



**HAL**  
open science

# Hot off the Press: The First Proven Performance Guarantees for the Non-Dominated Sorting Genetic Algorithm II (NSGA-II) on a Combinatorial Optimization Problem

Sacha Cerf, Benjamin Doerr, Benjamin Hebras, Yakob Kahane, Simon Wietheger

## ► To cite this version:

Sacha Cerf, Benjamin Doerr, Benjamin Hebras, Yakob Kahane, Simon Wietheger. Hot off the Press: The First Proven Performance Guarantees for the Non-Dominated Sorting Genetic Algorithm II (NSGA-II) on a Combinatorial Optimization Problem. GECCO '24 Companion: Proceedings of the Genetic and Evolutionary Computation Conference Companion, 2024, Melbourne, Australia. pp.27 - 28, 10.1145/3638530.3664080 . hal-04676956

**HAL Id: hal-04676956**

**<https://hal.science/hal-04676956>**

Submitted on 24 Aug 2024

**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.



# Hot off the Press: The First Proven Performance Guarantees for the Non-Dominated Sorting Genetic Algorithm II (NSGA-II) on a Combinatorial Optimization Problem

Sacha Cerf

École Polytechnique  
Institut Polytechnique de Paris  
Palaiseau, France

Benjamin Doerr

Laboratoire d'Informatique (LIX)  
École Polytechnique, CNRS  
Institut Polytechnique de Paris  
Palaiseau, France

Benjamin Hebras

École Polytechnique  
Institut Polytechnique de Paris  
Palaiseau, France

Yakob Kahane

École Polytechnique  
Institut Polytechnique de Paris  
Palaiseau, France

Simon Wietheger

Algorithms and Complexity Group  
TU Wien  
Vienna, Austria

## ABSTRACT

Recently, the first mathematical runtime guarantees have been obtained for the NSGA-II, one of the most prominent multi-objective optimization algorithms, however only for synthetic benchmark problems.

In this work, we give the first proven performance guarantees for a classic optimization problem, the NP-complete bi-objective minimum spanning tree problem. More specifically, we show that the NSGA-II with population size  $N \geq 4((n-1)w_{\max} + 1)$  computes all extremal points of the Pareto front in an expected number of  $O(m^2 n w_{\max} \log(n w_{\max}))$  iterations, where  $n$  is the number of vertices,  $m$  the number of edges, and  $w_{\max}$  is the maximum edge weight in the problem instance. This result confirms, via mathematical means, the good performance of the NSGA-II observed empirically. It also paves the way for analyses of the NSGA-II on complex combinatorial optimization problems.

As a side result, we also obtain a new analysis of the performance of the GSEMO algorithm on the bi-objective minimum spanning tree problem, which improves the previous best result by a factor of  $|F|$ , the number of points in the convex hull of the Pareto front, a set that can be as large as  $n w_{\max}$ . The main reason for this improvement is our observation that both algorithms find the different extremal points in parallel rather than sequentially, as assumed in the previous proofs.

This paper for the Hot-off-the-Press track at GECCO 2024 summarizes the work *Sacha Cerf, Benjamin Doerr, Benjamin Hebras, Jakob Kahane, and Simon Wietheger. 2023. The first proven performance guarantees for the Non-Dominated Sorting Genetic Algorithm II (NSGA-II) on a combinatorial optimization problem. In International Joint Conference on Artificial Intelligence, IJCAI 2023. ijcai.org, 5522–5530 [1].*

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).  
GECCO '24 Companion, July 14–18, 2024, Melbourne, VIC, Australia  
© 2024 Copyright held by the owner/author(s).  
ACM ISBN 979-8-4007-0495-6/24/07.  
<https://doi.org/10.1145/3638530.3664080>

## CCS CONCEPTS

• **Theory of computation** → **Theory and algorithms for application domains**; *Theory of randomized search heuristics.*

## KEYWORDS

NSGA-II, runtime analysis, multi-objective optimization, theory.

## ACM Reference Format:

Sacha Cerf, Benjamin Doerr, Benjamin Hebras, Jakob Kahane, and Simon Wietheger. 2024. Hot off the Press: The First Proven Performance Guarantees for the Non-Dominated Sorting Genetic Algorithm II (NSGA-II) on a Combinatorial Optimization Problem. In *Genetic and Evolutionary Computation Conference (GECCO '24 Companion)*, July 14–18, 2024, Melbourne, VIC, Australia. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3638530.3664080>

## SUMMARY OF OUR RESULTS

Many optimization problems consist of several conflicting objectives. In such scenarios, there typically does not exist solution that is optimal regarding all criteria. The most common solution concept therefore is to compute a set of Pareto optima (solutions which cannot be improved in one objective without accepting a worsening in another objective), and let a decision maker select the final solution based on their preference.

