Answers to the comments of the reviewers  
Reviewer 1:

The advice in this paper is important for practitioners who do not know better than to keep the order of runs provided by the coordinate exchange algorithm.

That the order of runs produced by the algorithm is not random is certainly known to the developer of the algorithm in JMP, because the default run order is to randomize. Of course, the user may choose to override this ordering. Your method for proving that the order provided by the algorithm is not random is clever. I note that you leave out the x2^2 term in your model to force the number of level combinations to be only 6.

You should mention that the algorithm could be altered to generate random orderings by randomly choosing the order of the treatment of the rows in the starting design instead of starting with the first row and moving sequentially through the rows. However, this approach would be somewhat wasteful of resources, since it is simple to randomize the run order after the fact.

Thank you for your supportive comments and the useful suggestion about changing the order in which the rows are considered. This would indeed make the outcome random again. In the revised version of our paper, we mention this alternative solution in addition to our original solution of permutating the run order after optimization.

Minor Edit:

Page 5 first full paragraph line 5 - change "extremely rarely" to "rarely"

Done.

Reviewer 2:

This is a very short note where the authors check the randomness of the order of  
the D-optimal design points obtained using a coordinate-exchange algorithm. This is  
checked in a particular example with two factors with 3 and 2 levels respectively. It  
turns out that after running the algorithm a number of times there is not a random  
ordering, what suggest the need of performing always a randomization of the points  
after the optimal design is computed with this algorithm. The idea is very interesting and relevant. The paper is correct and well written.

Thank you for this positive evaluation of our work.

My main concern is whether the topic without a real example related to the journal  
policy does actually fit here. Moreover, the note is really short and considers a  
very particular case with a particular criterion and a particular algorithm. A more  
profound study would be interesting using more examples in different scenarios.

We have specifically chosen to submit this work as a short communication, one of the types of contributions Chemometrics and Intelligent Laboratory Systems publishes. We do not believe that we can flesh the results out to a full-length publication without becoming too repetitious. Analyzing the technical details why the algorithm does not produce a random outcome might also be of lesser interest to the readership of Chemometrics and Intelligent Laboratory Systems. Also, while our work is certainly interesting, it is perhaps not important enough for a full paper on the subject. Our goals are to show with an easy to understand example that the commonly used coordinate exchange algorithm does not produce random run orders and to recommend practitioners to perform a randomization of the run order after the design has been produced.

Just D-optimality is considered, where a saturated optimal design has equal  
weights in all the design points. I wonder whether other criteria giving different  
weights for an optimal approximate design would give some light about this unbalance ordering in the algorithm.

We have added a short paragraph in the revised manuscript mentioning that our conclusion is valid for A-optimal designs as well.

The pseudo-code in Figure 1 should not refer to a particular sentence as fmincon from a particular language. Moreover, I believe this can be simplified and even made  
more clear.

We have removed the Matlab specific use of the term “fmincon” and simplified the pseudo-code in Figure 1.

Summarizing, in my opinion more examples and criteria would be needed to  
strengthen the conclusions.

See above.

## Reviewer 3:

This paper highlights an undesirable property of the coordinate exchange algorithm which is used in the construction of optimal experimental designs.

Thank you for your comments.

1. After reading this paper I constructed a D-optimal design in JMP (Pro 14.0.0) software  
and the default option is for the experimental runs to be randomized. Is this not an  
option (or obvious option) in other popular design software? Like SAS? Design Expert?  
It might be useful to Chemometrics readers to know which software randomizes the runs  
and which software does not.

SAS proc optex does not randomize the run orders in its standard output, and the same goes for Design Expert. We hesitate to mention this explicitly in the paper, because these implementation details differ from version to version in the software packages.

2. Does design size matter? For example, is the lack of randomization more of a concern  
for smaller designs than for larger?

Except for the smallest designs for the simplest models, the non-randomness always occurs in our experience. Only for a trivial example like y = ax + b with 2 runs, the design +1 -1 is found as often as the opposite order -1 +1. For more complicated models, the extent to which the orders are not random seems to be somewhat model and run size dependent. But analyzing these trends in detail does not seem extremely important to us, as any non-randomness in run order should be avoided, and the solution of randomly permuting the runs is computationally so cheap that there is no reason not to do it.

