Review of "Faithful variable screening for high-dimensional convex regression"

I am overall quite satisfied with the authors' work on the revised manuscript. The authors seem to have addressed the salient points raised in the previous round of AOS reviews. In particular, the additional discussion regarding the boundary flatness condition enhances the reader's understanding of the assumption made in the paper.

A few minor comments:

(i) In Section 2.3, the authors preview their main results concerning variable selection consistency in a finite-sample setting. However, at this point, the "true regression function" f_0 has not yet been defined (it is defined later in Section 5.2). It would be helpful if the authors were to state the model generating the data in this section.

(ii) In Example 3.3, the authors describe a method for approximating any bounded density over a hypercube arbitrarily well using boundary flat densities. The second paragraph mentions truncating the density to the hypercube $[\epsilon, 1-\epsilon]^p$—since the truncated function isn't a density, it should technically also be rescaled.

(iii) In the penultimate paragraph of the proof of Theorem 3.2, $\phi(\mathbf{x}_{-k})$ seems to be redefined incorrectly (it should be $\sum_{k' \neq k} f^*_{k'}(x_{k'})$).

(iv) In the first paragraph of Section 5, the authors state that variable screening consistency means that $\mathbb{P}(\hat{S} = S)$ approaches 0. It should approach 1.

(v) In Assumption A4 at the bottom of page 25, the moment generating function should read $\mathbb{E}e^{tW}$ rather than $\mathbb{E}e^{t\epsilon}$. There is some change in notation in Section 6.1, where W is replaced by ϵ; perhaps it would be better to keep the notation consistent.

(vi) The last sentence of the first paragraph on page 26 is a bit hard to follow. What does "if X is independent" mean (does this mean $p(\mathbf{x})$ is a product distribution)? I did not understand where the simplification came from.

(vii) The $\nu = 0.9$ curve in Figure 5(d) is rather ill-behaved compared to the other curves. However, if my understanding is correct, the probability of successful screening should still tend to 1 in this case. Perhaps the authors could consider plotting more samples to show eventual consistency for this curve, as well (and also take an average over more trials, since the behavior of the curve is somewhat noisy).

Thanks to the authors for their detailed attention to the points raised by the reviewers. Modulo the remaining small edits, the paper seems suitable to be published in AOS. In this reviewer's opinion, it provides a valuable contribution to the literature on shape-constrained regression.