General Comments

I thought the comments in between lemmas seemed to do a reasonable job conveying the ideas. In general, I felt like this relies on math that I'm not really familiar with, so I'm not sure how I'm missing because of that. I do think it would be helpful to have more of a plain English roadmap of the key ideas.

What I seemed to get from this section of the paper is that the estimation procedure calculates a path based on an assumed dimension. Then if that dimension is correct, the path will definitely be small, and if it's not correct the path will probably be large. The paper then proceeds to work out what the cutoff should be between "small" and "large" and what exactly "probably large" means. I think this kind of high level explanation, if it's correct, would be helpful to a reader. That way even if it takes me a lot of work to follow the math, I can quickly get a big picture of how the estimator works.

One other thing that was mentioned in the math-writing notes Brian posted: it may be helpful to give each lemma a name. Instead of saying "Lemma 2", maybe using a name that encapsulates what the lemma showed would make it easier on the reader to follow the argument without having to flip pages to look back.

Specific Comments

Section 3, Paragraph 2

> "the intrinsic dimension of data is assumed to be either d_1 or d_2." Are d_1 and d_2 previously defined? Or are we restricting to a case where the dimension is one of two known (but, for the purposes of argument, arbitrary) values? I think this needs clarification. Reading on, it appears it's this second case, but I think it would be better to say this immediately so that the reader doesn't pause due to confusion.

Paragraph 3, line 4

> I think the sentence transition here could be improved. In the previous sentence you define the estimator. You could use this by saying something like "This estimator $\hat{dim}_n$ is always correct when the intrinsic dimension is d_1. We will show this in Proposition 7. If the intrinsic dimension is d_2, then ..."

Paragraph 4

> Here it says that Lemma 5 gives a bound on the probability of error when the intrinsic dimension is d_2, but in the previous paragraph it says this is shown in Proposition 7. I found this confusing. The second sentence ends by mentioning a threshold L. It's not clear to me what this is referring to. Are you saying that the length of the path can be arbitrarily large with high probability?

Paragraph 5, line 4

> It says when $d_1 = 1$, the lemma is straightforward, but the ensuing explanation still struck me as rather technically complex. Depending on the expectations of the readership, maybe this could be made more "intuitive".