Because they naturally work with sets of solutions, evolutionary algorithms (EAs) are a standard approach to multi-objective problems with many successful applications [11]. The by far most prominent multi-objective evolutionary algorithm (MOEA) is the *Non-Dominated Sorting Genetic Algorithm II* (NSGA-II) proposed by Deb, Pratap, Agarwal, and Meyarivan [2] (with over 50,000 citations on Google Scholar).

Despite being very popular, this algorithm is only little understood from a fundamental perspective. In fact, it was only in 2022 that the first mathematical runtime analysis of the NSGA-II was presented [10]. The works that quickly followed analyze the performance of the NSGA-II on simple benchmark problems, mostly multi-objective variants of the ONEMAX, LEADINGONES, and JUMP benchmarks well-studied in the theory of single-objective randomized search heuristics [4, 9]. However, up until now, the NSGA-II was never theoretically studied on a practical optimization problem.

In this work, we propose the first mathematical runtime analysis of the NSGA-II on such a problem, namely the bi-objective minimum spanning tree problem, which is the multi-objective extension of the well-known minimum spanning tree problem (MST). In previous literature, the MST problem has been used to understand how EAs solve combinatorial optimization problems, the first result in this direction being an  $O(m^2 \log(nw_{\max}))$  runtime upper bound, shown in [8], where  $n$  and  $m$  are the number of nodes and edges of the graph, and  $w_{\max}$  is the maximum weight. Using a balanced mutation operator, this can be improved to  $O(mn \log(nw_{\max}))$ .

In the bi-objective variant, edges are assigned a vector of two non-negative weights. The weight of a spanning tree is then defined as the vector sum of the weights of its edges, and the task is to compute the set of all spanning trees such that no other spanning tree has a smaller weight on both coordinates. If multiple such *Pareto optima* have the same weight vector, one is allowed to find only one of them, effectively computing the *Pareto front*: the set of all optimal *weight vectors*.

This problem is NP-complete, but it is possible to compute the *extremal points* (points which are vertices of the polygonal line formed by convex hull of the Pareto front) in polynomial time [6]. The first result on how MOEAs solve this problem is [7]. It shows that the GSEMO algorithm, a multi-objective analogue of the basic  $(1+1)$  EA, computes the extremal points of the Pareto front in an expected number of  $O(m^2 n w_{\min} (|F| + \log(nw_{\max})))$  iterations, where  $F$  denotes the convex hull of the Pareto front and  $w_{\min}$  denotes the minimum between the maximums of the two weight functions. As in the single-objective case, balanced mutation improves the runtime by a factor of  $\Omega(m/n)$ .

Our results are both an extension of this line of work to the more practically used NSGA-II, and an improvement of the results in [7]. On one hand, we establish the first runtime upper bound for the NSGA-II on the bi-objective MST, assuming that the population size  $N$  is sufficiently large. More precisely, we show that if  $N \geq 4((n-1)w_{\min} + 1)$ , then the NSGA-II finds the extremal points of the Pareto front in  $O(Nm^2 \log(nw_{\max}))$  fitness evaluations, and that balanced mutation improves this number by an  $\Omega(m/n)$  factor. On the other hand, we bring the previous runtime bound shown in [7] down to  $O(m^2 n w_{\min} \log(nw_{\max}))$ , removing a factor of  $|F|$ , for which the only known upper bound is  $nw_{\min}$ . This improvement stems from our observation that, contrary to previous analyses, the NSGA-II and the GSEMO find the extremal points in parallel, rather than sequentially. We propose a new technique based on drift analysis that exploits this parallelism to derive our tighter bound. Even though we apply it only to these two algorithms, we believe that this approach is general enough to derive runtime guarantees also for other MOEAs.

