



**HAL**  
open science

# Multiobjective Hypervolume Based ISOOMOO Algorithms Converge with At Least Sublinear Speed to the Entire Pareto Front

Eugénie Marescaux, Anne Auger

► **To cite this version:**

Eugénie Marescaux, Anne Auger. Multiobjective Hypervolume Based ISOOMOO Algorithms Converge with At Least Sublinear Speed to the Entire Pareto Front. 2021. hal-03198414v2

**HAL Id: hal-03198414**

**<https://hal.science/hal-03198414v2>**

Preprint submitted on 16 Sep 2021 (v2), last revised 22 Apr 2022 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1                   **MULTIOBJECTIVE HYPERVOLUME BASED ISOOMOO**  
2   **ALGORITHMS CONVERGE WITH AT LEAST SUBLINEAR SPEED TO**  
3                   **THE ENTIRE PARETO FRONT**

4                   EUGÉNIE MARESCAUX AND ANNE AUGER

5       **Abstract.** In multiobjective optimization, one is interested in finding a good approximation of the  
6 Pareto set and the Pareto front, i.e the sets of best compromises in the decision and objective spaces,  
7 respectively. In this context, we introduce a new algorithm framework, Incremental SingleObjective Op-  
8 timization for MultiObjective Optimization (ISOOMOO) for approximating the Pareto front with an  
9 increasing number of points. We focus on HV-ISOOMOO, its instantiation with the hypervolume indi-  
10 cator, a set-quality indicator which is widely used for algorithms design and performance assessment.  
11 HV-ISOOMOO algorithms approximate the Pareto front by greedily maximizing the hypervolume. We  
12 study the convergence to the entire Pareto front of HV-ISOOMOO coupled with perfect singleobjec-  
13 tive optimization. The convergence is defined as the convergence of the hypervolume of the sets of all  
14 meta-iterations incumbents towards the hypervolume of the Pareto front. We prove tight lower bounds  
15 on the speed of convergence for convex and bilipschitz Pareto fronts in  $O(1/n^c)$  with  $c = 1$  and  $c \leq 1$ ,  
16 respectively. The index  $n$  denotes the number of meta-iterations of HV-ISOOMOO. For convex Pareto  
17 fronts, the convergence is in  $\Theta(1/n)$ , namely the fastest convergence achievable by a biobjective opti-  
18 mization algorithm. These are the first results on the speed of convergence of multiobjective optimization  
19 algorithms towards the entire Pareto front. We also analyze theoretically the asymptotic convergence  
20 behavior.

21       **Key words.** multiobjective optimization, convergence, hypervolume, Pareto front

22       **AMS subject classifications.** 90C29, 90C30

23       **1. Introduction.** Real-world problems often involve the optimization of several con-  
24 flicting objectives. The solution of such problems is the set of non-dominated *decision*  
25 *vectors* (vectors of the search space), the *Pareto set*. It is defined as the set of solutions  
26 that cannot be improved along one objective without degrading along another one. Its  
27 image in the objective space is the Pareto front. A decision maker is then often involved to  
28 choose, based on its preferences, a single best compromise. The shape of the Pareto front  
29 informs on the trade-off between objectives. Many algorithms such as evolutionary algo-  
30 rithms approximate the Pareto front with a number of points fixed in the beginning. But  
31 some algorithms, in particular stemming from direct search methods [1, 7, 10, 11] aim at  
32 approximating the Pareto set or Pareto front with as many well-distributed points as pos-  
33 sible. Ideally, the quality of the Pareto front approximation increases with time without  
34 stagnating and such algorithms are referred to in the sequel as *anytime* algorithms.

35       The speed of convergence towards a critical decision vector or a vector of the Pareto  
36 front have been examined for many algorithms such as  $(1+1)$  evolutionary multiobjective  
37 algorithms [5] or Newton’s method [14]. Convergence speeds are typically similar to the  
38 ones obtained for singleobjective optimization. They both apply to a convergence to-  
39 wards a single point. Their analysis is sometimes reduced to the study of the convergence  
40 of a singleobjective optimization algorithm. The convergence of anytime algorithms to-  
41 wards the whole Pareto set or front is of a different kind because these are sets and not  
42 points. It has already been theoretically investigated for some algorithms [10] and more  
43 abstract frameworks [19], but analysis of the speed of convergence are missing. Addition-  
44 ally, empirical studies typically focus on determining which algorithm is faster and do not  
45 provide precise information on the speed of convergence such as order of convergence or

46 complexity. Yet, while largely overlooked, investigating the speed of convergence either  
 47 theoretically or empirically is important. In this context, it has been proven that con-  
 48 vergence of biobjective optimization algorithms towards the whole Pareto front is always  
 49 sublinear in the number of function evaluations, at least when measuring convergence  
 50 with the hypervolume indicator [16] or the multiplicative  $\epsilon$ -indicator [8], and thus much  
 51 smaller than typical speeds of convergence to a single point. The hypervolume and the  
 52 multiplicative  $\epsilon$ -indicator are set-quality indicators widely used in multiobjective opti-  
 53 mization, both to guide algorithms and for performance assessment. The hypervolume is  
 54 at the core of all known strictly Pareto-compliant indicators [20].

55 In this paper, we introduce a new anytime algorithm framework, Incremental Sin-  
 56 gleObjective Optimization for MultiObjective Optimization (ISOOMOO). We focus on its  
 57 instantiation with the hypervolume indicator, namely HV-ISOOMOO algorithms. These  
 58 algorithms try to greedily maximize the hypervolume by adding points with the largest  
 59 hypervolume contribution. These points are obtained by running a singleobjective opti-  
 60 mizer. A greedy idea has already been used in the selection part of some multiobjective  
 61 optimization evolutionary algorithms such as SMS-EMOA to find a set of  $p$  points with a  
 62 large hypervolume [6]. The hypervolume of such discrete greedy approximation is proven  
 63 to be at least  $(e-1)/e$  times the one of the  $p$ -optimal distribution [17]. To the best of our  
 64 knowledge, we provide the first continuous equivalent of this result. We investigate the  
 65 speed of convergence of HV-ISOOMOO towards the whole Pareto front in the ideal case  
 66 of perfect singleobjective optimization, measuring the convergence with the hypervolume.  
 67 For convex and bilipschitz Pareto fronts, we prove that the convergence is in  $O(1/n^c)$  with  
 68  $c = 1$  and  $c \leq 1$ , respectively, with  $n$  being the number of singleobjective optimization  
 69 runs performed. For convex Pareto fronts, the convergence is exactly in  $\Theta(1/n)$  as no  
 70 biobjective algorithm can converge faster to the Pareto front [16]. Additionally, we prove  
 71 that for simultaneously bilipschitz and smooth enough Pareto fronts doubling the number  
 72 of points in the approximation divides the optimality gap by a factor which converges  
 73 asymptotically to two. In the proof process, we obtain interesting intermediary results  
 74 such as bounds on the normalized maximum hypervolume and a geometric interpretation  
 75 of optimality conditions.

76 The paper is organized as follows. In Section 2, we lay the foundations of the prob-  
 77 lem. In Section 3, we prove preliminary results later used to investigate convergence. In  
 78 Section 4, we derive lower bounds on the speed of convergence of HV-ISOOMOO coupled  
 79 with perfect singleobjective optimization and an insight on its asymptotic convergence  
 80 behavior.

81 *Notations and conventions.* For  $a, b \in \mathbb{N}$ , we note  $\llbracket a; b \rrbracket$  the set  $\{a, a+1, \dots, b-1, b\}$ .  
 82 For a vector  $u \in \mathbb{R}^2$ , we note  $u_1$  and  $u_2$  respectively its first and its second coordinate. If  
 83 the vector notation already contains an index, we separate the two indices with a comma.  
 84 For simplicity sake, we often replace the set  $\{u\}$  by  $u$  in the notations. We say that a  
 85 function  $f : \mathbb{R} \rightarrow \mathbb{R}$  is decreasing (resp. strictly decreasing) when for all  $x < y$ , we have  
 86  $f(x) \geq f(y)$  (resp.  $f(x) > f(y)$ ). We only consider two-dimensional objective spaces and  
 87 refer to the Lebesgue measure of a set as its area.

88 **2. Background, algorithm framework and assumptions.** We lay in this sec-  
 89 tion the foundations of the problem we analyze. First, we recall some classic concepts  
 90 of multiobjective optimization. Then, we introduce the ISOOMOO class of algorithms  
 91 and its hypervolume based instantiation HV-ISOOMOO. We also formalize a mathemat-  
 92 ical abstraction of HV-ISOOMOO coupled with perfect singleobjective optimization, the

93 greedy set sequences. Finally, we examine our assumptions on the biobjective optimization  
 94 problem.

95 **2.1. Biobjective optimization problems, the Pareto front and the hyper-**  
 96 **volume indicator.** We consider a biobjective minimization problem:

97 (2.1) 
$$\min_{x \in \Omega \subset \mathbb{R}^d} F(x)$$
  
 98

99 with  $F : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^2 : x \mapsto (F_1(x), F_2(x))$ . We define two dominance relations for  
 100 vectors in the objective space. We say that  $u$  weakly dominates  $v$  denoted by  $u \preceq v$  if  
 101  $u_1 \leq v_1$  and  $u_2 \leq v_2$  and that  $u$  dominates  $v$  denoted by  $u \prec v$  if  $u \preceq v$  and  $u \neq v$ . A  
 102 vector of the objective space  $\mathbb{R}^2$  is said *feasible* when it belongs to  $F(\Omega)$ . Solving the  
 103 optimization problem consists in finding a good approximation of the *Pareto front*, the  
 104 set of non-dominated feasible vectors,  $\{F(X) : X \in \Omega, \forall Y \in \Omega, F(Y) \not\prec F(X)\}$ . We  
 105 restrict ourselves to Pareto fronts with an explicit representation:

106 (2.2) 
$$\text{PF}_f = \{(x, f(x)) : x \in [x_{\min}, x_{\max}]\}$$

108 with  $f : \mathbb{R} \mapsto \mathbb{R}$  decreasing. We denote by  $u_{\min} := (x_{\min}, f(x_{\min}))$  and  $u_{\max} :=$   
 109  $(x_{\max}, f(x_{\max}))$  the extreme vectors of the Pareto front. Likewise, we denote by  $\tilde{u}_{\min,r} :=$   
 110  $(\tilde{x}_{\min,r}, f(\tilde{x}_{\min,r}))$  and  $\tilde{u}_{\max,r} := (\tilde{x}_{\max,r}, f(\tilde{x}_{\max,r}))$  the extremes vectors of the part of  
 111 the Pareto front dominating a reference point  $r$ , with  $\tilde{x}_{\min,r} := \max(x_{\min}, f^{-1}(r_2))$  and  
 112  $\tilde{x}_{\max,r} := \min(x_{\max}, r_1)$ . The vector  $(x_{\max}, f(x_{\min}))$  is called the *nadir point*. All these  
 notations are illustrated in Figure 1.

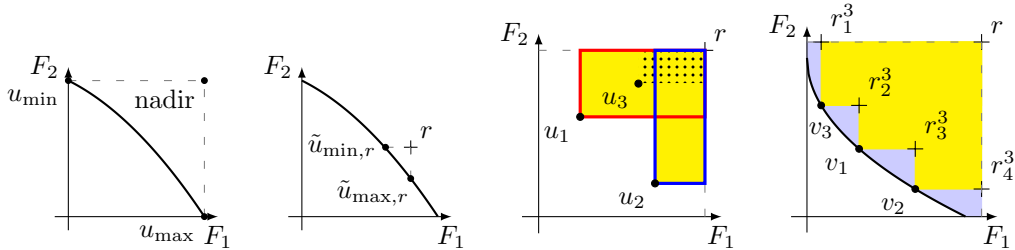


Fig. 1: Illustration of notations. The extreme vectors  $u_{\min}$  and  $u_{\max}$  and the nadir point  
 (leftmost) ; the extreme vectors relative to the reference point  $r$   $\tilde{u}_{\min,r}$  and  $\tilde{u}_{\max,r}$  (left) ;  
 three vectors  $u_1$ ,  $u_2$  and  $u_3$  and the regions weakly dominated by them and dominating  
 $r$ ,  $\mathcal{D}_{u_1}^r$  (red box),  $\mathcal{D}_{u_2}^r$  (blue box) and  $\mathcal{D}_{u_3}^r$  (dotted box) (right) ; the greedy set  $\mathcal{S}_3 = \{v_1, v_2, v_3\}$ , its four gap  
 regions (blue shaded) and the associated reference points (rightmost).

113  
 114 The *hypervolume* with respect to a reference point  $r$  of a set  $S$  of objective vectors  
 115 is the Lebesgue measure of the region of the objective space dominated by  $S$  and strictly  
 116 dominating the reference point  $r$ . We denote it  $\text{HV}_r(S)$ . When no vector of the Pareto  
 117 front dominates the reference point  $r$ ,  $\text{HV}_r(S) = 0$  for any set  $S$  of feasible points of the  
 118 objective space. Since this particular case is not interesting, we only consider reference  
 119 points dominated by at least one vector of the Pareto front from now on. We refer to  
 120 such reference points as *valid*.

121 The region of the objective space dominated by  $S$  and dominating  $r$  (see the righthand  
 122 plot of Figure 1) is denoted by  $\mathcal{D}_S^r$  and formally defined as:

$$123 \quad (2.3) \quad \mathcal{D}_S^r = \{w \in \mathbb{R}^2 : \exists u \in S : u \preceq w \prec r\} .$$

125 The hypervolume of a set  $S$  relative to the reference point  $r$  equals  $\lambda(\mathcal{D}_S^r)$  with  $\lambda(\cdot)$  being  
 126 the Lebesgue measure. The set  $\mathcal{D}_S^r$  is the union of the  $\mathcal{D}_u^r$  for  $u \in S$ ,  $\mathcal{D}_u^r$  being the  
 127 rectangle  $[u_1, r_1] \times [u_2, r_2]$  when  $u$  dominates  $r$  and  $\emptyset$  otherwise, see the righthand plot of  
 128 Figure 1. Note that the  $\mathcal{D}_u^r$  are not disjoint.

129 We use the hypervolume to characterize the convergence of a set  $S$  of objective vectors  
 130 to the entire Pareto front. For a fixed valid reference point  $r$ , a set  $S$  is said to converge  
 131 to the Pareto front when the hypervolume difference  $\text{HV}_r(\text{PF}_f) - \text{HV}_r(S)$  converges  
 132 to 0. We define the *optimality gap* of  $S$  with respect to a valid reference point  $r$  as  
 133  $\text{HV}_r(\text{PF}_f) - \text{HV}_r(S)$ . Another quantity of interest is how much adding a vector to a set  
 134 affects its hypervolume. The *hypervolume improvement* with respect to  $r$  of the vector  $u$   
 135 to the set  $S$  is  $\text{HVI}_r(u, S) = \text{HV}_r(S \cup \{u\}) - \text{HV}_r(S)$ . We also use the term hypervolume  
 136 improvement to refer to the hypervolume increase of an increasing<sup>1</sup> sequence of sets. More  
 137 precisely, the hypervolume improvement at iteration  $n$  of an increasing sequence  $(\mathcal{S}_n)_{n \in \mathbb{N}^*}$   
 138 is  $\text{HV}_r(\mathcal{S}_{n+1}) - \text{HV}_r(\mathcal{S}_n)$ .

139 **2.2. The ISOOMOO framework, its HV-ISOOMOO instantiation and the**  
 140 **associated greedy set sequences.** The Incremental SingleObjective Optimization for  
 141 MultiObjective Optimization (ISOOMOO) framework builds incrementally an increasing  
 142 sequence  $(\mathcal{I}_n)_{n \in \mathbb{N}^*}$  of sets of vectors of the objective space. The pseudocode of ISOOMOO  
 143 is given in Algorithm 2.1, where the current value of  $\mathcal{I}_n$  is denoted by  $\mathcal{I}$ . At each so-  
 144 called meta-iteration, a singleobjective maximization algorithm SOOPTIMIZER (line 2 in  
 145 Algorithm 2.1) is run on the criterion  $X \in \Omega \subset \mathbb{R}^d \mapsto J(\mathcal{I}, X)$  and the resulting solution is  
 146 added to  $\mathcal{I}$  (line 3 in Algorithm 2.1). We use the term meta-iteration to separate between  
 147 the (meta-)iterations of ISOOMOO and the iterations of SOOPTIMIZER. Since the set  
 148  $\mathcal{I}$  is composed of the final objective incumbents of previous runs of SOOPTIMIZER and  
 149 (ideally) provides an approximation of the Pareto front, we call it *final incumbents Pareto*  
 150 *front approximation*.

151 The singleobjective optimization procedure may vary between meta-iterations. More  
 152 precisely, the run of SOOPTIMIZER depends on data about precedent runs stored in  
 153  $D$  (line 4 in Algorithm 2.1). This allows to alternate between various singleobjective  
 154 optimization algorithms with different features, but also to adapt their initialization. This  
 155 could be done by storing in  $D$  an iteration index and the final search space incumbents  
 156 of SOOPTIMIZER runs.