3. Please define *fmincon* in Figure 1 in the caption or when you reference the figure. Most  
readers will probably not be familiar with this MATLAB function.

We have removed the Matlab specific use of the term “fmincon” from figure 1.

4. On that note, do all coordinate exchange algorithm use an interior point optimization  
method? Will using a different optimization method change your findings? I’m sure not  
all software the implementation of coordinate exchange algorithms vary across software  
packages.

The original coordinate exchange algorithm used grid optimization to optimize the individual coordinates, while modern implementations, such as Ruseckaite et al. (2017) and Huang et al. (2019), also use an interior point method. We cannot be sure about the details of commercial software packages, as these do not give such detailed descriptions of their implementations. The non-randomness remains when grid optimization is used, since it is due to the sequential order in which the coordinates are optimized.

5. I’m curious as to the motivation behind this work. How was this property discovered?

The non-randomness of the algorithm was verified by repeatedly generating a design for a fixed model and run order. This is indeed not something you would normally do, except specifically for verifying this property. We, however, noted when designing experiments for widely different models and run numbers that design points in the corners were often listed in the first positions of the run order. This led us to hypothesize that the run order might not be random.

Specific comments:  
1. Pg 2, second paragraph, should be “completely randomized experiments” not  
“experiment”.

Fixed

2. Pg 3, first sentence, remove “of the experiment to be conducted”.

Done

3. Pg 4, top sentence, remove “any more”.

Done

4. Pg 4, first paragraph, regarding the statement “but it is recommended in design of  
experiment textbooks to carry out all tests in a random order.” Yes, but this is a  
fundamental concept in design of experiments, it is not just recommended. I would use  
stronger language here to convey how important your findings are.

We have added to the text that the randomization is crucial to justify the traditional analyses of the experimental data and referred to an interesting paper by Cox (2009). This is substantially stronger than our original formulation.

5. Pg 7, first paragraph, should be “of 96 run orders which are

Done.

## Reviewer 4:

The content of this **short communication** is of interest to users of computer generated designs. The need to randomize the design obtained by the exchange algorithm is well justified. References are appropriate; the text is clearly worded and easy to read.

Thank you for your positive evaluation.

However, the content does not meet exactly the scope of the journal: “Chemometrics and Intelligent Laboratory Systems publishes original research papers, short communications, reviews, tutorials and software descriptions reporting on novel developments in techniques for chemistry and related disciplines that are characterised by the application of statistical and computer methods.”  This short communication does not include any application to the chemical field. But I think its content is of interest regardless of any application. In my opinion it should be published in Chemolab, but obviously my interpretation of  the scope of the journal  may not be shared by the editor.

Design of experiments is a key topic in many issues of Chemometrics and Intelligent Laboratory Systems. We specifically chose to try to publish our result as an easy to understand communication in this journal because it is valuable to experimental design practitioners, which are numerous in chemistry and chemometrics laboratories.

 Reviewer 5:  
In this short communication, the authors described how the well-known coordinate-exchange algorithm works. They presented the main features of the algorithm, than they pointed out a possible limitation. They showed that the order of the design points produced by the coordinate-exchange algorithm is not random as expected. An example is used to demonstrate the thesis, which is studied through a computational experiment that involves a large number of executions of the algorithm.

Thank you for your comments.

 In what follows, three main comments about what could be help the authors to improve the main quality of the communication, which is well written and easy to read.

 -      The computational experiment is based on a quadratic model in which the factor level combination orders does not impact the optimality of the experimental design. The example is a nice tool to demonstrate the main issue pointed out by the authors.  However, it would be interesting to investigate how this absence of randomness impacts the design matrix quality when optimality changes. The magnitude of the impact can substantially change the importance of this contribution.

When other optimality criteria are used for designing completely randomized experiments, the same problem occurs. When preparing the revision of our original paper, we also studied the results for A-optimality instead of D-optimality and observed the same kind of non-randomness of the output.