**Summary and conclusion:** In this first mathematical runtime analysis of the NSGA-II on a combinatorial optimization problem, we provided a general approach to proving runtime guarantees for MOEAs solving the bi-objective MST problem. As in most theoretical works on randomized search heuristics, our aim was not so much finding the best possible algorithm to solve this problem (for this problem-specific algorithms will usually be superior), but to gain a deeper understanding of an algorithm, here the NSGA-II, by revealing the mechanisms it uses to solve a certain problem. Overall, this work indicates that mathematical runtime analyses

for the NSGA-II are possible also for combinatorial optimization problems. We concentrated on proving performance guarantees at all, but it would be interesting to derive more insights on how to optimally use the NSGA-II on the particular problem (we only saw that balanced mutation is preferable). Such results are possible in principle: previous works on artificial benchmarks, e.g., [5], gave some indications on the right mutation rate. Clearly, runtime analyses of the NSGA-II on other combinatorial optimization problems would also be desirable to put this research direction on a broader basis.

**Subsequent work:** In [3], another multi-objective EA, the MOEA/D, is analyzed on the class of Multi-Objective Minimum Weight Base Problems (MOMWB), which the MOST problem belongs to. These problems consist in minimizing a weight function over the bases of any given matroid. It is shown that the MOEA/D finds the extremal points of any MOMWB in  $O(m^2 L (m \log(m-n) + 3 \log(w_{\max})) + \log n)$  function evaluations, and the entirety of the convex hull  $C$  in  $O(m^2 L (m \log(m-n) + 3 \log(w_{\max}) + |C|^2) + \log n)$ , where  $L = \sqrt{\min(n, m-n)}$ . This work is interesting for at least three reasons. First, results from [8] about the structure of the search space, which are central in our paper (in particular, Lemma 2), are generalized to MOMWB. Hence, further work could be conducted to see if our analysis of the NSGA-II can be adapted to MOMWB. Second, the factor  $L$  comes from a new upper bound on the number of extremal points for any MOMWB. This bound might also be useful to tighten our analysis. Third, they demonstrate the possibility to obtain runtime bounds for finding the full convex hull of the Pareto front, rather than the extremal points.

## REFERENCES

- [1] Sacha Cerf, Benjamin Doerr, Benjamin Hebras, Jakob Kahane, and Simon Wietheger. 2023. The first proven performance guarantees for the Non-Dominated Sorting Genetic Algorithm II (NSGA-II) on a combinatorial optimization problem. In *International Joint Conference on Artificial Intelligence, IJCAI 2023*. ijcai.org, 5522–5530.
- [2] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and T. Meyarivan. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6 (2002), 182–197.
- [3] Anh Viet Do, Aneta Neumann, Frank Neumann, and Andrew M. Sutton. 2023. Rigorous runtime analysis of MOEA/D for solving multi-objective minimum weight base problems. In *Advances in Neural Information Processing Systems, NeurIPS 2023*.
- [4] Benjamin Doerr and Frank Neumann (Eds.). 2020. *Theory of Evolutionary Computation—Recent Developments in Discrete Optimization*. Springer. Also available at [http://www.lix.polytechnique.fr/Labo/Benjamin.Doerr/doerr\\_neumann\\_book.html](http://www.lix.polytechnique.fr/Labo/Benjamin.Doerr/doerr_neumann_book.html).
- [5] Benjamin Doerr and Zhongdi Qu. 2023. A first runtime analysis of the NSGA-II on a multimodal problem. *IEEE Transactions on Evolutionary Computation* 27 (2023), 1288–1297.
- [6] Horst W. Hamacher and Günther Ruhe. 1994. On spanning tree problems with multiple objectives. *Annals of Operations Research* 52 (1994), 209–230.
- [7] Frank Neumann. 2007. Expected runtimes of a simple evolutionary algorithm for the multi-objective minimum spanning tree problem. *European Journal of Operational Research* 181 (2007), 1620–1629.
- [8] Frank Neumann and Ingo Wegener. 2007. Randomized local search, evolutionary algorithms, and the minimum spanning tree problem. *Theoretical Computer Science* 378 (2007), 32–40.
- [9] Frank Neumann and Carsten Witt. 2010. *Bioinspired Computation in Combinatorial Optimization – Algorithms and Their Computational Complexity*. Springer.
- [10] Weijie Zheng, Yufei Liu, and Benjamin Doerr. 2022. A first mathematical runtime analysis of the Non-Dominated Sorting Genetic Algorithm II (NSGA-II). In *Conference on Artificial Intelligence, AAAI 2022*. AAAI Press, 10408–10416.
- [11] Aimin Zhou, Bo-Yang Qu, Hui Li, Shi-Zheng Zhao, Ponnuthurai Nagarathnam Suganthan, and Qingfu Zhang. 2011. Multiobjective evolutionary algorithms: A survey of the state of the art. *Swarm and Evolutionary Computation* 1 (2011), 32–49.