Dear Editor in Chief,

Please find below our response to the associate editor and the reviewers.

We have addressed all comments in a thorough manner, and we take this opportunity to thank the reviewers and the associate editor for their thoughtful remarks. We have also made some additional minor corrections and modifications, in order to improve the quality of the paper. All our modifications can be seen in the submitted PDF file (Perez\_MOMCTS\_diff.pdf), which clearly highlights the differences between the previous version and this submission.

The revised paper is a significant improvement over the last version. We hope the reviewers will agree with our belief that the manuscript is now acceptable for publication.

Yours sincerely,

Diego Perez,

Corresponding Author.

**Response to Associate Editor and to Reviewers**

Note: we have put our responses in red to mark them out clearly. We have left the reviewers' comments exactly as they were. We use **red bold face** for our response within this document, and *red italic face* when quoting snippets of text from the paper.

**Associate Editor**

Comments to the Author

While I do not demand that you perform additional experiments with another MOEA, I think it is important that you clarify the limitations of NSGA-II and avoid giving the impression that it is the state of the art or that all MOEAs have the same characteristics as NSGA-II. It's fair to say that NSGA-II is the most well-known MOEA, but for several types of problems it has now been surpassed by other algorithms, as noted by reviewer 3.

**Done. Thanks for this comment. We have added references to new approaches such as NSGA-III, SMS-MOEA and MOEA/D. Additionally, we have changed the text related to NSGA-II:**

*NSGA-II is known to deliver solutions with very good diversity and convergence to the Pareto-front for 2-objective problems. However, for problems with more than 3 objectives, other approaches such as NSGA-III (Jain 2013), SMS-EMOA (Beume 2007) and MOEA/D (Zhang 2007) are shown to outperform NSGA-II.*

**These references are:**

*Q. Zhang and H. Li, “MOEA/D: A multiobjective evolutionary algorithm based on decomposition”, IEEE Transactions on Evolutionary Computation, vol. 11, no. 6, pp. 712–731, 2007.*

*Nicola Beume and Boris Naujoks and Michael Emmerich, “SMS-EMOA: Multiobjective selection based on dominated hypervolume", European Journal of Operational Research, vol. 181, no. 3, pp. 653-1669, 2007.*

*Himanshu Jain and Kalyanmoy Deb, “An Improved Adaptive Approach for Elitist Nondominated Sorting Genetic Algorithm for Many-Objective Optimization”, In Proceedings of Evolutionary Multi-Criterion Optimization (EMO 2013), Purshouse, RobinC. and Fleming, PeterJ. and Fonseca, CarlosM. and Greco, Salvatore and Shaw, Jane (Eds.), pp. 307-321, Springer, 2013.*

"Due to the very large branching factor of Go, this game is considered the drosophila of Game AI"

This is in, my mind, a somewhat misleading metaphor. The main advantage of the drosophila is that it is easy to study and affect all the myriad phenotypical traits of the fly; Go, on the other hand, has almost no variations at all (only the same game played on different board sizes).

**Done. We have removed this metaphor, highlighting the broad use of this game by researchers in the last years. This now reads as follows:**

*Due to the very large branching factor of the game and the absence of clear heuristics to tackle it, Go became object of study for many researchers. MCTS was the first algorithm able to reach professional level play in the reduced board size version (Lee, 2009).*

**Reviewer: 1**

Comments to the Author

1. Table's caption should be on the topic of the table.

**We have checked that all captions are in relation with the data shown in the tables across the paper.**

2. The authors should make an experiment to campare with other approaches.

**Following the associate editor comments, no additional experiments have been performed. We agree with this reviewer that a comparison with more approaches would be interesting, although we believe that we have incorporated those that are the most relevant to this study.**

3. The authors should give their motivation and contribution in the revised version.

**Regarding motivation, we believe the first three paragraphs of the paper define it properly. Regarding the contribution, we have added the following sentence at the end of the sixth paragraph (section I), in order to make it clearer:**

*Therefore, the main contribution of this paper is the proposal of a multi-objective version of MCTS that is applicable to real-time domains and is able to provide different solutions across the multi-objective spectrum.*

