# 3rd report on “Interpolation of High-Dimensional Data”

by Thomas C. H. Lux et al.

The manuscript has greatly improved from the last version. The authors have adequately answered almost all of my concerns, so I would suggest that the paper be accepted after a last round of minor revisions, to fix some small details left. I list them here below.

1. page 2: ref 36 is not about sparse grids, but rather about providing convergence estimates for high-dimensional regression. Please move it a few lines below, next to the sentence “similar uniform bounds cannot be constructed for regressors in general”.

2. page 2: the sentence “However, sparse grids are not studied here because they have exponential size in dimension” is still wrong. Sparse grids are precisely built so that the size in dimension grows less than exponentially. What you can say is something milder, like “However, sparse grids are not studied here because their size still grows quite fast with dimensions (although less than exponentially) making them not readily tractable for more than a few tens of dimensions, and [...]”.

3. MARS equations at page 4: what is the relation between $2j-1$ and $l$, and similarly between $2j$ and $k$? Also: each iteration considers one new point $x^{(p)}$, but one has more points than basis/iterations ($n > m$): how do you choose $p$ at each iteration? The formula has a subscript in $x_i^{(p)}$, but it is also unclear why one should consider only one coordinate of $x^{(p)}$, could it be that there is a typo?

4. Add reference to Figures 1,2,3 in the text.

5. Paragraph starting with “A hurdle when modeling”: the sentence “In this example the KS test null hypothesis is rejected at p-value 0.01, however it is not rejected at p-value 0.001.” is out of context here. Move it to the paragraph starting with “The two-sample KS test is”.

6. After the formula $KS > \sqrt{\ldots}$, mention that $n_1, n_2$ do not refer to the number of points used for function interpolation, but to the numerosity of the samples used to compute the empirical CDFs that are being compared by the KS test. Also, use $\mathbb{N}$, not $\mathcal{N}$.

7. page 13, “THE FACT That the theoretical error bound ...”.

8. Figure 20: the scatter plot for the forest fire dataset (top-left) looks weird, the clouds of point are not even slightly aligned to the bisector of the $X-Y$ plan. If this was correct, the interpolation methods would be working terribly bad in this example. Please clarify.

9. Appendix: saying that the p-value is “the probability of observing a given statistic if the null hypothesis is true” is imprecise. I’d rather say “the probability that other sets of data would produce a value of the given statistic at least as extreme as what obtained with the current data, if the null hypothesis were true”. I would also add the following comment, which could benefit a reader with little familiarity with the subject: “In other words, the p-value is a way of measuring the “likelihood” of a set of data assuming that the null hypothesis is true: a small p-value indicates a strong statistical evidence that the null hypothesis can be rejected”.