---

**Algorithm 2.1** ISOOMOO Framework

---

```

1: while not stopping criterion do
2:    $Y, d \leftarrow \text{SOOPTIMIZER}(X \mapsto J(\mathcal{I}, X), D)$ 
3:    $\mathcal{I} \leftarrow \mathcal{I} \cup \{F(Y)\}$  # update of the approximation of the Pareto front
4:    $D \leftarrow D \cup \{d\}$  # update of the data collected
5: end while
```

---

<sup>1</sup>A sequence of set  $\{A_n, n \geq 0\}$  is increasing if the following inclusions  $A_0 \subset A_1, \dots \subset A_n \subset \dots$  hold.

157 In this paper, we study HV-ISOOMOO, an instantiation of ISOOMOO for which the  
 158 criterion  $J(\mathcal{I}, \cdot)$  relates to the hypervolume improvement to  $\mathcal{I}$ . Formally, HV-ISOOMOO  
 159 is a class of algorithms derived from ISOOMOO for which the maximization of the cri-  
 160 terion  $J$  is *compliant* with the maximization of the hypervolume improvement as defined  
 161 below.

162 *Assumption 2.1.* (Compliance to  $\text{HVI}_r$  maximization) The maximization of a crite-  
 163 rion  $J$  as in ISOOMOO is *compliant* with the maximization of  $\text{HVI}_r$  if for any set  $\mathcal{I}$  of  
 164 objective vectors, we have

$$165 \quad (2.4) \quad \operatorname{argmax}_{X \in \mathbb{R}^d} J(\mathcal{I}, X) = \operatorname{argmax}_{X \in \mathbb{R}^d} \text{HVI}_r(F(X), \mathcal{I}) .$$

166 We define an HV-ISOOMOO algorithm relative to the reference point  $r$  as an ISOOMOO  
 167 algorithm as described in [Algorithm 2.1](#) where the criterion  $J$  satisfies [Assumption 2.1](#). At  
 168 each meta-iteration  $n$ , an HV-ISOOMOO algorithm seeks a feasible vector maximizing  
 169 the hypervolume improvement to the final incumbents Pareto front approximation  $\mathcal{I}_n$ .  
 170 Ideally, when  $n$  goes to infinity, the non-dominated subset of  $(\mathcal{I}_n)_{n \in \mathbb{N}^*}$  converges to the  
 171 (entire) Pareto front which maximizes the hypervolume. In other words, HV-ISOOMOO  
 172 algorithms try to approximate the Pareto front with a greedy approach.

173 **DEFINITION 2.2.** *We define the convergence of an HV-ISOOMOO algorithm as the*  
 174 *convergence of  $\text{HV}_r(\mathcal{I}_n)$  towards  $\text{HV}_r(\text{PF}_f)$ .*

175 The performance of a specific HV-ISOOMOO algorithm depends crucially on the  
 176 choice of the criterion  $J$ . In this respect,  $\text{HVI}_r(\mathcal{I}, F(\cdot))$  itself is not a good candidate for  
 177  $J(\mathcal{I}, \cdot)$ . Indeed, it is constant equal to zero in the region dominated by  $\mathcal{I}$ , which makes it  
 178 difficult to optimize. A criterion whose maximization is compliant with the maximization  
 179 of the hypervolume improvement and designed to be easier to optimize has already been  
 180 introduced in [\[18\]](#) under the name *uncrowded hypervolume improvement* (UHVI). For  
 181  $F(X)$  not dominated by  $\mathcal{I}$ ,  $\text{UHVI}_r$  and  $\text{HVI}_r$  are equal. Otherwise, in the region where  
 182 the hypervolume improvement is null,  $\text{UHVI}_r$  is negative and equals minus the distance  
 183 to the empirical non-dominated front of the set  $\mathcal{I}$  relative to  $r$ . It is easy to see that  
 184  $\text{UHVI}_r$  satisfies [\(2.4\)](#).

185 The choice of SOOPTIMIZER also plays a key role in the performance of an HV-  
 186 ISOOMOO algorithm. In this paper, we analyze the HV-ISOOMOO framework under  
 187 the assumption of perfect singleobjective optimization formalized below.

188 *Assumption 2.3* (Perfect singleobjective optimization). At every meta-iteration  $n$ ,  
 189 for any final incumbents Pareto front approximation  $\mathcal{I}_n$ , the run of SOOPTIMIZER (line 2  
 190 in [Algorithm 2.1](#)) returns  $Y \in \operatorname{argmax}_{X \in \Omega} J(\mathcal{I}_n, F(X))$ .

191 The assumption of perfect singleobjective optimization is reminiscent to the assump-  
 192 tion of perfect line search which is common in the analysis of gradient based methods [\[12\]](#).  
 193 Under this assumption, all choices of criterions verifying [Assumption 2.1](#) are equivalent.  
 194 The convergence of HV-ISOOMOO coupled with perfect singleobjective optimization is a  
 195 necessary condition for the soundness of the approach. Furthermore, lower bounds on the  
 196 speed of convergence of a real instantiation of HV-ISOOMOO could be obtained by com-  
 197 bining lower bounds on the speed of convergence of HV-ISOOMOO under [Assumption 2.3](#)  
 198 of perfect singleobjective optimization with the ones of singleobjective optimization algo-  
 199 rithms.

200 We introduce below mathematical abstractions of the HV-ISOOMOO framework un-  
 201 der [Assumption 2.3](#) of perfect singleobjective optimization, greedy sequences and greedy  
 202 set sequences.

203 **DEFINITION 2.4** (Greedy sequence and greedy set sequence). *Given a valid reference*  
 204 *point  $r$ , we define as greedy sequence relative to  $r$ , a sequence  $(v_n)_{n \in \mathbb{N}^*}$  satisfying*

$$205 \quad (2.5) \quad v_1 \in \arg \max_{v \in F(\Omega)} HV_r(v) \text{ and}$$

$$206 \quad (2.6) \quad v_{n+1} \in \arg \max_{v \in F(\Omega)} HV_r(\{v_1, \dots, v_n, v\}) \text{ for all } n \geq 1 .$$

207  
 208 *The greedy set sequence  $(\mathcal{S}_n)_{n \in \mathbb{N}^*}$  associated to the greedy sequence  $(v_n)_{n \in \mathbb{N}^*}$  is composed*  
 209 *of the greedy sets  $\mathcal{S}_n := \{v_k, k \leq n\}$ .*

210 There is a bijection between greedy sequences and greedy set sequences. The  $n$ -th element  
 211 of the greedy sequence  $(v_n)_{n \in \mathbb{N}^*}$  associated to a greedy set sequence  $(\mathcal{S}_n)_{n \in \mathbb{N}^*}$  is simply  
 212 the unique element of  $\mathcal{S}_n \setminus \mathcal{S}_{n-1}$  if  $n > 1$  and of  $\mathcal{S}_1$  if  $n = 1$ .

213 The recurrence relation of the greedy sequence (2.6) is equivalent to  $v_{n+1}$  belonging  
 214 to  $\arg \max_{v \in F(\Omega)} HVI_r(v, \mathcal{S}_n)$  for all  $n \geq 1$ . It is immediate to see that under [Assump-](#)  
 215 [tion 2.3](#), the final incumbents generated by HV-ISOOMOO constitute a greedy sequence  
 216 while the final incumbents Pareto front approximations form the associated greedy set  
 217 sequence  $(\mathcal{I}_n)_{n \in \mathbb{N}^*}$ . The indices  $n$  of both sequences correspond to HV-ISOOMOO meta-  
 218 iterations. In this paper, we derive convergence results for greedy set sequences, which  
 219 transfer to HV-ISOOMOO under [Assumption 2.3](#).

220 As we will see in [Subsection 3.1](#), the problem of maximizing the hypervolume improve-  
 221 ment to a fixed set can be rewritten as the maximum of a finite number of hypervolume  
 222 maximization problems. Therefore, we can infer from [4, Theorem 1] that as soon as the  
 223 Pareto front is lower semi-continuous, there exists a greedy sequence, and thus a greedy  
 224 set sequence.

225 **PROPOSITION 2.5.** *If the Pareto front is described by a lower semi-continuous func-*  
 226 *tion  $f$ , then there exists a greedy sequence  $(v_n)_{n \in \mathbb{N}^*}$  relative to any valid reference point.*

227 *Proof.* If  $f$  is lower semi-continuous, then for any valid reference point  $r$ , the maxi-  
 228 mum of  $HV_r(\cdot)$  exists, see [4, Theorem 1]. Therefore, there exists a vector verifying (2.5)  
 229 and the problem of maximizing the maximum of a finite number of hypervolume functions  
 230 defined in (3.3) admits a solution. Since (2.6) and (3.3) are equivalent by [Lemma 3.5](#), we  
 231 can build a sequence  $(v_n)_{n \in \mathbb{N}^*}$  verifying (2.5) and (2.6), namely a greedy sequence.  $\square$

232 Additionally, since the hypervolume indicator associated to a valid reference point is  
 233 strictly Pareto-compliant (see [15]), this sequence is composed of vectors of the Pareto  
 234 front.

235 **PROPOSITION 2.6.** *If the Pareto front is described by a lower semi-continuous func-*  
 236 *tion  $f$ , then any vector of a greedy sequence relative to a valid reference point  $r$  belongs to*  
 237 *the Pareto front. Consequently, for such Pareto front and reference point and under [As-](#)*  
 238 [s-](#)  
 239 *umption 2.3, all final incumbents Pareto front approximations  $\mathcal{I}_n$  of an HV-ISOOMOO*  
*algorithm relative to  $r$  are subsets of the Pareto front.*

240 *Proof.* Since for any valid reference point  $r$ ,  $HV_r(\cdot)$  is strictly Pareto-compliant [15],  
 241 its maximum is non-dominated and belongs to the Pareto front. Thus, in particular, a  
 242 vector  $v_1$  verifying (2.5) belongs to the Pareto front. Additionally, by [Lemma 3.5](#), every



243 solution of (2.6) verifies (3.3) and is solution of at least one hypervolume maximization  
 244 problem, and thus also belong to the Pareto front.  $\square$

245 Yet, in general, there exists more than one greedy sequence, and thus greedy set sequence.  
 246 For example, there are infinitely many greedy sequences associated to any affine Pareto  
 247 front with a reference point dominating the nadir point. This statement relies on the fact  
 248 that the unique maximizer of the hypervolume relative to a reference point  $r$  dominating  
 249 the nadir point is the middle of the section of the Pareto front dominating  $r$ , see [3,  
 250 Theorem 5]. As a consequence, the middle of the section of the Pareto front dominating  $r$   
 251 is the only candidate for  $v_1$  but  $v_2$  can be either at  $1/4$  or at  $3/4$  of this section. Similarly,  
 252  $v_3$  has to be in the position where  $v_2$  is not but  $v_4$  can be at  $1/8, 3/8, 5/8$  or  $7/8$  of the  
 253 section of the Pareto front dominating  $r$ . For any  $n$ , we can find an iteration  $m$  such  
 254 that  $v_m$  can be placed at  $2^n$  different points, whatever the  $m - 1$  first terms of the greedy  
 255 sequence are.

256 **2.3. Assumptions on the Pareto front and the objective functions.** We  
 257 present and discuss here the assumptions on the function  $f$  describing the Pareto front  
 258 under which we derive convergence results. We typically assume that the function  $f$  is  
 259 bilipschitz, convex or simultaneously bilipschitz and with a Hölder continuous derivative.  
 260 Under any of these three assumptions,  $f$  is continuous. For the sake of conciseness, we  
 261 transfer the properties of  $f$  to the Pareto front. For example, we call *convex Pareto*  
 262 *front* a Pareto front described by a convex function. We recall that a function  $f$  is  
 263 Hölder continuous with exponent  $\alpha$ , namely  $\mathcal{C}^{1,\alpha}$ , when there exists  $H \geq 0$  such that  
 264  $|f(x) - f(y)| \leq H \times |x - y|^\alpha$  for all  $x, y$  [13]. We note  $[f]_\alpha$  the *minimum Hölder coefficient*  
 265 with respect to the exponent  $\alpha$  of a  $\mathcal{C}^{1,\alpha}$  function  $f$ , that is  $[f]_\alpha := \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|^\alpha}$ .  
 266 When needed, we detail the bilipschitz constants and say that a bilipschitz function  $f$  is  
 267  $(L_{\min}, L_{\max})$ -bilipschitz if for all  $x, y \in [x_{\min}, x_{\max}]$ , we have  $L_{\min} \times |x - y| \leq |f(x) - f(y)| \leq$   
 268  $L_{\max} \times |x - y|$  where  $L_{\max} \geq L_{\min} > 0$ . We also talk of *affine Pareto fronts* when  
 269  $f(x) = ax + b$  with  $a < 0$  and  $b \in \mathbb{R}$ . As they form a line in the biobjective case, they are  
 270 usually referred to as linear Pareto fronts. They provide good examples to illustrate a  
 271 point and help to understand the results we prove on the asymptotic convergence behavior.

272 We remind below sufficient conditions on the search space and on the objective func-  
 273 tions which guarantee that  $f$  is convex and bilipschitz.

274 **PROPOSITION 2.7.** *Given a biobjective minimization problem as in (2.1) whose Pareto*  
 275 *front is described by a function  $f$ . If  $F_1$  and  $F_2$  are respectively  $(L_{\min,1}, L_{\max,1})$ -bilipschitz*  
 276 *and  $(L_{\min,2}, L_{\max,2})$ -bilipschitz, then  $f$  is  $(\frac{L_{\min,2}}{L_{\max,1}}, \frac{L_{\max,2}}{L_{\min,1}})$ -bilipschitz. If the search space*  
 277  *$\Omega$  and the objective functions  $F_1$  and  $F_2$  are convex, then  $f$  is convex.*

278 The proofs of this proposition can be found for instance in [16]. The conditions on  $F_1$ ,  
 279  $F_2$  and  $\Omega$  are sufficient but non-necessary conditions. Indeed, adding small discontinuity in  
 280 the objective functions far from the Pareto set makes them non-convex and non-bilipschitz  
 281 without modifying the Pareto front.

282 Representing  $F_1$  values on the absciss and  $F_2$  values on the ordinate instead of the  
 283 converse is an arbitrary choice. When  $f$  is a bijection, if we chose to represent the  $F_2$   
 284 values on the  $x$ -axis instead of on the  $y$ -axis, we would have another representation of  
 285 the Pareto front :  $\{(y, f^{-1}(y)) : y \in [f(x_{\max}); f(x_{\min})]\}$ . If so, the inverse function  $f^{-1}$   
 286 would play the role of  $f$ . It is interesting to notice that the choice of the objective function  
 287 represented on the horizontal axis does not impact whether the function characterizing



288 the Pareto front is bilipschitz or convex. Indeed,  $f$  being bilipschitz is equivalent to both  
 289  $f$  and  $f^{-1}$  being lipschitz. Additionally, we can prove that given that the function  $f$  is  
 290 decreasing,  $f$  being convex is equivalent to its inverse  $f^{-1}$  being convex. The proof of  
 291 this property is straightforward. If  $f$  is convex, then for all  $x, y \in \mathbb{R}$ ,  $f\left(\frac{f^{-1}(x)+f^{-1}(y)}{2}\right)$  is  
 292 smaller than  $\frac{f(f^{-1}(x))+f(f^{-1}(y))}{2}$ , that is  $\frac{x+y}{2}$ . Since  $f$  and therefore  $f^{-1}$  are decreasing,  
 293 by composing by  $f^{-1}$  each side of the inequality, we obtain a characterization of the  
 294 convexity of  $f^{-1}$ : for all  $x, y \in \mathbb{R}$ ,  $\frac{f^{-1}(x)+f^{-1}(y)}{2}$  is larger than  $f^{-1}\left(\frac{x+y}{2}\right)$ .

295 **3. Preliminary results.** In this section, we present preliminary results which are  
 296 crucial for the analysis of the convergence of HV-ISOOMOO. While we expose them as  
 297 tools for convergence analysis, they are also interesting for their own sake.

298 **3.1. Decomposition of the optimality gap using gap regions.** The optimality  
 299 gap is the Lebesgue measure of the *total gap region* introduced below.

300 DEFINITION 3.1. *The total gap region of  $S$  with respect to a fixed valid reference*  
 301 *point  $r$ ,  $\mathcal{G}_S^r$ , is defined as the region of the objective space which dominates  $r$  and is*  
 302 *weakly dominated by  $PF_f$  but not by  $S$ , namely  $\mathcal{D}_{PF_f}^r \setminus \mathcal{D}_S^r$ .*

303 We introduced  $\mathcal{D}_S^r$  in (2.3). Its Lebesgue measure is  $HV_r(S)$ .

304 When  $S$  is a subset of the Pareto front dominating the reference point  $r$ , the total  
 305 gap region  $\mathcal{G}_S^r$  has a particular shape which can be visualized in the rightmost plot of  
 306 Figure 1. It can be decomposed into the disjoint union of  $|S| + 1$  sets of the form  $\mathcal{D}_{S'}^r$   
 307 that are formally defined below.