4. The authors should cite much newer related journal papers.

**Done. We have added three new papers:**

*Himanshu Jain and Kalyanmoy Deb, “An Improved Adaptive Approach for Elitist Nondominated Sorting Genetic Algorithm for Many-Objective Optimization”, In Proceedings of Evolutionary Multi-Criterion Optimization (EMO 2013), Purshouse, RobinC. and Fleming, PeterJ. and Fonseca, CarlosM. and Greco, Salvatore and Shaw, Jane (Eds.), pp. 307-321, Springer, 2013.*

*Q. Zhang and H. Li, “MOEA/D: A multiobjective evolutionary algorithm based on decomposition”, IEEE Transactions on Evolutionary Computation, vol. 11, no. 6, pp. 712–731, 2007.*

***Nicola Beume and Boris Naujoks and Michael Emmerich, “SMS-EMOA: Multiobjective selection based on dominated hypervolume", European Journal of Operational Research, vol. 181, no. 3, pp. 653-1669, 2007.***

5. The authors should mention what is the difference between the proposed method and the published multi-objective optimization.

**We believe this reviewer refers to Wang's approach. Both approaches (especially ours) are described in depth in the paper. These two approaches are quite different, apart from using Pareto fronts at some point during the algorithm. We believe that stating that our algorithm is “influenced by Wang's approach” and that “some modifications are needed” may be a bit misleading, so we have rewritten this paragraph in Section IV, just before explaining our technique in detail:**

*As mentioned before, one of the objectives of this paper is to propose an MO-MCTS algorithm that is suitable for real-time domains. Therefore, an approach different from Wang's is needed, in order to overcome the high computational cost involved in their approach.*

**Reviewer: 2**

Comments to the Author

Abstract: the layout overflows.

**We believe that this is something that only happens in the PDF version from manuscript central (as well as some other font changes in the title and the authors' lines). We have checked that the new PDF submitted does not have this problem.**

Section II:

"MCTS uses Upper Confidence Bound (UCB1, see Equation 1) as a Tree Policy" UCB1 is a popular algorithm, but not the only one used in MCTS. Hence this sentence should be rephrased "Our implementation of MCTS uses Upper Confidence Bound... This tree policy was proposed in: L. Kocsis and C. Szepesvari. 2006 (Bandit based monte-carlo planning).

**Done. This paragraph now reads as follows:**

*[…] it chooses between taking an action that leads to states with the best outcome found so far, and performing a move to go to less explored game states, respectively. In order to achieve this, Kocsis and Szepesvari (Kocsis, 2006) applied Upper Confidence Bound (UCB1, see Equation 1) as a Tree Policy.*

The formalisation of MCTS that is used in this paper was proposed in Chaslot et al 2007 (Progressive strategies for Monte Carlo Tree search), as well as Chaslot 2010 (Monte Carlo Tree Search), one of which could also be cited

**Done. An inaccurate reference was given at the fourth paragraph of section II. We have changed it to cite Chaslot 2007, “Progressive Strategies for Monte-Carlo Tree Search”**.

"as it balances both facets of the search when the rewards are normalized between 0 and 1" This balance is application dependant. In Go for instance, coefficients much lower are used. Hence, the sentence should be reformulated to make clear that this coefficient is application dependant.

**Done. We have stressed that sqrt(2) is a typical value for single-player games, and added a note about its dependence to the game employed. This sentence reads as follows now:**

*[…] A commonly used value for single-player games is sqrt(2), as it balances both facets of the search when the rewards are normalized between 0 and 1. The value of C is application dependant, and it may vary from game to game.*

"A clear example" -> 'Clear' should be omited.

**Done.**

You mention SameGame for monte carlo tree search with a reference from 2010. Monte Carlo Tree Search was used in the game of SameGame by Schadd et al 2008 "Single-Player Monte-Carlo Tree Search", please also mention this reference.

**Done. We are now citing Schadd et al 2008 instead.**

Similarly, MCTS was applied to Morpion Solitaire first in "Reflexive Monte-Carlo Search", T. Cazenave. CGW 2007, pp. 165-173", please include this reference too.

