $i_y^{<} = i$, so (W) gives $y_2 \ge p_2^{i+1}$ and (N) gives $y_2 \le p_2^i$; (p^{i+1}, y) satisfies (A).
- $\mathcal{B}^3$: $y = (y_1, p_2^k)$, $y_1 > p_1^k$. Then $i_y^{\le} = i_y^{<} = k$, and $y_2 = p_2^k$ saturates both bounds; (p^k, y) satisfies (B).

In each case $y \notin \mathcal{F}$ follows from the strict inequality on the free coordinate.

($\subseteq$). Let y be weakly but not strictly dominated, $y \notin \mathcal{F}$. Combining (W) and (N) gives the pinch

$$p_2^{\,i_y^{\le}} \ \le\ y_2 \ \le\ p_2^{\,i_y^{<}} \qquad \text{(with the right side vacuous if } i_y^{<} = 0).$$

We split on the position of y_1 relative to $\{p_1^1, \ldots, p_1^k\}$.

- $y_1 < p_1^1$: $i_y^{\le} = 0$, so (W) fails. Impossible.
- $y_1 = p_1^i$, $i \in \{1, \ldots, k\}$: $i_y^{\le} = i$, $i_y^{<} = i - 1$. The pinch reads $p_2^i \le y_2 \le p_2^{i-1}$, with the upper bound vacuous when $i = 1$. Since $y \notin \mathcal{F}$ rules out $y_2 = p_2^i$:
  - $i = 1$: $y_2 > p_2^1$, so $y \in \mathcal{B}^0$.

– $i \geq 2$: $y_2 \in (p_2^i, p_2^{i-1}]$, so $y \in \mathcal{B}_{i-1}^2$.

- $y_1 \in (p_1^i, p_1^{i+1})$, $i \in \{1, \dots, k-1\}$: $i_y^{\leq} = i_y^{<} = i$, so the pinch forces $y_2 = p_2^i$, hence $y \in \mathcal{B}_i^1$.
- $y_1 > p_1^k$: $i_y^{\leq} = i_y^{<} = k$, so $y_2 = p_2^k$, hence $y \in \mathcal{B}^3$.

This is exhaustive. □

Proof of the negative boundary characterization. Let $\sigma : \mathbb{R}^2 \to \mathbb{R}^2$, $\sigma(z) = -z$. Then σ is an involution with

$$q \prec^- y \iff \sigma(q) \prec \sigma(y), \qquad q \prec\!\!\prec^- y \iff \sigma(q) \prec\!\!\prec \sigma(y),$$

so y is weakly but not strictly $\prec^-$-dominated by $\mathcal{F}^-$ if and only if $\sigma(y)$ is weakly but not strictly $\prec$-dominated by $\sigma(\mathcal{F}^-)$. Equivalently,

$$\mathcal{B}^-(\mathcal{F}^-) \;=\; \sigma\big(\mathcal{B}(\sigma(\mathcal{F}^-))\big) \;=\; -\,\mathcal{B}\big(-\mathcal{F}^-\big).$$

Reindexing the reflected front. Define $\tilde{p}^{\,j} := -q^{\,m-j+1}$ for $j = 1, \dots, m$. The hypotheses $q_1^1 < \cdots < q_1^m$ and $q_2^1 > \cdots > q_2^m$ give

$$\tilde{p}_1^1 < \cdots < \tilde{p}_1^m, \qquad \tilde{p}_2^1 > \cdots > \tilde{p}_2^m,$$

so $\tilde{\mathcal{F}} := \{\tilde{p}^{\,1}, \dots, \tilde{p}^{\,m}\} = \sigma(\mathcal{F}^-)$ is a positive Pareto front in the convention of the positive proposition, with $k = m$. By that proposition,

$$\mathcal{B}(\tilde{\mathcal{F}}) \;=\; \mathcal{B}^0 \;\cup\; \bigcup_{i=1}^{m-1} \mathcal{B}_i^1 \;\cup\; \bigcup_{i=1}^{m-1} \mathcal{B}_i^2 \;\cup\; \mathcal{B}^3,$$

with each piece defined in terms of $\tilde{p}^{\,j}$.