308 DEFINITION 3.2 (Gap regions, gaps and associated reference points). *Let  $S =$*   
 309  *$\{v_1, \dots, v_n\}$  be a set of  $n$  distinct vectors of the Pareto front dominating a valid refer-*  
 310 *ence point  $r$ . Let  $\sigma$  be the permutation ordering the  $v_i$  by increasing  $F_1$  values:  $v_{\sigma(1),1} <$   
 311  $v_{\sigma(2),1} < \dots < v_{\sigma(n),1}$ .*

- 312 • *For all  $i \in \llbracket 1, n+1 \rrbracket$ , the  $i$ -th gap region of the set  $S$ ,  $\mathcal{G}_{S,i}^r$ , is the set  $\mathcal{D}_{PF_f}^{r_i}$  with*  
 313 *the associated reference points  $r_i$  being  $r_1 = (v_{\sigma(1),1}, r_2)$ ,  $r_{n+1} = (r_1, v_{\sigma(n),2})$  and*  
 314  *$r_i = (v_{\sigma(i),1}, v_{\sigma(i-1),2})$  for all  $i \in \llbracket 2, n \rrbracket$ .*
- 315 • *We refer to  $\mathcal{G}_{S,1}^r$  and  $\mathcal{G}_{S,n+1}^r$  as the left and the right extreme gap region of  $S$ ,*  
 316 *respectively.*

317 The left (resp. right) extreme gap region is empty when the left (resp. right) extreme  
 318 vector of the Pareto front belongs to  $S$ . Non-extreme gap regions are never empty.  
 319 The total gap region is the disjoint union of the gap regions:  $\mathcal{G}_{S_n}^r = \dot{\cup}_{i=1}^{n+1} \mathcal{G}_{S_n,i}^r$ . This  
 320 decomposition of the total gap region, and thus of the optimality gap, is the cornerstone  
 321 of the convergence analysis. Since the area of a gap region  $\mathcal{G}_{S_n,i}^r$  is  $HV_{r_i}^n(PF_f)$ , we can  
 322 write the optimality gap as the sum of  $n+1$  hypervolumes of the Pareto front with respect  
 323 to different reference points.

324 LEMMA 3.3. *At any iteration  $n$ , the optimality gap of a greedy set sequence with*  
 325 *respect to a valid reference point can be decomposed as*

326 (3.1) 
$$HV_r(PF_f) - HV_r(S_n) = \sum_{i=1}^{n+1} HV_{r_i}^n(PF_f) .$$
  
 327

328 *Proof.* The optimality gap at iteration  $n$  is the Lebesgue measure of the total gap  
 329 region  $\mathcal{G}_{\mathcal{S}_n}^r$ , which is the disjoint union of the gap regions  $\mathcal{G}_{\mathcal{S}_n,i}^r = \mathcal{D}_{\text{PF}_f}^{r_i^n}$ . Therefore, the  
 330 optimality gap equals  $\sum_{i=1}^{n+1} \lambda(\mathcal{D}_{\text{PF}_f}^{r_i^n}) = \sum_{i=1}^{n+1} \text{HV}_{r_i^n}(\text{PF}_f)$ .  $\square$

331 Additionally, we can express the hypervolume improvement of any vector to  $\mathcal{S}_n$  as an  
 332 hypervolume. It is a trivial assertion for vectors which do not dominate  $\mathcal{S}_n$ . For other  
 333 vectors, the reference point depends on the gap region to which the vector belongs.

334 **LEMMA 3.4.** *Let  $(\mathcal{S}_n)_{n \in \mathbb{N}^*}$  be a greedy set sequence relative to a valid reference point*  
 335  *$r$ . At any iteration  $n$ , for any  $u$  belonging to the  $i$ -th gap region of  $\mathcal{S}_n$ ,  $\mathcal{G}_{\mathcal{S}_n,i}^r$ , we have*

336 (3.2) 
$$\text{HVI}_r(u, \mathcal{S}_n) = \text{HV}_{r_i^n}(u) .$$

337 *Proof.* The hypervolume improvement of any  $u \in \mathcal{G}_{\mathcal{S}_n,i}^r$  is the Lebesgue-measure of  
 338 the intersection between  $\mathcal{G}_{\mathcal{S}_n,i}^r$  and  $\mathcal{D}_u^r$ . Therefore, it is equal to  $\lambda(\mathcal{D}_u^{r_i^n})$ , that is  $\text{HV}_{r_i^n}(u)$ .  $\square$   
 339 We can now reformulate the recurrence relation defining a greedy sequence at iteration  
 340  $n + 1$ . Indeed, picking a vector maximizing the hypervolume improvement to  $\mathcal{S}_n$  is  
 341 equivalent to pick a vector where the largest value of the maximum of the hypervolumes  
 342 with respect to the  $r_i^n$  is reached.

343 **LEMMA 3.5.** *At any iteration  $n$ , the recurrence relation satisfied by  $v_{n+1}$ , i.e. (2.6),*  
 344 *can be reformulated as*

346 (3.3) 
$$v_{n+1} \in \arg \max_{u \in \text{PF}_f} \max_{i \in [1, n+1]} \text{HV}_{r_i^n}(u) .$$

347 *Proof.* The hypervolume improvement of any vector  $u$  to  $\mathcal{S}_n$  is  $\max_{i \in [1, n+1]} \text{HV}_{r_i^n}(u)$ .  
 348 It is a consequence of [Lemma 3.4](#) and of the fact that the hypervolume with respect to  $r_i^n$   
 349 is null outside the  $i$ -th gap region of  $\mathcal{S}_n$ . Additionally,  $v_{n+1}$  belongs to the Pareto front  
 350 by [Proposition 2.6](#). Thus, (2.6) is equivalent to (3.3).  $\square$   
 351

352 Similarly, we can express the decrease of the optimality gap at iteration  $n + 1$ ,  
 353  $\text{HV}_r(\mathcal{S}_{n+1}) - \text{HV}_r(\mathcal{S}_n)$ , as the maximum of  $n + 1$  hypervolume maximization problems.

354 **LEMMA 3.6.** *Let  $(\mathcal{S}_n)_{n \in \mathbb{N}^*}$  be a greedy set sequence relative to a valid reference point*  
 355  *$r$ . The hypervolume improvement at iteration  $n + 1$  equals*

356 (3.4) 
$$\text{HV}_r(\mathcal{S}_{n+1}) - \text{HV}_r(\mathcal{S}_n) = \max_{u \in \text{PF}_f} \max_{i \in [1, n+1]} \text{HV}_{r_i^n}(u) .$$

357 *Proof.* The hypervolume improvement  $\text{HV}_r(\mathcal{S}_{n+1}) - \text{HV}_r(\mathcal{S}_n)$  is the hypervolume  
 358 improvement of  $v_{n+1}$  to  $\mathcal{S}_n$ . With the same arguments as in the proof of [Lemma 3.5](#), we  
 359 can prove that it equals  $\max_{u \in \text{PF}_f} \max_{i \in [1, n+1]} \text{HV}_{r_i^n}(u)$ .  $\square$   
 360

361 **3.2. Lower bound of the normalized maximum hypervolume for convex**  
 362 **Pareto fronts.** In this section and the next one, we provide bounds on the maximum  
 363 hypervolume achievable by a single feasible vector normalized by the maximum hypervol-  
 364 ume of a feasible set:  $\frac{\max_{u \in \text{PF}_f} \text{HV}_r(u)}{\text{HV}_r(\text{PF}_f)}$ . We refer to this ratio as the *normalized maximum*  
 365 *hypervolume* with respect to  $r$ . Bounds on the normalized maximum hypervolume are ex-  
 366 ploited in [Section 4](#) to provide bounds on the speed of convergence of greedy set sequences  
 367 towards the Pareto front.

368 The hypervolume relative to a reference point  $r$  of a vector  $u = (x, f(x))$  of the Pareto  
 369 front is  $\text{HV}_r(u) = (r_1 - x) \times (r_2 - f(x))$ . From this simple formula, we derive in the next  
 370 proposition necessary conditions for a vector of the Pareto front to be an hypervolume  
 371 maximizer when  $f$  has at least left and right derivatives in  $x^*$ .

372 **PROPOSITION 3.7.** *Let  $x^* \in ]x_{\min}, x_{\max}[$  such that  $u^* := (x^*, f(x^*))$  maximizes the  
 373 hypervolume with respect to a valid reference point  $r$ . If the function  $f$  describing the  
 374 Pareto front admits left and right derivatives in  $x^*$ , respectively  $f'_-(x^*)$  and  $f'_+(x^*)$ , then*

375 (3.5) 
$$-f'_+(x^*) \leq \frac{r_2 - f(x^*)}{r_1 - x^*} \leq -f'_-(x^*) .$$
  
 376

377 *Proof.* We define the function  $\text{HV}_{x,r}(\cdot)$  as  $x \mapsto \text{HV}_r((x, f(x)))$ . If  $x^*$  maximizes  
 378  $\text{HV}_{x,r}(\cdot)$ , then the left and the right derivatives of  $\text{HV}_{x,r}(\cdot)$  are positive and negative,  
 379 respectively. By replacing the left and right derivatives of  $\text{HV}_{x,r}(\cdot)$  by their explicit  
 380 formulas and reorganizing the terms we obtain (3.5).  $\square$

381 Equation (3.5) states that the slope of the diagonal of the rectangle  $\mathcal{D}_{u^*}^r$  is between the  
 382 absolute values of the slopes of the right and the left derivatives of  $f$  at  $x^*$  (see the middle  
 383 plot of Figure 2). To the best of our knowledge, this geometric interpretation is new. It  
 384 becomes simpler when  $f$  is differentiable. Then, the absolute value of the slope of the  
 385 tangent of the front at a non-extreme vector  $u^*$  is equal to the slope of the diagonal of  
 386 the rectangle  $\mathcal{D}_{u^*}^r$  (see the lefthand plot of Figure 2).

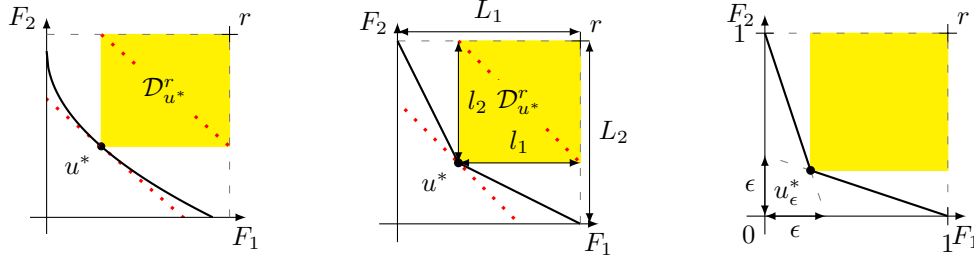


Fig. 2: Left and middle: Two convex Pareto fronts and their respective hypervolume maximizers  $u^*$ , one differentiable (left) and one non-differentiable (middle). The slopes of the two dotted lines, namely  $\text{PF}_g$  and the diagonal of  $\mathcal{D}_{u^*}^r$ , are equal. Right: The Pareto front  $\text{PF}_\epsilon$  and the hypervolume maximizer  $u_\epsilon^*$  for  $\epsilon = 1/3$  and  $r = (1, 1)$ .

387 **COROLLARY 3.8.** *Let  $x^* \in ]x_{\min}, x_{\max}[$  be such that  $u^* := (x^*, f(x^*))$  maximizes the  
 388 hypervolume with respect to a valid reference point  $r$ . If the Pareto front is described by  
 389 a differentiable function  $f$  in  $x^*$ , then  $f'(x^*)$  satisfies*

390 (3.6) 
$$-f'(x^*) = \frac{r_2 - f(x^*)}{r_1 - x^*} .$$
  
 391

392 *Proof.* It is a direct consequence of Proposition 3.7  $\square$

393 A convex function may not be differentiable, but it always has left and right der-  
 394 ivatives. It is also above its left and right tangent lines respectively on the left and

395 on the right of  $x^*$ . Therefore, [Proposition 3.7](#) implies that the affine function  $g : x \mapsto$   
396  $f(x^*) - \frac{r_2 - f(x^*)}{r_1 - x^*} \times (x - x^*)$  is a minorant of  $f$ . It is the key idea of the proof of the  
397 following lower bound on the normalized maximum hypervolume.

398 **PROPOSITION 3.9.** *If the Pareto front is described by a convex function  $f$ , then the*  
399 *following lower bound on the normalized maximum hypervolume with respect to any valid*  
400 *reference point  $r$  holds:*

$$401 \quad (3.7) \quad \frac{\max_{u \in PF_f} HV_r(u)}{HV_r(PF_f)} \geq \frac{1}{2}$$

402  
403 *where the inequality is an equality if and only if the Pareto front is affine and  $r$  dominates*  
404 *the nadir point.*

405 *Proof.* As explained in the above paragraph, the convexity of  $f$  implies that the  
406 affine function  $g : x \mapsto f(x^*) - \frac{r_2 - f(x^*)}{r_1 - x^*} \times (x - x^*)$  is a minorant of  $f$ . Therefore,  
407  $PF_g := \{g(x) : x \in [x_{\min}, x_{\max}]\}$  dominates  $PF_f$ , and thus has a higher hypervolume.  
408 We denote  $L_1 := r_1 - \tilde{x}_{\min, r}$  and  $L_2 := r_2 - f(\tilde{x}_{\max, r})$  the lengths of the rectangle  
409  $\mathcal{R} := [\tilde{x}_{\min, r}, r_1] \times [f(\tilde{x}_{\max, r}), r_2]$ . We denote  $l_1 := r_1 - x^*$  and  $l_2 := r_2 - f(x^*)$  the lengths  
410 of the rectangle  $\mathcal{D}_{u^*}^r$ . The region of  $\mathcal{R}$  which dominates  $PF_g$  is a right-angled triangle.  
411 Additionally, by definition, the slope of its hypotenuse is  $l_2/l_1$ , and thus the lengths of  
412 the other sides are  $L_1 - l_1 + (L_2 - l_2) \times \frac{l_1}{l_2}$  and  $L_2 - l_2 + (L_1 - l_1) \times \frac{l_2}{l_1}$  (see the middle  
413 plot of [Figure 2](#)). Therefore, we have

$$414 \quad HV_r(PF_g) = \lambda(\mathcal{R}) - \lambda(\{u \in \mathbb{R}^2 : u \in \mathcal{R}, u \preceq PF_g\})$$

$$415 \quad = L_1 L_2 - \frac{1}{2} \times (L_1 - l_1 + (L_2 - l_2) \times \frac{l_1}{l_2}) \times (L_2 - l_2 + (L_1 - l_1) \times \frac{l_2}{l_1})$$

$$416 \quad = l_1 l_2 \times \left[ -2 + 2 \times \frac{L_2}{l_2} - \frac{1}{2} \times \left(\frac{L_2}{l_2}\right)^2 + 2 \times \frac{L_1}{l_1} - \frac{1}{2} \times \left(\frac{L_1}{l_1}\right)^2 \right].$$

417  
418 For all  $x$ , we have  $(x - 2)^2 \geq 0$  and thus  $2x - \frac{1}{2}x^2 \leq 2$ . Therefore, we can conclude that  
419  $HV_r(PF_g)$ , and thus  $HV_r(PF_f)$  is smaller than  $2 \times l_1 l_2$ , that is  $2 \times HV_r(u^*)$ . If either  
420  $L_1/l_1 \neq 2$  or  $L_2/l_2 \neq 2$ , the inequality is strict. Thus, when the inequality is an equality,  
421 the center of  $\mathcal{R}$  belongs to the Pareto front. Since  $f$  is convex, it happens only when  $f$  is  
422 affine and the reference point  $r$  dominates the nadir point. Conversely, if both conditions  
423 are met, we know that the optimum is in the middle of the Pareto front and that we have  
424 the equality (see [[3](#), Theorem 5]).  $\square$

425 We just proved that one half is a tight lower bound on the normalized maximum  
426 hypervolume for convex Pareto fronts. However, except for the trivial upper bound 1,  
427 there is no upper bound valid for every convex Pareto front, even when  $r$  dominates the  
428 nadir point. Here is a simple example which illustrates this. Let consider the convex  
429 Pareto front  $PF_\epsilon := \{\max(1 - \frac{x}{\epsilon}, \epsilon - \epsilon \times x) : x \in [0, 1]\}$  represented in the righthand  
430 plot of [Figure 2](#) and the reference point  $r = (1, 1)$ . When  $\epsilon \leq 1$ ,  $PF_\epsilon$  is convex and  
431 [\(3.6\)](#) implies that  $u_\epsilon^* = (\epsilon \times (1 - \epsilon), \epsilon \times (1 - \epsilon))$  is the unique hypervolume minimizer.  
432 Thus, the normalized maximum hypervolume of  $PF_\epsilon$  for this reference point is equal to  
433  $\frac{(1 - \epsilon + \epsilon^2)^2}{1 - \epsilon \times (1 - \epsilon)^2 + (\epsilon - \epsilon^2)^2}$  and converges to 1 when  $\epsilon$  goes to 0.