**Done.**

Citation 31: what are the authors?

**The dash line indicates that the authors are the same as the ones from the previous reference. We believe this is an accepted IEEE reference format.**

Page 8, top of the page:

"1) Macro-actions for MO-PTSP

2) Heuristics for MO-PTSP:"

Should there be a title before? Those two sections seems to come out of the blue.

**Done. These are subsubsections from V.B. Similar sections were also included in V.A. In both cases, these subsections described heuristics applied to each game. We have reorganized the subsections of this section to make things clearer:**

*V.A Deep Sea Treasure*

*V.B Heuristics for DST*

*V.C Multi-objective PTSP*

*V.D Heuristics for MO-PTSP*

If possible, please put Table II in the previous page, so the results shown relate with the text.

**We have re-organized figure 9 and tables I-III, so they show up in the correct order. The distances between each table and the texts that refer to them have been minimized as much as possible.**

**Reviewer: 3**

Comments to the Author

First of all, I am not too happy about the choice of the NSGA-II as comparison algorithm. It is true that many people still use it, but cannot really state that it is state-of-the-art. It may still be good for comparisons in 2 objectives, but allready in 3 objectives, it is well-known for years that other multi-objective algorithms are better (there are several newer developments, one of these that is especially near to the NSGA-II concerning implementation issues is the SMS-EMOA that is just more greedy).

**Done. You are right. We added citations to the NSGA-III, SMS-EMOA and MOEA/D. They are known in the literature to outperform NSGA-II, particularly for problems with more than 3 objectives. However, for 3-objective problems, NSGA-II is more or less outperformed in terms of the diversity of the solutions. For our 3-Objective MO-PTSP problem, the spread of our solutions along the Pareto-front doesn’t play a critical role and we indeed obtain desirable solutions in our approximated Pareto-front. Nevertheless, NSGA-II can be replaced by any other multi-objective optimization method for future studies.**

Section III (the MO section) uses several simplifications, this could surely be improved. According to my knowledge, the generated front an MO-algorithm provides is usually called Pareto front approximation, especially if the real front is not known. Only the real front itself is called Pareto front.

**Done. On Page 3, we refer to the true Pareto-front. All over the paper we deal with the approximated Pareto-front or the non-dominated front.**

And if 2 solutions are not dominating each other, they are incomparable, aren't them?

**Done. We changed this to the standard notation:**

*In this case, it is said that they are indifferent to each other.*

Next, the Elitism description on page 3 says that the first front is always kept. But what if the first front gets larger than the population size? Shouldn't that be possible when the number of offspring individuals is equal to the population size? Don't we still need to remove some individuals?

**Done. Yes, it is possible and contemplated by the algorithm. Every new generation, the population is truncated to its defined size. The individuals are sorted according to their rank (first) and their crowding distance (second). For a population of size N, only the first N are kept, while the others are discarded. We have incorporated the following changes to make this clearer: first, at the end of the paragraph before the three bullet points:**

*[…] they are ranked according to a dominance criterion and a crowding distance measure, which are used to maintain a good diversity of solutions.* *After each iteration of the algorithm, only the best N individuals according to this ranking are maintained to the next generation.*

**Secondly, we have reworded the “Elitism” bullet point:**

*Elitism: The individuals from the first front F0 are always ranked first, according to the dominance criterion and their crowding distance. This ensures that the best solutions always survive to the next generation.*

In the beginning of the right col of page 5, I am not sure if the case is considered that the new node r is possibly dominating parts of the front (so that previously found non-dominated solutions would have to be cut out). Is it? Or isn't it possible that this happens?

**Done. Yes, the algorithm contemplates this possibility. This paragraph has been modified and it now reads as follows:**

*Here, if r is not dominated by the local Pareto front, it is added to the front and r is propagated to its parent. It may also happen that r dominates all or some solutions of the local front, in which case the new solution is included and the dominated solutions are removed from the Pareto front. If r is dominated by the node's Pareto front, the local Pareto front does not change and there is no need to maintain this propagation up the tree.*

Concerning the experimental part: on page 9 you use the term "experiment" as a synonym for "runs" or "repeats". That is somewhat confusing, to me an experiment is a higher level concept, and you partly use the term as such. Just repeating for statistical reasons is not really a new experiment.