Piece-by-piece reflection. We compute σ applied to each piece, using $\tilde{p}_t^{\,j} = -q_t^{\,m-j+1}$ for $t \in \{1, 2\}$ and the identity $\sigma(I \times J) = (-I) \times (-J)$ on rectangles, which sends (a, b) to $(-b, -a)$, $(a, b]$ to $[-b, -a)$, and $(a, +\infty)$ to $(-\infty, -a)$.

- $\sigma(\mathcal{B}^0) = \sigma\big(\{\tilde{p}_1^1\} \times (\tilde{p}_2^1, +\infty)\big) = \{q_1^m\} \times (-\infty, q_2^m) = \mathcal{B}^{-,3}$.
- $\sigma(\mathcal{B}^3) = \sigma\big((\tilde{p}_1^m, +\infty) \times \{\tilde{p}_2^m\}\big) = (-\infty, q_1^1) \times \{q_2^1\} = \mathcal{B}^{-,0}$.
- For $i \in \{1, \dots, m-1\}$, set $\ell := m - i$:
$$\sigma(\mathcal{B}_i^1) = \sigma\big((\tilde{p}_1^i, \tilde{p}_1^{i+1}) \times \{\tilde{p}_2^i\}\big) = (q_1^\ell, q_1^{\ell+1}) \times \{q_2^{\ell+1}\} = \mathcal{B}_\ell^{-,2}.$$
- Similarly, with $\ell := m - i$:
$$\sigma(\mathcal{B}_i^2) = \sigma\big(\{\tilde{p}_1^{i+1}\} \times (\tilde{p}_2^{i+1}, \tilde{p}_2^i]\big) = \{q_1^\ell\} \times [q_2^{\ell+1}, q_2^\ell) = \mathcal{B}_\ell^{-,1}.$$

The map $i \mapsto m - i$ is a bijection of $\{1, \dots, m-1\}$ onto itself, so the unions over i in $\mathcal{B}(\tilde{\mathcal{F}})$ are carried by σ onto the unions over ℓ in $\mathcal{B}^-(\mathcal{F}^-)$. Assembling the four bullets,

$$\mathcal{B}^-(\mathcal{F}^-) = \sigma(\mathcal{B}(\tilde{\mathcal{F}})) = \mathcal{B}^{-,0} \;\cup\; \bigcup_{\ell=1}^{m-1} \mathcal{B}_\ell^{-,1} \;\cup\; \bigcup_{\ell=1}^{m-1} \mathcal{B}_\ell^{-,2} \;\cup\; \mathcal{B}^{-,3}.$$ □

E.3 Local Ideal Points

Definition 4 (Local Ideal Points and Ideal Fronts)**.** Let $\mathcal{F}^+ = \{y^1, \dots, y^k\}$ with $k \geq 2$ and $y_1^1 < \cdots < y_1^k$ (implying $y_2^1 > \cdots > y_2^k$), and let $\mathcal{F}^- = \{w^1, \dots, w^m\}$ with $m \geq 2$ and $w_1^1 < \cdots < w_1^m$ (implying $w_2^1 > \cdots > w_2^m$). The *positive* and *negative ideal fronts* are the sets of *local positive* and *local negative ideal points*, respectively, given by

$$\mathcal{L}^+ \;=\; \{\ell_+^i\}_{i=1}^{k-1}, \qquad \ell_+^i \;:=\; (y_1^i,\; y_2^{i+1}),$$
$$\mathcal{L}^- \;=\; \{\ell_-^i\}_{i=1}^{m-1}, \qquad \ell_-^i \;:=\; (w_1^{i+1},\; w_2^i).$$

Proposition 5 (Local Ideal Points Dominate Their Neighbors)**.** For each $i \in \{1, \dots, k-1\}$:

$$\ell_+^i \prec^+ y^i \ \wedge \ \ell_+^i \prec^+ y^{i+1},$$

and for each $i \in \{1, \dots, m-1\}$:

$$\ell_-^i \prec^- w^i \ \wedge \ \ell_-^i \prec^- w^{i+1}.$$

Proof. $\ell_+^i \prec^+ y^i$: $y_1^i \leq y_1^i$ and $y_2^{i+1} < y_2^i$, with the second inequality strict.