434 **3.3. Lower and upper bounds of the normalized maximum hypervolume**  
435 **for bilipschitz Pareto fronts.** In this section, we examine lower and upper bounds on  
436 the normalized maximum hypervolume in the case of bilipschitz Pareto fronts.

437 We consider two affine fronts with the same left extreme vector as  $\text{PF}_f$  and slopes  
 438  $-L_{\min}$  and  $-L_{\max}$ , see the lefthand plot of [Figure 3](#). We call them  $\text{PF}_{\min}$  and  $\text{PF}_{\max}$ ,  
 439 respectively. Formally:

440 (3.8)  $\text{PF}_{\max} := \{(x, f_{\max}(x)) : x \in [x_{\min}, x_{\max}]\}$  and

441 (3.9)  $\text{PF}_{\min} := \{(x, f_{\min}(x)) : x \in [x_{\min}, x_{\max}]\}$

443 with  $f_{\min}(x) = f(x_{\min}) - (x - x_{\min}) \times L_{\min}$  and  $f_{\max}(x) = f(x_{\min}) - (x - x_{\min}) \times$   
 444  $L_{\max}$ . For a  $(L_{\min}, L_{\max})$ -bilipschitz function  $f$ ,  $f_{\min}(x) \leq f(x) \leq f_{\max}(x)$  for all  $x \in$   
 445  $[x_{\min}, x_{\max}]$ , and thus the Pareto front is dominated by  $\text{PF}_{\max}$  and dominates  $\text{PF}_{\min}$ .  
 These two affine fronts provide bounds on both the hypervolume of the Pareto front

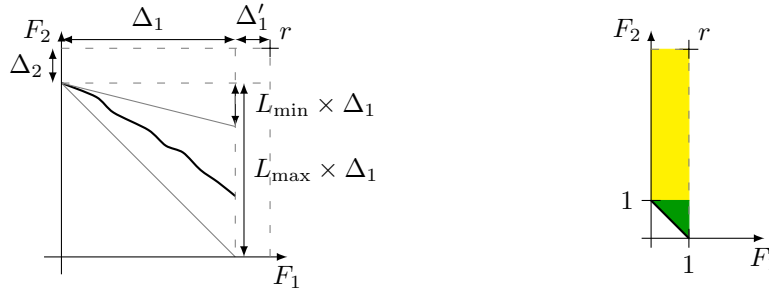


Fig. 3: Left : The Pareto front  $\text{PF}_f$  surrounded by  $\text{PF}_{\max}$  (below) and  $\text{PF}_{\min}$  (above).  
 Right : An illustration that  $\text{HV}_r(u_{\min}) - \text{HV}_r(\text{PF}_f)$  (■) becomes negligible compared to  
 $\text{HV}_r(u_{\min})$  (■) for  $r_1 = 1$  and  $r_2 \rightarrow \infty$ .

446 and the largest hypervolume of a vector on the Pareto front. They are key to prove  
 447 the following lower bound on the normalized maximum hypervolume of a  $(L_{\min}, L_{\max})$ -  
 448 bilipschitz Pareto front.  
 449

450 **PROPOSITION 3.10.** *If the Pareto front is described by a  $(L_{\min}, L_{\max})$ -bilipschitz func-*  
 451 *tion  $f$ , then for any valid reference point  $r$ , we have*

452 (3.10) 
$$\frac{\max_{u \in \text{PF}_f} \text{HV}_r(u)}{\text{HV}_r(\text{PF}_f)} \geq \frac{1}{2} \times \frac{L_{\min}}{L_{\max}} .$$

453

454 *Proof.* The fronts  $\text{PF}_{\max}$  and  $\text{PF}_{\min}$  are defined respectively in (3.8) and (3.9). We  
 455 note  $\Delta_1 := \tilde{x}_{\max, r} - \tilde{x}_{\min, r}$ ,  $\Delta'_1 := r_1 - \tilde{x}_{\max, r}$ ,  $\Delta_2 := r_2 - f(\tilde{x}_{\min, r})$  and  $V := \Delta_2 \times (r_1 -$   
 456  $\tilde{x}_{\min, r})$ , see the lefthand plot of [Figure 3](#). Since the front  $\text{PF}_{\max}$  dominates the Pareto  
 457 front, the hypervolume of  $\text{PF}_f$  is smaller than the hypervolume of  $\text{PF}_{\max}$ ,  $V + L_{\max} \times$   
 458  $\Delta_1 \times \Delta'_1 + \frac{1}{2} \times L_{\max} \times \Delta_1^2$ . Additionally, since each vector of  $\text{PF}_{\min}$  is dominated by a  
 459 vector of  $\text{PF}_f$ , the maximum hypervolume of a vector of  $\text{PF}_f$  is larger than the maximum  
 460 hypervolume of a vector of  $\text{PF}_{\min}$ . The front  $\text{PF}_{\min}$  being an affine and therefore convex  
 461 front, we know by [Proposition 3.9](#) that the maximum hypervolume of a vector of  $\text{PF}_{\min}$  is  
 462 larger than half of  $\text{HV}_r(\text{PF}_{\min})$ , which is equal to  $\frac{1}{2} \times (V + L_{\min} \times \Delta_1 \times \Delta'_1 + \frac{1}{2} \times L_{\min} \times \Delta_1^2)$ .  
 463 To summarize, the maximum hypervolume of a vector of  $\text{PF}_f$  is larger than  $\frac{1}{2} \times (V + L_{\min} \times$   
 464  $\Delta_1 \times \Delta'_1 + \frac{1}{2} \times L_{\min} \times \Delta_1^2)$ . Combining the upper bound on the hypervolume of  $\text{PF}_f$  and the  
 465 lower bound on the maximum hypervolume of a vector of  $\text{PF}_f$ , the normalized maximum

466 hypervolume is larger than  $\frac{\frac{1}{2} \times (V + L_{\min} \times \Delta_1 \times \Delta'_1 + \frac{1}{2} \times L_{\min} \times \Delta_1^2)}{V + L_{\max} \times \Delta_1 \times \Delta'_1 + \frac{1}{2} \times L_{\max} \times \Delta_1^2}$ . This quantity is itself larger  
 467 than  $\frac{1}{2} \times \frac{L_{\min} \times \Delta_1 \times \Delta'_1 + \frac{1}{2} \times L_{\min} \times \Delta_1^2}{L_{\max} \times \Delta_1 \times \Delta'_1 + \frac{1}{2} \times L_{\max} \times \Delta_1^2}$ . As  $V \geq 0$  and  $0 < \frac{L_{\min} \times \Delta_1 \times \Delta'_1 + \frac{1}{2} \times L_{\min} \times \Delta_1^2}{L_{\max} \times \Delta_1 \times \Delta'_1 + \frac{1}{2} \times L_{\max} \times \Delta_1^2} < 1$ , we  
 468 conclude that the normalized maximum hypervolume is larger than  $\frac{1}{2} \times \frac{L_{\min}}{L_{\max}}$ .  $\square$

469 We cannot guarantee any upper bound strictly inferior to 1 on the normalized maximum  
 470 hypervolume without adding an assumption on the reference point. Indeed, for a given  
 471 bounded Pareto front, it is easy to show that the normalized maximum hypervolume goes  
 472 to 1 for  $r_1 = x_{\max}$  and  $r_2 \rightarrow \infty$  (see the righthand plot of [Figure 3](#)). However, if  $f$  is  
 473  $(L_{\min}, L_{\max})$ -bilipschitz and  $r$  dominates the nadir point, we can prove that the normalized  
 474 maximum hypervolume is larger than  $\frac{1}{2} \times \frac{L_{\max}}{L_{\min}}$ . The proof relies on the fact that if the  
 475 reference point  $r$  dominates the nadir point, the vector of an affine front with the largest  
 476 hypervolume relative to  $r$  is its middle (see [3, Theorem 5]), whose hypervolume is half  
 477 of the hypervolume of the entire Pareto front.

478 **PROPOSITION 3.11.** *If the Pareto front is described by a  $(L_{\min}, L_{\max})$ -bilipschitz func-*  
 479 *tion  $f$  and the reference point  $r$  is valid and dominates the nadir point, the following*  
 480 *upper-bound on the normalized maximum hypervolume with respect to  $r$  holds*

$$481 \quad (3.11) \quad \frac{\max_{u \in PF_f} HV_r(u)}{HV_r(PF_f)} \leq \frac{1}{2} \times \frac{L_{\max}}{L_{\min}} .$$

483 *Proof.* We use the same notations as in the proof of [Proposition 3.10](#). Since  $r$  domi-  
 484 nates the nadir point, both  $\Delta'_1$ ,  $\Delta_2$  and  $V$  equal 0, and thus the hypervolumes of  $PF_{\max}$   
 485 and  $PF_{\min}$  equal  $\frac{1}{2} \times L_{\max} \times \Delta_1^2$  and  $\frac{1}{2} \times L_{\min} \times \Delta_1^2$ , respectively. The domination of  
 486  $PF_{\min}$  by  $PF_f$  implies that the hypervolume of the Pareto front is below  $\frac{1}{2} \times L_{\min} \times \Delta_1^2$ .  
 487 Since  $PF_{\max}$  is an affine front whose extremes dominate  $r$ , its middle is the unique hyper-  
 488 volume maximizer (see [2, Theorem 5]) with an hypervolume equal to  $\frac{1}{4} \times L_{\max} \times \Delta_1$ . The  
 489 domination of  $PF_f$  by  $PF_{\max}$  implies that the maximum hypervolume of a vector of  $PF_f$   
 490 is larger than  $\frac{1}{4} \times L_{\max} \times \Delta_1^2$ . Gathering the lower bound on  $HV_r(PF_f)$  and the upper  
 491 bound on the maximum hypervolume of a vector of  $PF_f$ , we retrieve (3.11).  $\square$

492 This upper bound is only relevant for  $L_{\max}/L_{\min} < 2$  and is the tightest for  $L_{\max} = L_{\min}$ ,  
 493 where it achieves the value 1/2. In this paper, we use this upper bound for  $L_{\max}/L_{\min}$   
 494 close to 1 to analyze the asymptotic convergence behavior of HV-ISOOMOO.

495 **4. Convergence of HV-ISOOMOO coupled with perfect singleobjective**  
 496 **optimization.** We prove in this section various convergence results for HV-ISOOMOO  
 497 algorithms coupled with perfect singleobjective optimization. We first prove that when  
 498 the Pareto front is either convex or bilipschitz, these algorithms converge to the entire  
 499 Pareto front. We transform the bounds on the normalized maximum hypervolume proven  
 500 in [Section 3](#) into lower bounds on the speed of convergence. Then, we analyze the asymp-  
 501 totic convergence behavior when the Pareto front is bilipschitz with a Hölder continuous  
 502 derivative.

503 To analyze the decrease of the optimality gap with  $n$ , we track in which gap regions  
 504 the vectors of the greedy sequence are inserted over multiple iterations. Naturally, a gap  
 505 region of  $\mathcal{S}_n$  persists in being a gap region as long as no greedy vector is added in this  
 506 specific gap region. The greedy vector  $v_{n+1}$  is said to *fill* the gap region of  $\mathcal{S}_n$  to which it  
 507 belongs. At iteration  $n + 1$ , this gap region disappears, replaced by two gap regions that  
 508 we call its *children*. More generally, we say that a gap region is a *descendant* of another  
 509 gap region when it is a proper subset of this gap region.

510 **4.1. Convergence of HV-ISOOMOO with guaranteed speed of conver-**  
511 **gence.** We prove some upper bounds on the relation between the optimality gap at  
512 iteration  $2n + 1$  and at iteration  $n$ . These bounds translate into lower bounds on the  
513 speed of convergence of HV-ISOOMOO under [Assumption 2.3](#) of perfect singleobjective  
514 optimization. The proof relies on inequalities of the form

$$515 \quad (4.1) \quad \max_{u \in \text{PF}_f} \text{HV}_{r'}(u) \geq C \times \text{HV}_{r'}(\text{PF}_f)$$

517 stated in [Propositions 3.9](#) and [3.10](#) and equations regarding optimality gaps, areas of gap  
518 regions and hypervolume improvement presented in [Subsection 3.1](#). A consequence of  
519 (4.1) being true for any reference point  $r'$  is that the optimality gap at iteration  $2n + 1$   
520 is at most  $(1 - C)$  times the optimality gap at iteration  $n$ .

521 We sketch the proof idea in the simple case where each of the  $v_k$  ( $k \in \llbracket n + 1, 2n + 1 \rrbracket$ )  
522 is inserted in a distinct gap region of  $\mathcal{S}_n$ , see the lefthand plot of [Figure 4](#). Inserting  $v_k$  in  
523 a gap region leads to an hypervolume improvement larger than  $C$  times the area of  $k$   
524 gap region by (4.1). Thus, the hypervolume improvement from iteration  $n$  to  $2n + 1$   
525 is larger than  $C$  times the area of the union of all gap regions of  $\mathcal{S}_n$ , namely the optimality  
gap at iteration  $n$ . A detailed proof is presented after the theorem statement.

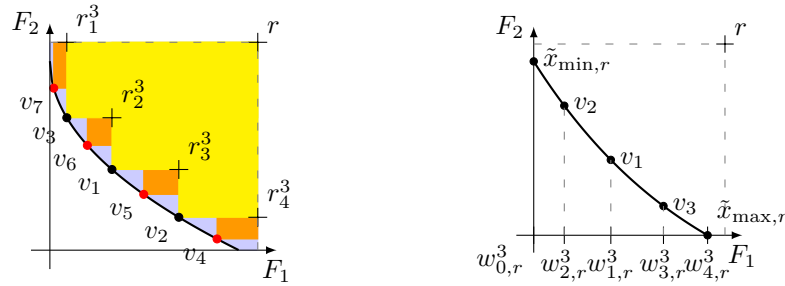


Fig. 4: Left: A Pareto front where each of the gap regions of  $\mathcal{S}_3$  is filled by one of the greedy vectors  $v_k$  for  $k \in \llbracket 4, 7 \rrbracket$ . It is described by  $f(x) = 1 - \sqrt{x}$  for  $x \in [0, 1]$ . We represent the region  $\mathcal{D}_{\mathcal{S}_3}^r$  (yellow), the gap regions of  $\mathcal{S}_3$  (blue) and the regions corresponding to  $\text{HVI}_r(v_k, \mathcal{S}_{k-1})$  for  $k \in \llbracket 4, 7 \rrbracket$  (orange). Right: The ordered greedy set  $F_1$ -values  $w_{i,r}^n$  corresponding to the greedy set  $\mathcal{S}_3$ . The Pareto front is described by  $f(x) = \frac{e}{e-1} \times e^{-x} + 1 - \frac{e}{e-1}$  for  $x \in [0, 1]$ .