**Done. Two modifications have been made. In the second paragraph of Section VI:**

*100 experiments have been run for each pair --> 100 runs have been performed for each pair*

**Also, in Section VI.B, fifth paragraph:**

*the same number of experiments is run for --> the same number of runs is performed for*

In table I, you imply that something is significantly better than something else in a statistical sense. How do you test that (details)?

**Done. We have changed this table to reflect the results in a better way. We indicate the average (and the standard error) and those results that are significant (p-value<0.01) are in bold. The comparison is done between the results from the objective with the highest weight and the results from the default weight vector (0.33,0.33,0.33). We have changed the caption of this table, and it now reads as follows:**

*MO-PTSP averages (plus standard error) with different weight vectors. Results in bold obtained an independent t-test p-value <0.01.*

**And also the main text has also been changed to explain this better:**

*Table I shows the results of executing the MO-MCTS controller during 30 games in each one of the 10 maps of the MO-PTSP, for every weight vector. This table includes the averages and standard errors obtained. Results in bold are those with an independent t-test p-value smaller than 0.01, comparing the results of the objective with the highest weight with the result from the default weight vector W1 = (0.33,0.33,0.33) in the same map. It can be seen that the highest weight in W leads, in most of the cases, to the best solution in that objective in the map.*

The comparison between 3 algorithms (using U-tests) is done pairwise, right? There are means to do that differently (all at once), but as it is only 3, this may be acceptable. But your concept of a domination test is not entirely clear to me. You do a pairwise comparison between the repeats for the 2 algorithms and count how often the solution obtained by one algorithm dominate the one of the second algorithm? The description is a bit fuzzy here.

**Done. We have enhanced this explanation for clarity:**

*Extending this comparison to all maps, each pair of algorithms ends with a triplet (D,∅,d). D is the number of maps where the first algorithm dominates the second. ∅ is the amount of maps where no dominance can be established, either because the Mann-Whitney-Wilcoxon non-parametric test failed, or because there is no dominance according to the dominance rules described in Section III. Finally, d states the number of maps where the first is dominated by the second. For example, a triplet (D,∅,d) = (8,2,0) comparing algorithms A and B would mean that the results obtained by A dominate those from B in 8 of the 10 maps, and that it is not possible to derive any dominance in the other 2. Table II summarizes these results for all the algorithms tested.*

Visualization: numbers are nice, but I imagine that especially table II would make a nice figure. Probably much easier to understand.

**We initially attempted to put this information into a figure, but it proved to be not easy to convey. After all, we are comparing four different approaches in a pairwise manner. For each comparison, there are four different weight vectors to check and, for each one of them, three numbers to show. This could be achieved with a 3D bars plot, but the number of bars to draw is too large. We believe the current table shows, in a clear snapshot, all the different comparisons made. We also believe that the most important piece of information is present and easy to understand: the number of maps where the approaches dominate each other.**

The subsection VI/C is a bit different, it somehow appears when I was expecting some conclusions. I am not sure if it is really well placed here as it just touches on interesting extensions but does not really cover them.

**We have left this section as it was. We believe its place in the paper is appropriate, after the MO-PTSP results discussion, as this is an extension of the research done in this game. We agree with the reviewer that this subsection is a little bit different from what is explained in the rest of the section, but we still believe it makes an interesting contribution, as it also motivates future work on this topic.**

Minor things:

- figure 2: the picture has an embedded copyright notice. I am not sure if this is ok with the publisher, maybe it should be put into the caption? It is very hard to read anyway.

**Done. The copyright notice has been removed from the figure, and it is now cited as a reference:**

*Brockhoff, D. Tutorial on Evolutionary Multiobjective Optimization. Proceedings of the 15th Annual Conference Companion on Genetic and Evolutionary Computation, 2013, 307-334*

- page 3, left col: litrature -> literature

**Done.**

- hardware named in the beginning of VI (experimentation): you really have a 230GHz server? Where can I order that? :-)

**Done. The actual value is 2.90GHz.**