$\ell_+^i \prec^+ y^{i+1}$: $y_1^i < y_1^{i+1}$ and $y_2^{i+1} \leq y_2^{i+1}$, with the first inequality strict.

It goes analogously for the local negative ideal points:

$\ell_-^i \prec^- w^i$: $w_1^{i+1} > w_1^i$ and $w_2^i \geq w_2^i$, with the first inequality strict.

$\ell_-^i \prec^- w^{i+1}$: $w_1^{i+1} \geq w_1^{i+1}$ and $w_2^i > w_2^{i+1}$, with the second inequality strict. □

Proposition 6 (Local ideal points are non-dominated)**.** The ideal fronts $\mathcal{L}^+$ and $\mathcal{L}^-$ are mutually non-dominated and non-dominated by the original Pareto fronts:

(i) For every $\ell_+^i \in \mathcal{L}^+$, there exists no $y \in \mathcal{F}^+ \cup (\mathcal{L}^+ \setminus \{\ell_+^i\})$ such that $y \prec^+ \ell_+^i$.

(ii) For every $\ell_-^i \in \mathcal{L}^-$, there exists no $w \in \mathcal{F}^- \cup (\mathcal{L}^- \setminus \{\ell_-^i\})$ such that $w \prec^- \ell_-^i$.

Proof. **(i) Positive case.** Let $\ell_+^i = (y_1^i, y_2^{i+1}) \in \mathcal{L}^+$.

Non-domination by $\mathcal{F}^+$. Suppose $y^j \prec^+ \ell_+^i$ for some j, i.e. $y_1^j \leq y_1^i$ and $y_2^j \leq y_2^{i+1}$, with at least one strict inequality. Since y_1 is strictly increasing: $y_1^j \leq y_1^i \Rightarrow j \leq i$. Since y_2 is strictly decreasing: $y_2^j \leq y_2^{i+1} \Rightarrow j \geq i+1$. These two conditions contradict each other, so no such j exists.

Non-domination by $\mathcal{L}^+$. Suppose $\ell_+^j \prec^+ \ell_+^i$ for some $j \neq i$, i.e. $y_1^j \leq y_1^i$ and $y_2^{j+1} \leq y_2^{i+1}$, with at least one strict inequality. Since y_1 is strictly increasing: $y_1^j \leq y_1^i \Rightarrow j \leq i$. Since y_2 is strictly decreasing: $y_2^{j+1} \leq y_2^{i+1} \Rightarrow j \geq i$. Hence $j = i$, so both inequalities hold with equality, contradicting the requirement of at least one being strict.

(ii) Negative case. Let $\ell_-^i = (w_1^{i+1}, w_2^i) \in \mathcal{L}^-$. Recall $n \prec^- m \iff m \prec^+ n$.

Non-domination by $\mathcal{F}^-$. Suppose $w^j \prec^- \ell_-^i$ for some j, equivalently $\ell_-^i \prec^+ w^j$, i.e. $w_1^{i+1} \leq w_1^j$ and $w_2^i \leq w_2^j$, with at least one strict inequality. Since w_1 is strictly increasing: $w_1^{i+1} \leq w_1^j \Rightarrow j \geq i+1$. Since w_2 is strictly decreasing: $w_2^i \leq w_2^j \Rightarrow j \leq i$. These two conditions contradict each other, so no such j exists.

Non-domination by $\mathcal{L}^-$. Suppose $\ell_-^j \prec^- \ell_-^i$ for some $j \neq i$, equivalently $\ell_-^i \prec^+ \ell_-^j$, i.e. $w_1^{i+1} \leq w_1^{j+1}$ and $w_2^i \leq w_2^j$, with at least one strict inequality. Since w_1 is strictly increasing: $w_1^{i+1} \leq w_1^{j+1} \Rightarrow j \geq i$. Since w_2 is strictly decreasing: $w_2^i \leq w_2^j \Rightarrow j \leq i$. Hence $j = i$, so both inequalities hold with equality, contradicting the requirement of at least one being strict. □