526

527 **PROPOSITION 4.1.** Consider a biobjective optimization problem with a Pareto front  
528 described by a function  $f$ . Any greedy set sequence  $(\mathcal{S}_n)_{n \in \mathbb{N}^*}$  relative to a valid reference  
529 point  $r$  satisfies for all  $n$

$$530 \quad (4.2) \quad \frac{\text{HV}_r(\text{PF}_f) - \text{HV}_r(\mathcal{S}_{2n+1})}{\text{HV}_r(\text{PF}_f) - \text{HV}_r(\mathcal{S}_n)} \leq 1 - \frac{1}{2} \times \frac{L_{\min}}{L_{\max}} \text{ if } f \text{ is } (L_{\min}, L_{\max})\text{-bilipschitz and}$$

$$531 \quad (4.3) \quad \frac{\text{HV}_r(\text{PF}_f) - \text{HV}_r(\mathcal{S}_{2n+1})}{\text{HV}_r(\text{PF}_f) - \text{HV}_r(\mathcal{S}_n)} \leq \frac{1}{2} \text{ if } f \text{ is convex.}$$

533 *Proof.* Fix  $n \geq 1$ . We note  $\sigma$  a permutation of  $\llbracket 1, n + 1 \rrbracket$  such that  $n + \sigma(i)$  is the  
534 index of the first greedy vector  $v_k$  inserted in  $\mathcal{G}_{\mathcal{S}_n, i}^r$  when possible. With this choice of  $\sigma$ ,



535 the  $i$ -th gap region of  $\mathcal{S}_n$  is a gap region of  $\mathcal{S}_{n+\sigma(i)-1}$ . As a consequence, the hypervolume  
536 improvement to  $\mathcal{S}_{n+\sigma(i)-1}$  of any vector  $u$  belonging to the  $i$ -th gap region of  $\mathcal{S}_n$ ,  $\mathcal{G}_{\mathcal{S}_n,i}^r$ ,  
537 is equal to  $\text{HV}_{r_i^n}(u)$  by [Lemma 3.5](#). The hypervolume improvement of the greedy vector  
538  $v_{n+\sigma(i)}$  to  $\mathcal{S}_{n+\sigma(i)-1}$  being maximal, it is in particular larger than the one of any vector  
539 of  $\mathcal{G}_{\mathcal{S}_n,i}^r$  and thus than  $\frac{1}{2} \times \frac{L_{\min}}{L_{\max}} \times \text{HV}_{r_i^n}(\text{PF}_f)$  by [Proposition 3.10](#). In other words, the  
540 hypervolume improvement at any iteration  $n + \sigma(i)$  is larger than  $\frac{1}{2} \times \frac{L_{\min}}{L_{\max}} \times \text{HV}_{r_i^n}(\text{PF}_f)$ .  
541 By adding these inequations for all  $i \in \llbracket 1, n + 1 \rrbracket$ , we deduce that the hypervolume  
542 improvement from iteration  $n$  to  $2n + 1$  is larger than  $\frac{1}{2} \times \frac{L_{\min}}{L_{\max}} \times \sum_{i=1}^{n+1} \text{HV}_{r_i^n}(\text{PF}_f)$ .  
543 Since the sum of the  $\text{HV}_{r_i^n}(\text{PF}_f)$  is the optimality gap at iteration  $n$ , we have [\(4.2\)](#). If  
544  $f$  is convex instead of bilipschitz, we use [Proposition 3.9](#) instead of [Proposition 3.10](#) and  
545 obtain [\(4.3\)](#).  $\square$

546 Since the optimality gaps form a decreasing sequence, such lower bounds on the relation  
547 between the optimality gaps at iteration  $2n+1$  and at iteration  $n$  imply that the optimality  
548 gap associated to a greedy set sequence converges asymptotically to 0. Equivalently, HV-  
549 ISOOMOO algorithms coupled with perfect singleobjective optimization converge to the  
550 entire Pareto front as stated formally below.

551 **THEOREM 4.2.** *Consider a biobjective optimization problem with a Pareto front de-*  
552 *scribed by a bilipschitz or convex function  $f$ .*

553 *The hypervolume of a greedy set sequence relative to a valid reference point  $r$  converges*  
554 *to the hypervolume of the entire Pareto front, i.e.  $\text{HV}_r(\mathcal{S}_n) \xrightarrow[n \rightarrow \infty]{} \text{HV}_r(\text{PF}_f)$ .*

555 *Equivalently, for such Pareto fronts and under [Assumption 2.3](#) of perfect singleobjec-*  
556 *tive optimization, HV-ISOOMOO algorithms relative to a valid reference point  $r$  converge*  
557 *to the Pareto front in the sense of [Definition 2.2](#).*

558 From the lower bounds on the relation between the optimality gaps at iteration  $2n + 1$   
559 and at iteration  $n$ , we deduce the following upper bounds on the normalized optimality  
560 gap at any iteration.

561 **COROLLARY 4.3.** *Consider a biobjective optimization problem with a Pareto front*  
562 *described by a  $(L_{\min}, L_{\max})$ -bilipschitz function. A greedy set sequence  $(\mathcal{S}_n)_{n \in \mathbb{N}^*}$  relative*  
563 *to a valid reference point  $r$  satisfies for all  $n$*

$$564 \quad (4.4) \quad \frac{\text{HV}_r(\text{PF}_f) - \text{HV}_r(\mathcal{S}_n)}{\text{HV}_r(\text{PF}_f)} \leq \left(1 - \frac{1}{2} \times \frac{L_{\min}}{L_{\max}}\right)^{\lfloor \log_2(n+1) \rfloor} \leq (2n+2)^{\log_2(1 - \frac{1}{2} \times \frac{L_{\min}}{L_{\max}})} .$$

566 *If the function  $f$  is convex, then any greedy set sequence relative to a valid reference point*  
567  *$r$  satisfies for all  $n$*

$$568 \quad (4.5) \quad \frac{\text{HV}_r(\text{PF}_f) - \text{HV}_r(\mathcal{S}_n)}{\text{HV}_r(\text{PF}_f)} \leq \left(\frac{1}{2}\right)^{\lfloor \log_2(n+1) \rfloor} \leq \frac{1}{2n+2} .$$

570 *Hence, for such reference points and under [Assumption 2.3](#) of perfect singleobjective op-*  
571 *timization, HV-ISOOMOO algorithms relative to  $r$  satisfy [\(4.4\)](#) if  $f$  is  $(L_{\min}, L_{\max})$ -*  
572 *bilipschitz and [\(4.5\)](#) if  $f$  is convex where  $\mathcal{S}_n$  is replaced by  $\mathcal{I}_n$ , the final incumbents Pareto*  
573 *front approximation at iteration  $n$ .*

574 *Proof.* The  $k$ -th term of the sequence defined by  $u_0 = 1$  and  $u_{n+1} = 2 \times u_n + 1$  for  
575 all  $n \geq 1$  is  $2^k - 1$ . Thus, [\(4.2\)](#) and [\(4.3\)](#) imply that when  $f$  is  $(L_{\min}, L_{\max})$ -bilipschitz or  
576 convex, the normalized optimality gap at iteration  $2^k - 1$  is inferior to  $(1 - C)^k$  with  $C$

577 equal to  $\frac{1}{2} \times \frac{L_{\min}}{L_{\max}}$  and  $\frac{1}{2}$ , respectively. Since the hypervolume of the greedy set increases  
 578 with  $n$ , and thus the optimality gap decreases with  $n$ , we deduce the first inequalities in  
 579 (4.4) and (4.5) via the change of variable  $k = \lfloor \log_2(n+1) \rfloor$ .

580 Additionally, for every  $n$ ,  $\lfloor \log_2(n+1) \rfloor$  is smaller than  $\log_2(n+1)+1$ , that is  $\log_2(2n+2)$ .  
 581 For every  $C$ ,  $\log_2(2n+2)$  equals  $\log_C(2n+2) \times \log_2(C)$ , and thus  $C^{\log_2(2n+2)}$  equals  
 582  $(2n+2)^{\log_2(C)}$ . Therefore, we can infer that  $(2n+2)^{\log_2(C)}$  is an upper bound of the  
 583 normalized optimality gap with  $C = 1 - \frac{1}{2} \times \frac{L_{\min}}{L_{\max}}$  and  $C = \frac{1}{2}$  when  $f$  is  $(L_{\min}, L_{\max})$ -  
 584 bilipschitz and convex, respectively.  $\square$

585 We focus here on the relation between the optimality gap at iteration  $n$  and at  
 586 iteration  $2n+1$ . We could similarly examine the relation between the optimality gap at  
 587 iteration  $n$  and at any later iteration. For example, we could prove that if  $f$  is  $(L_{\min}, L_{\max})$ -  
 588 bilipschitz, then for all  $n$ , for all  $k \leq n+1$ ,  $\frac{\text{HV}_r(\text{PF}_f) - \text{HV}_r(\mathcal{S}_{n+k})}{\text{HV}_r(\text{PF}_f) - \text{HV}_r(\mathcal{S}_n)}$  is smaller than  $1 - \frac{1}{2} \times$   
 589  $\frac{L_{\min}}{L_{\max}} \times \frac{k}{n+1}$ .

590 Consider the  $k$  gap regions of  $\mathcal{S}_n$  with the largest areas. The hypervolume improve-  
 591 ment from iteration  $n$  to  $n+k$  is at least  $\frac{1}{2} \times \frac{L_{\min}}{L_{\max}}$  times the area of the union of these  
 592 gap regions, which is at least  $\frac{k}{n+1}$  times the optimality gap at iteration  $n$ .

#### 593 4.2. Asymptotical behavior of the convergence of $\text{HV}_r(\mathcal{S}_n)$ to $\text{HV}_r(\text{PF}_f)$ .

594 In this section, we analyze the asymptotic convergence behavior for a Pareto front de-  
 595 scribed by a bilipschitz function with a Hölder continuous derivative. We prove that, in  
 596 this case, doubling the number of vectors in the greedy set divides the optimality gap by a  
 597 factor which converges asymptotically to two as stated in [Theorem 4.10](#). This asymptotic  
 598 limit corresponds to the case of affine Pareto fronts with a reference point dominating  
 599 the nadir point. For such Pareto fronts and reference points, the optimality gap is always  
 600 halved when the number of vectors in the greedy set goes from  $n$  to  $2n+1$ , see [Figure 5](#).

601 First, we study the properties of the part of the Pareto front corresponding to a  
 602 specific gap region of  $\mathcal{S}_n$ . For all  $n$ , let note  $\sigma_n$  the permutation of  $\llbracket 1, n \rrbracket$  which orders  
 603 the vectors of  $\mathcal{S}_n$  by increasing  $F_1$ -values and the so-called *ordered greedy set  $F_1$ -values*:

$$604 \quad (4.6) \quad w_{i,r}^n := v_{\sigma_n(i),1} \text{ for } i \in \llbracket 1, n \rrbracket, w_{0,r}^n := \tilde{x}_{\min,r} \text{ and } w_{n+1,r}^n := \tilde{x}_{\max,r} .$$

606 Naturally, we have  $w_{0,r}^n \leq w_{1,r}^n \leq \dots \leq w_{n+1,r}^n$ , and the intervals  $[w_{i-1,r}^n, w_{i,r}^n[$  for  $i \in$   
 607  $\llbracket 1, n+1 \rrbracket$  form a partition of  $[\tilde{x}_{\min,r}, \tilde{x}_{\max,r}[$ , see the righthand plot of [Figure 4](#). The  
 608 interval  $[w_{i-1,r}^n, w_{i,r}^n[$  corresponds to the part of the Pareto front belonging to the  $i$ -th  
 609 gap region of  $\mathcal{S}_n$ . When the Pareto front is bilipschitz, the lengths of these intervals  
 610 converge asymptotically to 0 as stated in the next lemma. It is a direct consequence of  
 611 the convergence of  $\text{HV}_r(\mathcal{S}_n)$  to  $\text{HV}_r(\text{PF}_f)$  stated in [Theorem 4.2](#).

612 **LEMMA 4.4.** *If the Pareto front is described by a bilipschitz function  $f$  and the greedy*  
 613 *set sequence is associated to a valid reference point  $r$ , then the ordered greedy set  $F_1$ -values*  
 614 *satisfy  $\max_{i \in \llbracket 1, n+1 \rrbracket} w_{i,r}^n - w_{i-1,r}^n \xrightarrow[n \rightarrow \infty]{} 0$  with the  $w_{i,r}^n$  defined in (4.6).*

615 *Proof.* Let  $L_{\min}$  and  $L_{\max}$  be constants such that  $f$  is  $(L_{\min}, L_{\max})$ -bilipschitz. The  
 616 area of the  $i$ -th gap region of  $\mathcal{S}_n$  is  $\int_{w_{i-1,r}^n}^{w_{i,r}^n} (f(x) - f(w_{i,r}^n)) dx$ . This is larger than  
 617  $\int_{w_{i-1,r}^n}^{w_{i,r}^n} L_{\min} \times (w_{i,r}^n - x) dx$ , which equals  $\frac{1}{2} \times L_{\min} \times (w_{i,r}^n - w_{i-1,r}^n)^2$ . Since the area  
 618 of any gap region of  $\mathcal{S}_n$  is inferior to the optimality gap at iteration  $n$ , this implies that  
 619 the difference  $w_{i,r}^n - w_{i-1,r}^n$  is inferior to  $\sqrt{2} \times (\text{HV}_r(\text{PF}_f) - \text{HV}_r(\mathcal{S}_n))$  for all  $n$ , for all

620  $i \in \llbracket 1, n+1 \rrbracket$ . Therefore, the convergence of  $\text{HV}_r(\mathcal{S}_n)$  to  $\text{HV}_r(\text{PF}_f)$  stated in [Theorem 4.2](#)  
 621 implies that the maximum over  $i$  of  $w_{i,r}^n - w_{i-1,r}^n$  converges to 0.  $\square$

622 We prove in the next lemma that if the Pareto front is described by a bilipschitz  
 623 function  $f$  with a Hölder continuous derivative, then the the part of the Pareto front  
 624 belonging to a specific gap region of  $\mathcal{S}_n$  is bilipschitz for some constants whose ratio con-  
 625 verges asymptotically to 1. Affine functions being the only functions to be  $(L_{\min}, L_{\max})$ -  
 626 bilipschitz with  $L_{\min}/L_{\max} = 1$ , it supports the interpretation that the convergence of  
 627 a greedy set sequence for such Pareto fronts and for affine Pareto fronts share some  
 628 asymptotic similarities.

629 When  $f$  is bilipschitz, its restriction to the part of the Pareto front dominating  $r_i^n$ ,  
 630 that is  $[w_{i-1,r}^n, w_{i,r}^n]$ , is  $(L_{\min}^{i,n}, L_{\max}^{i,n})$ -bilipschitz with

$$631 \quad (4.7) \quad \begin{aligned} L_{\min}^{i,n} &:= \inf \left\{ \left| \frac{f(x) - f(y)}{x - y} \right|, x, y \in [w_{i-1,r}^n, w_{i,r}^n], x \neq y \right\} \text{ and} \\ L_{\max}^{i,n} &:= \sup \left\{ \left| \frac{f(x) - f(y)}{x - y} \right|, x, y \in [w_{i-1,r}^n, w_{i,r}^n], x \neq y \right\}. \end{aligned}$$

632 At iteration  $n$ , the ratio between  $L_{\max}^{i,n}$  and  $L_{\min}^{i,n}$ , the bilipschitz constants on the  $i$ -th gap  
 633 region of  $\mathcal{S}_n$ , is by definition smaller than

$$634 \quad (4.8) \quad q_n := \max \left\{ \frac{L_{\max}^{i,n}}{L_{\min}^{i,n}}, i \in \llbracket 1, n+1 \rrbracket : [w_{i-1,r}^n, w_{i,r}^n] \neq \emptyset \right\}.$$

635

636 The proof of the convergence of  $q_n$  to 1 relies on the fact that a derivable function can  
 637 be approximated locally by an affine function. The quality of this approximation is  
 638 guaranteed by the Hölder continuity of the derivative.

639 **LEMMA 4.5.** *We consider a greedy set sequence  $(\mathcal{S}_n)_{n \in \mathbb{N}^*}$  relative to a valid reference*  
 640 *point  $r$ . If the Pareto front is described by a bilipschitz function with a Hölder continuous*  
 641 *derivative, then  $q_n$  defined in (4.8) converges asymptotically to 1.*

642 *Proof.* We take  $\alpha$  such that  $f'$  is Hölder continuous with exponent  $\alpha$ , i.e  $f$  is  $\mathcal{C}^{1,\alpha}$ ,  
 643 and  $L_{\min}, L_{\max} > 0$  such that the function  $f$  describing the Pareto front is  $(L_{\min}, L_{\max})$ -  
 644 bilipschitz. We recall that  $f$  is decreasing, and thus for all  $x < y$ , we have  $f(x) - f(y) \geq 0$ .  
 645 Since  $f$  is  $\mathcal{C}^{1,\alpha}$  and therefore  $\mathcal{C}^1$ , the Taylor formula with Lagrange remainder states that  
 646 for all  $x < y$ , there exists  $\xi \in [x, y]$  such that  $f(y) = f(x) + (y-x) \times f'(\xi)$ . Since  $f$  is  $\mathcal{C}^{1,\alpha}$ ,  
 647 this implies that for all  $x < y$ ,  $|f(y) - f(x) - (y-x) \times f'(x)| \leq (y-x)^{1+\alpha} \times [f']_{\mathcal{C}^\alpha}$ . Thus,  
 648  $\frac{f(y)-f(x)}{x-y}$  is smaller than  $-f'(x) + [f']_{\mathcal{C}^\alpha} \times (y-x)^\alpha$ . We now restrict ourselves to  $x$  and  
 649  $y$  belonging to the non-empty interval  $[w_{i-1,r}^n, w_{i,r}^n]$ . Our goal is to find an upper bound  
 650 depending on  $i$  but not on either  $x$  or  $y$ . Since  $f$  is  $\mathcal{C}^{1,\alpha}$ , the difference between  $-f'(x)$   
 651 and  $-f'(w_{i-1,r}^n)$  is smaller than  $[f']_{\mathcal{C}^\alpha} \times (x - w_{i-1,r}^n)^\alpha$ , and thus  $[f']_{\mathcal{C}^\alpha} \times (w_{i,r}^n - w_{i-1,r}^n)^\alpha$   
 652 . Additionally, the difference between  $x$  and  $y$  is smaller than  $w_{i,r}^n - w_{i-1,r}^n$ . We conclude  
 653 that for  $x, y \in [w_{i-1,r}^n, w_{i,r}^n]$ ,  $\frac{f(y)-f(x)}{x-y}$  is smaller than  $-f'(w_{i-1,r}^n) + 2[f']_{\mathcal{C}^\alpha} \times (w_{i,r}^n -$   
 654  $w_{i-1,r}^n)^\alpha$ , and thus so is  $L_{\max}^{i,n}$  defined in (4.7).