-      The coordinate-exchange algorithm is based on an initial random generation of feasible designs.  Usually, the random number generator has effects on final performance of an optimization algorithm. While comparing algorithms is strongly recommended to set the seed to a unique value. This is not the case but some comments about the impact of the initial random generation of feasible designs phase could be useful and appreciated.

We have used the standard number generator in Matlab, a Mersenne twister, and have added this detail to the paper. As mentioned in the comments for reviewers 1 and 3, the non-random order of the coordinate-exchange algorithm’s output is due the sequential consideration of coordinates. To investigate whether the random number generator affects the randomness of the output, we have also tested the multiplicative congruential generator that Matlab also offers for generating random numbers. This also led to a very small p-value, rejecting the hypothesis that the coordinate exchange algorithm provides random run orders.

-      Lastly, a more complete analysis of the impact from a practical point of view could improve the quality of the communication. Expect for the violation of the random hypothesis, which other limitations and issues can derive from the main finding presented in the communication? In addition, it could be interesting to add some comments about the relation between chemistry, experimental design and the issue pointed out by the authors. This could be also important to better fit the aims and scope of the journal.

In the revised manuscript, we have added a reference to Cox (2009), which highlights the importance of randomization. We do not want to iterate all the arguments from that paper in our work, because we submitted our work as a short communication.

Design of experiments is a key topic in many issues of Chemometrics and Intelligent Laboratory Systems. This is witnessed by several papers we cite. Also, famous handbooks on chemometrics spend many pages on experimental design. Therefore, we believe that the relation between chemistry and experimental design should not be discussed in more detail, also because we submitted our paper as a short communication.

Just a minor comment about page 2. The “parameters” n and m are correctly presented in the right position. I would suggest introducing the meaning of fmincon in the same paragraph even if it is explained later on the paper. Maybe a introducing sentence will be enough.

Other reviewers have also pointed out that using Matlab specific function “fmincon” hinders the readability of Figure 1. We have rephrased things to prevent this from happening, and now only mention “fmincon” as a technical detail necessary for reproducibility of the work.

## Reviewer 6

In this paper, the authors check the outcomes of the coordinate exchange algorithm, when it is used for the construction of fully randomized designs. The results clearly showed a pattern on the outcomes of the algorithm, with respect to the sequence of the runs.

 The main contribution of this paper is to convince the readers for the necessity of randomisation every time before running an experiment. Although this result is quite important, several points need further clarification.

Thank you for your comments.

 Firstly, the authors mention that this algorithm is being used by JMP and Design Expert. I would have believed that the randomise order that the packages give in the end, it is an extra randomization and not the direct outcome of the coordinate exchange algorithm. If this is not true, this should be made more clear in the paper.

We cannot be sure of the randomness of the designs generated by commercial packages such as JMP and Design Expert, as they do not fully describe how their algorithm and user interface work. In JMP, the generated designs, as they first appear in the output, are not random, but they are randomized as soon as the user hits the “Make Table” button and does not switch off the default randomization option. To the best of our knowledge, other software packages merely report the algorithm’s output, without explicit randomization.

In the same sense, I believe the authors should motivate better their work. Do they know cases, in which the experimenter uses directly the outcome of the coordinate-exchange algorithm without a proper randomisation?

In the second paragraph of our work we have mentioned multiple studies that used the coordinate exchange algorithm. None of them mentioned whether they randomized the output of the algorithm. One logical interpretation of that is that they did not randomize.

 Moreover, there are different versions of the coordinate-exchange algorithm. In the paper, do the authors choose to change each coordinate by column? The coordinates could change by rows or even randomly. Does the observed phenomenon appear in all of these versions of the coordinate exchange algorithm?

The non-random order is introduced by the systematic sequential optimization of the coordinates. Optimizing the coordinates in random order, as reviewer 1 suggested, is the only way in which we can prevent the non-randomness from occurring.

# References

Ruseckaite, Aiste, Peter Goos, and Dennis Fok. "Bayesian D‐optimal choice designs for mixtures." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 66.2 (2017): 363-386.

Huang, Yuanzhi, et al. "Optimal design of experiments for non‐linear response surface models." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 68.3 (2019): 623-640.