655 Following the same approach, we can also infer that  $L_{\max}^{i,n}$  defined in (4.7) is greater  
 656 than the symmetric quantity  $-f'(w_{i-1,r}^n) - 2[f']_{\mathcal{C}^\alpha} \times (w_{i,r}^n - w_{i-1,r}^n)^\alpha$ . The quantity  
 657  $-f'(w_{i-1,r}^n)$  is greater than  $L_{\min}$  and  $(w_{i,r}^n - w_{i-1,r}^n)^\alpha$  is smaller than  $\max_{i \in \llbracket 1, n+1 \rrbracket} (w_{i,r}^n -$

658  $w_{i-1,r}^n)^\alpha$ . As a consequence,  $q_n$  is smaller than  $\frac{L_{\min}+2[f']_{C^\alpha} \times \max_{i \in [1, n+1]} (w_{i,r}^n - w_{i-1,r}^n)^\alpha}{L_{\min}-2[f']_{C^\alpha} \times \max_{i \in [1, n+1]} (w_{i,r}^n - w_{i-1,r}^n)^\alpha}$ . By  
659 **Lemma 4.4**,  $\max_{i \in [1, n+1]} w_{i,r}^n - w_{i-1,r}^n$  converges to 0 and thus, this upper bound on  $q_n$   
660 converges to 1. Since  $q_n$  is always larger than 1, it converges to 1.  $\square$

661 A consequence of the previous lemma is that the bounds on the hypervolume improvement  
662 of  $v_{n+1}$  to  $\mathcal{S}_n$  normalized by the area of the gap region filled by  $v_{n+1}$  that we can infer  
663 from **Propositions 3.10** and **3.11** converge asymptotically to 1/2, see (4.9). Similarly,  
664 the bounds on the normalized area of the child of a gap region that we can infer from  
665 **Lemma A.2** converge to 1/4, see (4.10). These asymptotic values correspond to the case  
666 of an affine Pareto front with a reference point dominating the nadir point, see **Figure 5**.

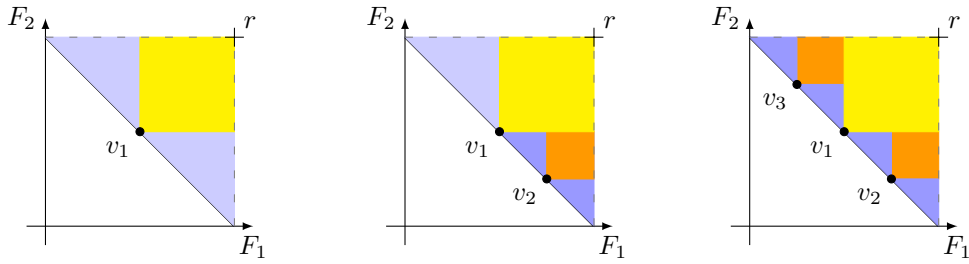


Fig. 5: The three greedy sets  $\mathcal{S}_1$  (left),  $\mathcal{S}_2$  (middle) and  $\mathcal{S}_3$  (right) and their gap regions for an affine Pareto front with a reference point  $r$  dominating the nadir point. The area of any of the gap regions of  $\mathcal{S}_1$  are half of  $\text{HV}_r(\text{PF}_f)$  (left). The area of any of the new gap regions of  $\mathcal{S}_2$  is a quarter of the area of their parents (middle). The optimality gap of  $\mathcal{S}_3$  (right) is half of the optimality gap of  $\mathcal{S}_1$ .

667

668 **LEMMA 4.6.** *We consider a greedy set sequence  $(\mathcal{S}_n)_{n \in \mathbb{N}^*}$  relative to a valid reference*  
669 *point  $r$ . If the Pareto front is described by a bilipschitz function  $f$  with a Hölder continuous*  
670 *derivative, then for all  $\epsilon > 0$ , for  $n$  large enough, for every non-empty gap region  $\mathcal{G}_{\mathcal{S}_n, i}^r$*   
671 *and every child  $\mathcal{G}_{\mathcal{S}_n, j}^r$  of  $\mathcal{G}_{\mathcal{S}_n, i}^r$ , we have*

$$672 \quad (4.9) \quad \frac{1}{2} \times (1 - \epsilon) \leq \frac{\max_{u \in \mathcal{G}_{\mathcal{S}_n, i}^r} \text{HVI}_r(u, \mathcal{S}_n)}{\lambda(\mathcal{G}_{\mathcal{S}_n, i}^r)} \leq \frac{1}{2} \times (1 + \epsilon) \text{ and}$$

$$673 \quad (4.10) \quad \frac{1}{4 \times (1 + \epsilon)} \leq \frac{\lambda(\mathcal{G}_{\mathcal{S}_n, j}^r)}{\lambda(\mathcal{G}_{\mathcal{S}_n, i}^r)} \leq \frac{1}{4 \times (1 - \epsilon)} .$$

674

675 *Proof.* The interval  $[w_{i-1,r}^n, w_{i,r}^n]$  is the set of the first coordinates of the vectors of  
676 the Pareto front which dominate  $r_i^n$ . The restriction to  $[w_{i-1,r}^n, w_{i,r}^n]$  of  $f$  is  $(L_{\min}, L_{\max})$ -  
677 bilipschitz for some  $L_{\min}$  and  $L_{\max}$  such that  $\frac{L_{\max}}{L_{\min}} = q_n$  with  $q_n$  defined in (4.8). Ad-  
678 ditionally, as stated in **Proposition B.2**, for  $n$  large enough, all the  $r_i^n$  corresponding to  
679 non-empty gap regions dominate the nadir point. As a consequence, the conditions to  
680 apply **Lemma A.2** and **Proposition 3.11** are met by non-extremes gap regions.

681 By **Propositions 3.10** and **3.11**,  $\frac{\max_{u \in \text{PF}_f} \text{HVI}_{r_i^n}(u)}{\text{HV}_{r_i^n}(\text{PF}_f)}$  is between  $\frac{1}{2} \times \frac{1}{q_n}$  and  $\frac{1}{2} \times q_n$ . Ad-  
682 ditionally, by **Lemma A.2**,  $\frac{\lambda(\mathcal{G}_{\mathcal{S}_n, j}^r)}{\text{HV}_{r_i^n}(\text{PF}_f)}$  is between  $\frac{1 - \frac{1}{2} \times q_n}{1 + q_n^2}$  and  $\frac{1 - \frac{1}{2} \times \frac{1}{q_n}}{1 + \frac{1}{q_n^2}}$ . The maximum

683 over the vectors  $u$  belonging to the Pareto front of  $\text{HV}_{r_i^n}(u)$  is equal to the maximum  
684 over  $u$  belonging to the  $i$ -th gap region of  $\mathcal{S}_n$  of  $\text{HVI}_r(u, \mathcal{S}_n)$ . Indeed,  $\text{HV}_{r_i^n}(\cdot)$  is null for  
685 vectors outside the  $i$ -th gap region of  $\mathcal{S}_n$  while it is nonnegative, equal to  $\text{HVI}_r(\cdot, \mathcal{S}_n)$ ,  
686 otherwise. Additionally,  $\text{HV}_{r_i^n}(\text{PF}_f)$  equals  $\lambda(\mathcal{G}_{\mathcal{S}_n, i}^r)$ . The convergence of  $q_n$  to 1 stated  
687 in [Lemma 4.5](#) imply that the bounds proven so far converge to a half and a quarter,  
688 respectively. Thus, we have [\(4.9\)](#) and [\(4.10\)](#) for  $n$  large enough.  $\square$

689 The following lemma states that for  $n$  large enough, the area of two non-empty gap regions  
690 relative to the same greedy set cannot be too different. More precisely, the area of any  
691 gap region of  $\mathcal{S}_n$  cannot be more than  $4 \times (1 + o(\epsilon))$  times greater than the area of another  
692 gap region of  $\mathcal{S}_n$ . The proof relies on considering the parents of the gap regions.

693 **LEMMA 4.7.** *We consider a greedy set sequence  $(\mathcal{S}_n)_{n \in \mathbb{N}^*}$  relative to a valid reference*  
694 *point  $r$ . If the Pareto front is described by a bilipschitz function with a Hölder continuous*  
695 *derivative, then for all  $\epsilon > 0$ , for  $n$  large enough and for any non-empty gap regions of*  
696  *$\mathcal{S}_n$ ,  $\mathcal{G}_{\mathcal{S}_n, i}^r$  and  $\mathcal{G}_{\mathcal{S}_n, j}^r$  with  $i, j \in \llbracket 1, n + 1 \rrbracket$ , we have*

$$697 \quad (4.11) \quad \frac{\lambda(\mathcal{G}_{\mathcal{S}_n, i}^r)}{\lambda(\mathcal{G}_{\mathcal{S}_n, j}^r)} \leq 4 \times \frac{(1 + \epsilon)^2}{1 - \epsilon} .$$

698

699 *Proof.* Fix  $\epsilon > 0$ . By [Lemma 4.6](#), there exists  $N_1 \in \mathbb{N}^*$  such that for all  $n$  greater  
700 than  $N_1$ , [\(4.9\)](#) and [\(4.10\)](#) are verified for any non-empty gap region of  $\mathcal{S}_n$  and its children.  
701 Since  $\max_{i \in \llbracket 1, n + 1 \rrbracket} w_{i, r}^n - w_{i-1, r}^n$  converges to 0 by [Lemma 4.4](#), every non-empty gap region  
702 is filled at some point. Take  $N_2$  such that all the non-empty gap regions of  $\mathcal{S}_{N_1}$  are filled  
703 at iteration  $N_2$ . For all  $n$  greater than  $N_2$ , [\(4.9\)](#) and [\(4.10\)](#) are true for any non-empty  
704 gap region of  $\mathcal{S}_n$  and its children, but also for its parents.

705 Take  $n \geq N_2$ . We note  $\mathcal{G}_1 := \mathcal{G}_{\mathcal{S}_n, i}^r$  and  $\mathcal{G}_2 := \mathcal{G}_{\mathcal{S}_n, j}^r$  two distinct non-empty gap regions  
706 of  $\mathcal{S}_n$ , and  $\mathcal{P}_1$  and  $\mathcal{P}_2$  their respective parents. When two sets correspond to gap regions  
707 relative to the same greedy set  $\mathcal{S}_m$ , we say that they cohabit at iteration  $m$ . Since only  
708 one vector is added to  $\mathcal{S}_n$  at a time, the cohabitation of  $\mathcal{G}_1$  and  $\mathcal{G}_2$  implies that either  $\mathcal{G}_1$   
709 and  $\mathcal{P}_2$  or  $\mathcal{G}_2$  and  $\mathcal{P}_1$  cohabit at some earlier iteration. In the first case, there necessarily  
710 exists  $m \geq N_2$  such that  $\mathcal{P}_2$  and  $\mathcal{G}_1$  are gap regions relative to  $\mathcal{S}_m$  and  $v_{m+1}$  belongs  
711 to  $\mathcal{P}_2$ , otherwise,  $\mathcal{G}_1$  and  $\mathcal{G}_2$  would not cohabit. By [\(4.9\)](#), the maximum hypervolume  
712 improvement to  $\mathcal{S}_m$  of a vector of  $\mathcal{G}_1$  and of a vector of  $\mathcal{P}_2$  are at least  $\frac{1}{2} \times (1 - \epsilon) \times \lambda(\mathcal{G}_1)$   
713 and at most  $\frac{1}{2} \times (1 + \epsilon) \times \lambda(\mathcal{P}_2)$ , respectively. Since a vector of  $\mathcal{P}_2$ ,  $v_{m+1}$ , maximizes the  
714 hypervolume improvement to  $\mathcal{S}_m$ , we have  $\lambda(\mathcal{G}_1) \times \frac{1}{2} \times (1 - \epsilon) \leq \lambda(\mathcal{P}_2) \times \frac{1}{2} \times (1 + \epsilon)$ . Since  
715  $\lambda(\mathcal{P}_2)$  is smaller than  $4 \times (1 + \epsilon)$  times the area of its child  $\lambda(\mathcal{G}_2)$  by [\(4.10\)](#), this inequality  
716 implies [\(4.11\)](#). In the second case,  $\mathcal{P}_2$  is filled before  $\mathcal{P}_1$ . Thus, there exists  $m \geq N_2$  such  
717 that  $\mathcal{P}_1$  and  $\mathcal{P}_2$  cohabit at iteration  $m$  and  $v_{m+1}$  belongs to  $\mathcal{P}_2$ . Since the area of  $\mathcal{P}_1$  is  
718 larger than the one of its child  $\mathcal{G}_1$ , the hypervolume improvement of  $v_{m+1}$  to  $\mathcal{S}_m$  is still  
719 larger than  $\frac{1}{2} \times (1 - \epsilon) \times \lambda(\mathcal{G}_1)$ . The rest of the argumentation remains valid.  $\square$

720 We now have all the results needed to analyze the asymptotic impact of doubling the  
721 number of points in the greedy set. To prove the following asymptotic upper bound,  
722 we rely on similar arguments as for its nonasymptotic counterpart, [Proposition 4.1](#). The  
723 previous lemma guarantees that the impact of doubling the number of points in the greedy  
724 set is asymptotically similar to the impact of passing from  $n$  points to  $2n + 1$ .

725 **PROPOSITION 4.8.** *Let  $(\mathcal{S}_n)_{n \in \mathbb{N}^*}$  be a greedy set sequence relative to valid reference*  
726 *point  $r$ . If the Pareto front is described by a bilipschitz function  $f$  with a Hölder continuous*

727 derivative, then for all  $\epsilon > 0$ , we have for  $n$  large enough

$$728 \quad (4.12) \quad \frac{HV_r(PF_f) - HV_r(\mathcal{S}_{2n})}{HV_r(PF_f) - HV_r(\mathcal{S}_n)} \leq \frac{1}{2} + o(\epsilon) .$$

729

730 *Proof.* Fix  $\epsilon > 0$ . Fix  $n$  large enough to verify (4.9) and (4.11) for this particular  $\epsilon$ .  
731 Let  $\sigma$  be a permutation of  $\llbracket 1, n+1 \rrbracket$  such that the  $i$ -th gap region of  $\mathcal{S}_n$  is filled by  
732  $v_{n+\sigma(i)}$  when it is filled before iteration  $2n+1$ . With this choice of permutation,  $\mathcal{G}_{\mathcal{S}_n, i}^r$   
733 is always a gap region of  $\mathcal{S}_{n+\sigma(i)-1}$ . Thus,  $HVI_r(v_{n+\sigma(i)}, \mathcal{S}_{n+\sigma(i)-1})$  is superior to the  
734 maximum hypervolume improvement of a vector of  $\mathcal{G}_{\mathcal{S}_n, i}^r$  to  $\mathcal{S}_{n+\sigma(i)-1}$ , which is superior  
735 to  $\frac{1}{2} \times (1-\epsilon) \times \lambda(\mathcal{G}_{\mathcal{S}_n, i}^r)$  by (4.9). It is equivalent to say that the hypervolume improvement  
736 at iteration  $n+\sigma(i)$  is larger than  $\frac{1}{2} \times (1-\epsilon) \times \lambda(\mathcal{G}_{\mathcal{S}_n, i}^r)$ . Summing over  $i \in \llbracket 1, n+1 \rrbracket$ , we  
737 obtain that the hypervolume improvement between iteration  $n$  and  $2n+1$  is larger than  
738 the sum over  $i$  of  $\frac{1}{2} \times (1-\epsilon) \times \lambda(\mathcal{G}_{\mathcal{S}_n, i}^r)$ , that is  $\frac{1}{2} \times (1-\epsilon)$  times the optimality gap at  
739 iteration  $n$ .

740 Now, we need to bound the hypervolume improvement at iteration  $2n+1$ , that is  
741  $HVI_r(v_{2n+1}, \mathcal{S}_{2n})$ . It is smaller than  $\frac{1}{2} \times (1+\epsilon) \times \max_{i \in \llbracket 1, 2n+1 \rrbracket} \lambda(\mathcal{G}_{\mathcal{S}_{2n}, i}^r)$  by (3.3) and  
742 (4.9). Since the area of a gap region is smaller than the one of its parent, the maximum  
743 area of a gap region is lower at iteration  $2n$  than at iteration  $n$ . The maximum area of  
744 one of the more than  $n-1$  gap regions of  $\mathcal{S}_n$  is itself smaller than  $\frac{1}{n-1} \times \frac{4 \times (1+\epsilon)^2}{1-\epsilon}$  times  
745 the optimality gap at iteration  $n$  by (4.11).

746 We conclude that the relation between the optimality gap at iteration  $2n$  and at  
747 iteration  $n$  is smaller than  $1 - \frac{1}{2} \times (1-\epsilon) + \frac{1-\epsilon}{2 \times (n-1)}$ .  $\square$

748 We roughly follow the same approach to obtain the following asymptotic lower bound on  
749 the impact of doubling the number of points in the greedy set. Lemmas 4.6 and 4.7 are  
750 key to prove an upper bound on the hypervolume improvement at iteration  $k$ . They allow  
751 to prove that filling a gap region of  $\mathcal{S}_n$  more than once gives, up to a factor  $1+o(\epsilon)$ , a  
752 lower hypervolume improvement than filling a gap region which was not filled. Indeed,  
753 the area of a descendant of a gap region of  $\mathcal{S}_n$  is at most  $\frac{1}{4} + o(\epsilon)$  times the area of its  
754 parent by Lemma 4.6, which is itself at most  $4 + o(\epsilon)$  times the area of any other gap  
755 region of  $\mathcal{S}_n$  by Lemma 4.7.

756 **PROPOSITION 4.9.** *Let  $(\mathcal{S}_n)_{n \in \mathbb{N}^*}$  be a greedy set sequence relative to a valid reference*  
757 *point  $r$ . If the Pareto front is described by a bilipschitz function  $f$  with a Hölder continuous*  
758 *derivative, then for all  $\epsilon > 0$ , we have for  $n$  large enough*

$$759 \quad (4.13) \quad \frac{HV_r(PF_f) - HV_r(\mathcal{S}_{2n})}{HV_r(PF_f) - HV_r(\mathcal{S}_n)} \geq \frac{1}{2} + o(\epsilon) .$$

760

761 *Proof.* Fix  $\epsilon > 0$ . Fix  $n$  large enough to verify (4.9), (4.10) and (4.11) for this  
762 particular  $\epsilon$ . Let  $\delta \in \{-1, 0, 1\}$  be such that  $\mathcal{S}_n$  has  $n+\delta$  non-empty gap regions. Let  
763  $i_0 := 1$  when the left extreme gap region is empty and  $i_0 := 0$  otherwise.

764 Let  $\sigma$  be a permutation of  $\llbracket 1, n+\delta \rrbracket$  such that the  $i$ -th non-empty gap region of  
765  $\mathcal{S}_n$ ,  $\mathcal{G}_{\mathcal{S}_n, i_0+i}^r$ , is filled by the vector  $v_{n+\sigma(i)}$  when it is filled before iteration  $2n+\delta$ . We  
766 distinguish two cases. In the first case,  $v_{n+\sigma(i)}$  is the child of the  $i$ -th non-empty gap  
767 region of  $\mathcal{S}_n$ , and consequently its hypervolume improvement to  $\mathcal{S}_{n+\sigma(i)-1}$  is at most  
768  $\frac{1}{2} \times (1+\epsilon) \times \lambda(\mathcal{G}_{\mathcal{S}_n, i_0+i}^r)$  by (4.9). In the second case,  $v_{n+\sigma(i)}$  belongs to  $\mathcal{G}_{\mathcal{S}_n, i_0+j}^r$ , the  
769  $j$ -th non-empty gap region of  $\mathcal{S}_n$ , with  $j \neq i$  and, by definition of  $\sigma$ , fills a descendant of



770 this gap region not  $\mathcal{G}_{\mathcal{S}_n, i_0+j}^r$  itself. By (4.9), the hypervolume improvement of  $v_{n+\sigma(i)}$  to  
771  $\mathcal{S}_{n+\sigma(i)-1}$  is still at most  $\frac{1}{2} \times (1 + \epsilon)$  times the area of the gap region it fills. By (4.10),  
772 the area of a descendant of  $\mathcal{G}_{\mathcal{S}_n, i_0+j}^r$  is smaller than  $\frac{1}{4 \times (1-\epsilon)}$  times the area of its ancestor.  
773 By (4.11), we also know that the area of the  $i$ -th non-empty gap region of  $\mathcal{S}_n$  is at most  
774  $4 \times \frac{(1+\epsilon)^2}{1-\epsilon}$  times the area of any other gap region of  $\mathcal{S}_n$ , in particular its  $i$ -th non-empty  
775 gap region. We conclude that the hypervolume improvement of  $v_{n+\sigma(i)}$  to  $\mathcal{S}_{n+\sigma(i)-1}$  is  
776 smaller than  $\frac{1}{2} \times \frac{(1+\epsilon)^3}{(1-\epsilon)^2} \times \lambda(\mathcal{G}_{\mathcal{S}_n, i_0+i}^r)$ . To summarize, since  $1 + \epsilon$  is smaller than  $\frac{(1+\epsilon)^3}{(1-\epsilon)^2}$ , the  
777 hypervolume improvement at any iteration  $n + \sigma(i)$  is smaller than  $\frac{1}{2} \times \frac{(1+\epsilon)^3}{(1-\epsilon)^2} \times \lambda(\mathcal{G}_{\mathcal{S}_n, i}^r)$ .  
778 Summing over  $i \in [1, n + \delta]$ , the hypervolume improvement from iteration  $n$  to  $2n + \delta$   
779 is smaller than  $\frac{1}{2} \times \frac{(1+\epsilon)^3}{(1-\epsilon)^2}$  times the sum over  $i$  of  $\lambda(\mathcal{G}_{\mathcal{S}_n, i}^r)$ , that is the optimality gap at  
780 iteration  $n$ .

781 Now, it is left to prove an upper bound on  $\text{HV}_r(\mathcal{S}_{2n}) - \text{HV}_r(\mathcal{S}_{2n+\delta})$ . This quantity  
782 is maximal for  $\delta = -1$ , where it is simply the hypervolume improvement at iteration  $2n$ .  
783 As in the previous proof, it is smaller than  $\frac{1+\epsilon}{2 \times (n-1)}$  times the optimality gap at iteration  
784  $n$ . Therefore, the relation between the optimality gap at iteration  $2n$  and at iteration  $n$   
785 is larger than  $1 - \frac{1}{2} \times \frac{(1+\epsilon)^3}{(1-\epsilon)^2} - \frac{1+\epsilon}{2 \times (n-1)}$ .  $\square$

786 We combine the lower and upper asymptotic bounds to obtain the following theorem.

787 **THEOREM 4.10.** *Consider a biobjective optimization problem and a greedy set se-*  
788 *quence  $(\mathcal{S}_n)_{n \in \mathbb{N}^*}$  relative to a valid reference point  $r$ . If the Pareto front is described by a*  
789 *bilipschitz function  $f$  with a Hölder continuous derivative, we have*

$$790 \quad (4.14) \quad \frac{\text{HV}_r(\text{PF}_f) - \text{HV}_r(\mathcal{S}_{2n})}{\text{HV}_r(\text{PF}_f) - \text{HV}_r(\mathcal{S}_n)} \xrightarrow{n \rightarrow \infty} \frac{1}{2} .$$

792 *Consequently, for such Pareto front and reference point and under [Assumption 2.3](#) of*  
793 *perfect singleobjective optimization, HV-ISOOMOO algorithms relative to  $r$  satisfy (4.14)*  
794 *where  $\mathcal{S}_n$  is replaced by  $\mathcal{I}_n$ , the final incumbents Pareto front approximation at iteration*  
795  *$n$ .*

796 **5. Conclusion.** We prove that HV-ISOOMOO algorithms coupled with a singleob-  
797 jective optimizer converge in  $O(1/n)$  on convex Pareto fronts and in  $O(1/n^c)$  on bilipschitz  
798 Pareto fronts with  $c \leq 1$  depending on the bilipschitz constants where  $n$  is the number  
799 of meta-iterations. Each meta-iteration corresponds to a singleobjective optimization  
800 run. Both bounds are tight over the class of Pareto fronts and reference points consid-  
801 ered. They are reached for affine Pareto fronts and reference points dominating the nadir  
802 point. On convex Pareto fronts, the convergence is exactly in  $\Theta(1/p)$ , the fastest con-  
803 vergence achievable by biobjective optimization algorithms [16]. It shows that greedily  
804 adding points maximizing the hypervolume contribution as in HV-ISOOMOO algorithms  
805 is an effective way to quickly increase the hypervolume. Additionally, we prove that for  
806 bilipschitz Pareto fronts with a Hölder continuous derivative, doubling the number of  
807 meta-iterations divides the optimality gap by a factor which converges asymptotically to  
808 two. This asymptotic behavior resembles what we would observe with an affine Pareto  
809 front and a reference point dominating the nadir point. Yet, it does not guarantee con-  
810 vergence in  $\Theta(1/n)$ . Both  $\left(\frac{\log(n)}{n}\right)_{n \in \mathbb{N}^*}$  and  $\left(\frac{1}{n \times \log(n)}\right)_{n \in \mathbb{N}^*}$  are examples of sequences  
811 verifying this property without converging in  $\Theta(1/n)$ . The convergence on nonconvex



812 Pareto fronts could theoretically be slower than in  $\Theta(1/n)$ , but not faster [16].

813 **Appendix A. Normalized areas of the gap regions relative to an hyper-**  
 814 **volume maximizer.** The goal of this section is to prove bounds on the normalized  
 815 areas of the gap regions  $\mathcal{G}_{\text{left}}^{u^*}$  and  $\mathcal{G}_{\text{right}}^{u^*}$  relative to an hypervolume maximizer  $u^*$  (see the  
 816 lefthand plot of Figure 6) in the case of a bilipschitz Pareto front and of a reference point  
 817  $r$  dominating the nadir point. These bounds are stated in Lemma A.2. The proof relies  
 818 on the bounds on the normalized maximum hypervolume proven in Subsection 3.3 and  
 the following lower and upper bounds on the relation between  $\lambda(\mathcal{G}_{\text{left}}^{u^*})$  and  $\lambda(\mathcal{G}_{\text{right}}^{u^*})$ .

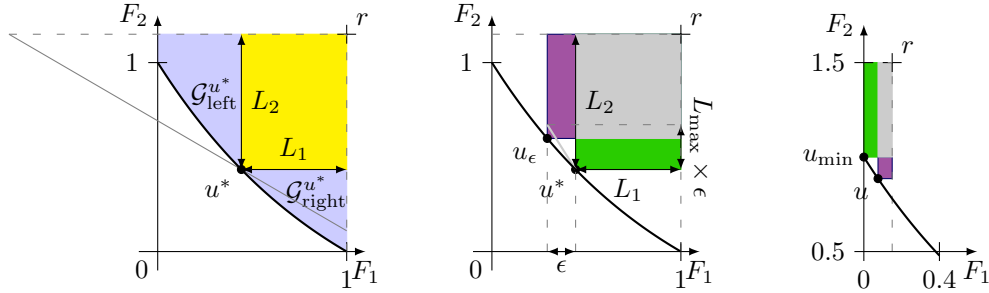


Fig. 6: Illustration of elements of the proofs of Proposition A.1 in the case  $r_1 \leq x_{\max}$  (left and middle) and of Lemma B.1 (right). The Pareto front is described by  $f(x) = \frac{e}{e-1} \times e^{-x} + 1 - \frac{e}{e-1}$  for  $x \in [0, 1]$ . Left: the gap regions  $\mathcal{G}_{\text{left}}^{u^*}$  and  $\mathcal{G}_{\text{right}}^{u^*}$  with a segment of slope  $-L_{\min}$  passing through  $u^*$ . Middle: the hypervolume improvements  $\text{HVI}_r(u_\epsilon, u^*)$  (■) and  $\text{HVI}_r(u_\epsilon, u^*)$  (■). Right: the hypervolume improvement  $\text{HVI}_r(u_{\min}, u)$  (■) and its counterpart  $\text{HVI}_r(u, u_{\min})$  (■) where  $u$  is a vector of the Pareto front which dominates  $r$ .

819

820 **PROPOSITION A.1.** *We assume that the Pareto front is described by a  $(L_{\min}, L_{\max})$ -*  
 821 *bilipschitz function  $f$ . Let  $u^*$  be a non-extreme vector of the Pareto front which maximizes*  
 822 *the hypervolume with respect to a valid reference point  $r$ . If  $r_1 \leq x_{\max}$ , we have  $\lambda(\mathcal{G}_{\text{right}}^{u^*}) \geq$*   
 823  *$\frac{L_{\min}^2}{L_{\max}^2} \times \lambda(\mathcal{G}_{\text{left}}^{u^*})$ . If  $r_2 \leq f(x_{\min})$ , we have  $\lambda(\mathcal{G}_{\text{left}}^{u^*}) \geq \frac{L_{\min}^2}{L_{\max}^2} \times \lambda(\mathcal{G}_{\text{right}}^{u^*})$ .*

824 *Proof.* We consider the case where  $r_1 \leq x_{\max}$ . Let  $x^*$  be the first coordinate of  $u^*$ .  
 825 We denote  $L_1 := r_1 - x^*$  and  $L_2 := r_2 - f(x^*)$  the lengths of the sides of the rectangle  
 826  $\mathcal{D}_{u^*}^r$ . For all  $x, y \in [x_{\min}, x_{\max}]$ , we have  $|f(x) - f(y)| \geq L_{\min} \times |x - y|$ . Additionally,  
 827 since  $r_1 \leq x_{\max}$ , the segment  $[x^*, x^* + L_1]$  is included in  $[x_{\min}, x_{\max}]$ . As a consequence,  
 828 the section of the Pareto front on the right of  $u^*$  dominates the segment between  $u^*$  and  
 829  $u^* + L_1 \times (1, -L_{\min})$ , see the lefthand plot of Figure 6. Therefore,  $\lambda(\mathcal{G}_{\text{right}}^{u^*})$  is larger than  
 830 the area of the region of the objective space dominated by this segment, not dominated  
 831 by  $u^*$  and dominating  $r$ , that is  $\frac{1}{2} \times L_{\min} \times L_1^2$ . For all  $x, y \in [x_{\min}, x_{\max}]$ , we also have  
 832  $|f(x) - f(y)| \leq L_{\max} \times |x - y|$ . Therefore, the part of the Pareto front on the left of  $u^*$  is  
 833 dominated by the segment between  $u^*$  and  $u^* + L_2 \times (-\frac{1}{L_{\min}}, 1)$ , and  $\lambda(\mathcal{G}_{\text{left}}^{u^*})$  is smaller  
 834 than  $\frac{1}{2} \times \frac{1}{L_{\min}} \times L_2^2$ . We have yet to prove a lower bound on  $\frac{L_1}{L_2}$ . The vector  $u^*$  being  
 835 different from  $u_{\min}$ , for  $\epsilon > 0$  small enough, the vector  $u_\epsilon := (x^* - \epsilon, f(x^* - \epsilon))$  belongs  
 836 to the Pareto front. As we can see in the middle plot of Figure 6,  $\text{HVI}_r(u^*, u_\epsilon)$  is smaller

837 than  $L_1 \times L_{\max} \times \epsilon$  and  $\text{HVI}_r(u_\epsilon, u^*)$  is larger than  $\epsilon \times (L_2 - \epsilon \times L_{\max})$ . Additionally,  
 838  $u^*$  being an hypervolume maximizer,  $\text{HVI}_r(u^*, u_\epsilon)$  is larger than  $\text{HVI}_r(u_\epsilon, u^*)$ , and thus  
 839  $L_1 \times L_{\max} \geq L_2 - \epsilon \times L_{\max}$  for all  $\epsilon > 0$ . Taking the limit of this inequality when  $\epsilon \rightarrow 0$ ,  
 840 we obtain that  $L_1 \times L_{\max} \geq L_2$ . Combining the bounds on  $\lambda(\mathcal{G}_{\text{left}}^{u^*})$  and  $\lambda(\mathcal{G}_{\text{right}}^{u^*})$  with the  
 841 lower-bound on  $\frac{L_1}{L_2}$ , we obtain the desired lower bound on  $\lambda(\mathcal{G}_{\text{right}}^{u^*})$ . We can obtain the  
 842 symmetric inequality when  $r_2 \geq f(x_{\min})$  by following the same approach.  $\square$

843 In particular, when  $f$  is bilipschitz and  $r$  dominates the nadir point, both bounds hold.  
 844 We now prove the desired bounds on the normalized area of the gap regions  $\mathcal{G}_{\text{left}}^{u^*}$  and  
 845  $\mathcal{G}_{\text{right}}^{u^*}$ .

846 **LEMMA A.2.** *Let  $u^*$  be a vector which maximizes the hypervolume with respect to*  
 847 *a valid reference point  $r$ . If the Pareto front is described by a  $(L_{\min}, L_{\max})$ -bilipschitz*  
 848 *function  $f$  and the reference point  $r$  dominates the nadir point, both  $\lambda(\mathcal{G}_{\text{left}}^{u^*})$  and  $\lambda(\mathcal{G}_{\text{right}}^{u^*})$*   
 849 *are between  $(1 - \frac{1}{2} \times \frac{L_{\max}}{L_{\min}}) / (1 + \frac{L_{\max}^2}{L_{\min}^2})$  and  $(1 - \frac{1}{2} \times \frac{L_{\min}}{L_{\max}}) / (1 + \frac{L_{\min}^2}{L_{\max}^2})$ .*

850 *Proof.* Let note arbitrarily  $\mathcal{G}_1$  and  $\mathcal{G}_2$  the two gap regions of the set  $S = \{u^*\}$ .  
 851 By [Proposition A.1](#),  $\lambda(\mathcal{G}_2)$  is between  $\frac{L_{\min}^2}{L_{\max}^2} \times \lambda(\mathcal{G}_1)$  and  $\frac{L_{\max}^2}{L_{\min}^2} \times \lambda(\mathcal{G}_1)$ . Additionally,  
 852 by [Propositions 3.10](#) and [3.11](#), the normalized maximum hypervolume  $\max_{u \in \text{PF}_f} \text{HV}_r(u)$   
 853 over  $\text{HV}_r(\text{PF}_f)$  is between  $\frac{1}{2} \times \frac{L_{\min}}{L_{\max}}$  and  $\frac{1}{2} \times \frac{L_{\max}}{L_{\min}}$ . These bounds can be transformed into  
 854 bounds on  $\text{HV}_r(\text{PF}_f) - \max_{u \in \text{PF}_f} \text{HV}_r(u)$ , that is  $\lambda(\mathcal{G}_1) + \lambda(\mathcal{G}_2)$ . As a consequence,  $\lambda(\mathcal{G}_1)$   
 855 is between  $(1 - \frac{1}{2} \times \frac{L_{\max}}{L_{\min}}) \times \text{HV}_r(\text{PF}_f) - \frac{L_{\max}^2}{L_{\min}^2} \times \lambda(\mathcal{G}_1)$  and  $(1 - \frac{1}{2} \times \frac{L_{\min}}{L_{\max}}) \times \text{HV}_r(\text{PF}_f) -$   
 856  $\frac{L_{\min}^2}{L_{\max}^2} \lambda(\mathcal{G}_1)$ . Moving all the  $\lambda(\mathcal{G}_1)$  terms on the same side and re-normalizing this side, we  
 857 obtain the desired bounds for  $\mathcal{G}_1$ , which can be chosen to be either  $\mathcal{G}_{\text{left}}^{u^*}$  or  $\mathcal{G}_{\text{right}}^{u^*}$ .  $\square$

858 **Appendix B. The nadir point is dominated by all the  $r_i^n$  corresponding to**  
 859 **non-empty gap regions for  $n$  large.** We show in this section that for bilipschitz  
 860 Pareto fronts, the nadir point is dominated by all the  $r_i^n$  corresponding to non-empty  
 861 gap regions, for  $n$  large enough. This result is stated in [Proposition B.2](#) and used in  
 862 [Subsection 4.2](#). It is equivalent to prove that the extreme vectors which dominate the  
 863 reference point belong to the greedy set for  $n$  large enough.

864 First, we prove in the next proposition that if  $r_1 > x_{\max}$  (resp.  $r_2 > f(x_{\min})$ ), then  
 865 for  $r_2$  (resp.  $r_1$ ) close enough to  $f(x_{\max})$  (resp.  $x_{\min}$ ) the extreme vector  $u_{\max}$  (resp.  
 866  $u_{\min}$ ) is the only hypervolume maximizer, see the righthand plot of [Figure 6](#). There are  
 867 similar statements in [\[9\]](#) for the set of  $\mu$  points maximizing the hypervolume, but they  
 868 only apply to  $\mu \geq 2$ .

869 **LEMMA B.1.** *We assume that the Pareto front is described by a function  $f$  which is*  
 870  *$(L_{\min}, L_{\max})$ -bilipschitz and that the reference point  $r$  is valid. If  $r_1 > x_{\max}$  and  $f(x_{\max}) <$*   
 871  *$r_2 < f(x_{\max}) + L_{\min} \times (r_1 - x_{\max})$ , the right extreme of the Pareto front  $u_{\max}$  is the only*  
 872 *maximizer of  $\text{HV}_r(\cdot)$ . Additionally, if  $r_2 > f(x_{\min})$  and  $x_{\min} < r_1 < x_{\min} + \frac{r_2 - f(x_{\min})}{L_{\max}}$ ,*  
 873 *the vector  $u_{\min} = (x_{\min}, f(x_{\min}))$  is the only maximizer of  $\text{HV}_r(\cdot)$ .*

874 *Proof.* This proof is illustrated in the righthand plot of [Figure 6](#). Let  $r$  be a reference  
 875 point such that  $r_2 > f(x_{\min})$  and  $x_{\min} < r_1 < x_{\min} + \frac{r_2 - f(x_{\min})}{L_{\max}}$ . Let  $u = (x, f(x)) \neq u_{\min}$   
 876 be a vector of the Pareto front which dominates  $r$ . The hypervolume improvement of  $u_{\min}$   
 877 to  $\{u\}$  is  $(r_2 - f(x_{\min})) \times (x - x_{\min})$ . The hypervolume improvement of  $u$  to  $\{u_{\min}\}$  is  
 878  $(f(x_{\min}) - f(x)) \times (r_1 - x)$ , which is smaller than  $L_{\max} \times (x - x_{\min}) \times (r_1 - x_{\min})$  since  $u$

879 dominates  $r$  and  $f$  is  $(L_{\min}, L_{\max})$ -bilipschitz. Since we assume that  $L_{\max} \times (r_1 - x_{\min}) <$   
 880  $r_2 - f(x_{\min})$ , the upper bound on  $\text{HVI}_r(u, u_{\min})$  is strictly smaller than  $\text{HVI}_r(u_{\min}, u)$ .  
 881 Therefore, the hypervolume of  $u_{\min}$  is strictly larger than the one of  $u$ . We conclude that  
 882  $u_{\min}$  is the unique hypervolume maximizer. The symmetric result can be obtained with  
 883 the same approach.  $\square$

884 It is left to prove that when  $r_1 > x_{\max}$  (resp.  $r_2 > f(x_{\min})$ ), the second coordinate of  
 885  $r_{n+1}^n$  (resp. the first coordinate of  $r_0^n$ ) indeed converge to  $f(x_{\max})$  (resp.  $x_{\min}$ ). It is a  
 886 straightforward consequence of [Lemma 4.4](#). Therefore, we are able to conclude.

887 **PROPOSITION B.2.** *We assume that the Pareto front is described by a bilipschitz func-*  
 888 *tion  $f$ . Let  $(\mathcal{S}_n)_{n \in \mathbb{N}^*}$  be a greedy set sequence relative to a valid reference point  $r$ . For*  
 889  *$n$  large enough, every reference point  $r_i^n$  corresponding to a non-empty gap region  $\mathcal{G}_{\mathcal{S}_n, i}^r$*   
 890 *dominates the nadir point.*

891 *Proof.* By [Lemma 4.4](#),  $w_{n,r}^n$  converges to  $x_{\max}$ , and thus the right extreme reference  
 892 point  $r_{n+1}^n := (r_1, f(w_{n,r}^n))$  converges to  $(r_1, f(x_{\max}))$  by continuity of  $f$ . Therefore, if  $r_1$   
 893 is strictly larger than  $x_{\max}$ , then there exists  $N$  such that for all  $n \geq N$ ,  $r_{n+1}^n$  verifies  
 894 the assumptions on the reference point of [Lemma B.1](#) which guarantee that  $u_{\max}$  is the  
 895 unique maximizer of  $\text{HV}_r(\cdot)$  over the right extreme gap region  $\mathcal{G}_{\mathcal{S}_n, n+1}^r$ . Let assume that  
 896  $u_{\max}$  does not belong to  $\mathcal{S}_n$ . Then,  $w_{N,r}^N \neq x_{\max}$ , and since  $w_{n,r}^n$  converges to  $x_{\max}$ , the  
 897 left extreme gap region  $\mathcal{G}_{\mathcal{S}_n, i}^r$  is necessarily filled at some later iteration. When the right  
 898 extreme gap region is filled,  $u_{\max}$ , the unique minimizer of  $\text{HV}_r(\cdot)$  over this gap region,  
 899 is added to the greedy set. To summarize, if  $r_1 > x_{\max}$ , then for  $n$  large enough  $\mathcal{S}_n$   
 900 contains  $u_{\max}$ , and thus the right extreme gap region is empty. We can prove with the  
 901 same approach that for  $r_2 > f(x_{\min})$ ,  $\mathcal{S}_n$  contains  $u_{\min}$  for  $n$  large enough.

902 At any iteration, the non-extreme reference points dominate the nadir point. Addi-  
 903 tionally, we proved that either  $r_1 < x_{\max}$  (resp.  $r_2 < f(x_{\min})$ ), and thus the left (resp.  
 904 right) extreme reference point dominates the nadir point or for  $n$  large enough, the left  
 905 (resp. right) extreme gap region is empty.  $\square$

906 **Acknowledgments.** We would like to thank Dimo Brockhoff and Nikolaus Hansen  
 907 for several helpful discussions, notably about the importance of decomposing the opti-  
 908 mality gap and for the conjecture of [\(4.14\)](#).

909

## REFERENCES

- 910 [1] C. AUDET, G. SAVARD, AND W. ZGHAL, *Multiobjective Optimization Through a Series of Single-*  
 911 *Objective Formulations*, SIAM Journal on Optimization, 19 (2008), pp. 188–210, [https://doi.](https://doi.org/10.1137/060677513)  
 912 [org/10.1137/060677513](https://doi.org/10.1137/060677513).  
 913 [2] A. AUGER, J. BADER, AND D. BROCKHOFF, *Theoretically investigating optimal  $\mu$ -distributions for*  
 914 *the hypervolume indicator: first results for three objectives*, in Proceedings of the 11th interna-  
 915 *tional conference on Parallel problem solving from nature: Part I, PPSN'10, Berlin, Heidelberg,*  
 916 *Sept. 2010, Springer-Verlag, pp. 586–596.*  
 917 [3] A. AUGER, J. BADER, D. BROCKHOFF, AND E. ZITZLER, *Theory of the Hypervolume Indicator:*  
 918 *Optimal  $\mu$ -Distributions and the Choice of the Reference Point*, in Proceedings of the Tenth  
 919 *ACM SIGEVO Workshop on Foundations of Genetic Algorithms, FOGA '09, New York, NY,*  
 920 *USA, Jan. 2009, Association for Computing Machinery, pp. 87–102, [https://doi.org/10.1145/](https://doi.org/10.1145/1527125.1527138)*  
 921 [1527125.1527138](https://doi.org/10.1145/1527125.1527138).  
 922 [4] A. AUGER, J. BADER, D. BROCKHOFF, AND E. ZITZLER, *Hypervolume-based multiobjective opti-*  
 923 *mization: Theoretical foundations and practical implications*, Theoretical Computer Science,  
 924 425 (2012), pp. 75–103, <https://doi.org/10.1016/j.tcs.2011.03.012>.

- 925 [5] N. BEUME, M. LAUMANN, AND G. RUDOLPH, *Convergence Rates of (1+1) Evolutionary Multiobjec-*  
926 *tive Optimization Algorithms*, in Parallel Problem Solving from Nature, PPSN XI, R. Schaefer,  
927 C. Cotta, J. Kołodziej, and G. Rudolph, eds., Lecture Notes in Computer Science, Berlin,  
928 Heidelberg, 2010, Springer, pp. 597–606, [https://doi.org/10.1007/978-3-642-15844-5\\_60](https://doi.org/10.1007/978-3-642-15844-5_60).
- 929 [6] N. BEUME, B. NAUJOKS, AND M. EMMERICH, *SMS-EMOA: Multiobjective selection based on dom-*  
930 *inated hypervolume*, European Journal of Operational Research, 181 (2007), pp. 1653–1669,  
931 <https://doi.org/10.1016/j.ejor.2006.08.008>.
- 932 [7] J. BIGEON, S. LE DIGABEL, AND L. SALOMON, *DMulti-MADS: Mesh adaptive direct multisearch*  
933 *for bound-constrained blackbox multiobjective optimization*, Computational Optimization and  
934 Applications, 79 (2021), pp. 301–338, <https://doi.org/10.1007/s10589-021-00272-9>.
- 935 [8] K. BRINGMANN AND T. FRIEDRICH, *The maximum hypervolume set yields near-optimal approx-*  
936 *imation*, in Proceedings of the 12th Annual Conference on Genetic and Evolutionary Compu-  
937 *tation - GECCO '10*, Portland, Oregon, USA, 2010, ACM Press, pp. 511 – 518, <https://doi.org/10.1145/1830483.1830576>.
- 938 [9] D. BROCKHOFF, *Optimal  $\mu$ -distributions for the hypervolume indicator for problems with linear bi-*  
939 *objective fronts: exact and exhaustive results*, in Proceedings of the 8th international conference  
940 *on Simulated evolution and learning, SEAL'10*, Berlin, Heidelberg, Dec. 2010, Springer-Verlag,  
941 pp. 24–34.
- 942 [10] A. L. CUSTÓDIO AND J. F. A. MADEIRA, *MultiGLODS: Global and local multiobjective optimization*  
943 *using direct search*, Journal of Global Optimization, 72 (2018), pp. 323–345, [https://doi.org/](https://doi.org/10.1007/s10898-018-0618-1)  
944 [10.1007/s10898-018-0618-1](https://doi.org/10.1007/s10898-018-0618-1).
- 945 [11] A. L. CUSTÓDIO, J. F. A. MADEIRA, A. I. F. VAZ, AND L. N. VICENTE, *Direct Multisearch for*  
946 *Multiobjective Optimization*, SIAM Journal on Optimization, 21 (2011), pp. 1109–1140, [https://doi.org/](https://doi.org/10.1137/10079731X)  
947 [10.1137/10079731X](https://doi.org/10.1137/10079731X).
- 948 [12] E. DE KLERK, F. GLINEUR, AND A. B. TAYLOR, *On the worst-case complexity of the gradient method*  
949 *with exact line search for smooth strongly convex functions*, Optimization Letters, 11 (2017),  
950 pp. 1185–1199, <https://doi.org/10.1007/s11590-016-1087-4>.
- 951 [13] R. FIORENZA, *Hölder and locally Hölder Continuous Functions, and Open Sets of Class  $C^k$ ,  $C^{k,\lambda}$* ,  
952 *Frontiers in Mathematics*, Birkhäuser Basel, 2016, <https://doi.org/10.1007/978-3-319-47940-8>.
- 953 [14] E. H. FUKUDA AND L. M. G. DRUMMOND, *On the convergence of the projected gradient method*  
954 *for vector optimization*, Optimization, 60 (2011), pp. 1009–1021, [https://doi.org/10.1080/](https://doi.org/10.1080/02331934.2010.522710)  
955 [02331934.2010.522710](https://doi.org/10.1080/02331934.2010.522710).
- 956 [15] J. KNOWLES AND D. CORNE, *On metrics for comparing nondominated sets*, in Proceedings of the  
957 *2002 Congress on Evolutionary Computation. CEC'02*, vol. 1, May 2002, pp. 711–716 vol.1,  
958 <https://doi.org/10.1109/CEC.2002.1007013>.
- 959 [16] E. MARESCAUX AND N. HANSEN, *Hypervolume in biobjective optimization cannot converge faster*  
960 *than  $\Omega(1/p)$* , in Proceedings of the Genetic and Evolutionary Computation Conference, GECCO  
961 '21, New York, NY, USA, June 2021, Association for Computing Machinery, pp. 430–438,  
962 <https://doi.org/10.1145/3449639.3459371>.
- 963 [17] G. L. NEMHAUSER, L. A. WOLSEY, AND M. L. FISHER, *An analysis of approximations for max-*  
964 *imizing submodular set functions—I*, Mathematical Programming, 14 (1978), pp. 265–294,  
965 <https://doi.org/10.1007/BF01588971>.
- 966 [18] C. TOURÉ, N. HANSEN, A. AUGER, AND D. BROCKHOFF, *Uncrowded hypervolume improvement:*  
967 *COMO-CMA-ES and the sofomore framework*, in Proceedings of the Genetic and Evolutionary  
968 *Computation Conference, GECCO '19*, New York, NY, USA, July 2019, Association for  
969 *Computing Machinery*, pp. 638–646, <https://doi.org/10.1145/3321707.3321852>.
- 970 [19] D. A. V. VELDHUIZEN AND G. B. LAMONT, *Evolutionary Computation and Convergence to a Pareto*  
971 *Front*, in Proceedings of the Third Annual Conference on Genetic Programming, San Francisco,  
972 1998, Stanford University Bookstore, pp. 221–228.
- 973 [20] E. ZITZLER, J. KNOWLES, AND L. THIELE, *Quality Assessment of Pareto Set Approximations*, in  
974 *Multiobjective Optimization: Interactive and Evolutionary Approaches*, J. Branke, K. Deb,  
975 K. Miettinen, and R. Słowiński, eds., Lecture Notes in Computer Science, Springer, Berlin,  
976 Heidelberg, 2008, pp. 373–404, [https://doi.org/10.1007/978-3-540-88908-3\\_14](https://doi.org/10.1007/978-3-540-88908-3_14).
- 977