Comments on "On the Wilks phenomenon of Bayes factors and the integrated likelihood ratio test"

The authors investigated the asymptotic null distribution of generalized fractional Bayes factors as an alternative of LRTs. The paper is clearly written, and the meat of the paper is well-explained. **However, this work has several limitations to be accepted by a reputable journal like JRSSB.** This work is somewhat bogus in a sense that it is lacking applications on real problems, and the examples considered by the authors (without simulation studies) are too simple to be practical. They do not consider or examine any realistic situations, but just list asymptotic results. For example, they assume the posterior contracts at $n^{-1/2}$, but this setting is too simple in modern statistics. They should have considered more theoretically interesting situations where the posterior contraction rate is much slower that $n^{-1/2}$ even under optimal settings (some examples are discussed in list number 5 below). Let me list some concerns or comments as follows:

1. Some related reference is missing such as Zhou and Guan (2018) JASA paper "On the Null Distribution of Bayes Factors in Linear Regression". In this paper, they showed by considering the Bayes factor as a test statistics (like this submitted paper), they derived the null distribution of the Bayes factor for linear models, which is a weighted sum of chi-squared random variables. Even though this work is restricted to linear models under simple assumptions, it is worth to cite. Also, I believe that you can find more references that consider a Bayes factors as a test statistics in frequentist settings.

2. On page 4, "The prior $\pi(\theta)$ and $\pi_0(\nu)$ may be improper...". This is wrong in general. If they are improper, we cannot calculate the Bayes factor. Of course, I understand that even when the priors are improper, the fractional Bayes factor, introduced later, circumvents this issue. But, I think that it would be better to briefly note this point at the place of the sentence "The prior $\pi(\theta)$ and $\pi_0(\nu)$ may be improper...", to avoid a confusion.

3. On line 31 on page 6, what is f?

4. On line 38 on page 6, you may want to note an extra weakness of the traditional approach. When the prior density is proportional to a function of the determinant of the Fisher information, its prior normalizing constant is difficult to evaluate, and the normalizing constant is critical for the Bayes factor. Although you noted "...Fisher information matrix has a complicated form... undesirable...", it would be better to be more specific.

5. This theoretical work is investigated under assumed that the posterior achieves $\sqrt{n}$-consistency (or the posterior contraction rate is $n^{-1/2}$). In many practical statistical models, the optimal posterior contraction rate would be slower than $n^{-1/2}$. These examples include high-dimensional sparse problem like $(\log p/n)^{1/2}$ or nonparametric function estimation $n^{-\beta/(2\beta+p)}$, where β is a smoothing factor for the true function. Under these interesting models, can you extend this result to more general models?

6. You proposed theoretical results on asymptotic null distribution of Bayes factor which is a linear transformation of a chi-square distribution. But, you have not considered any simulation works to examine finite sample behavior of the approximated distribution. In practice, the accuracy of the asymptotic null distribution with finite samples is of interest. You may want to show that this asymptotic null distribution is useful in practice, especially for complicated and practical models. The examples you considered (without simulation studies) are too simple and far from a practical point of view.

7. This paper is lack of tuning parameter selection which will be critical in the hypothesis testing result. Theoretically, it can satisfy some simple conditions, but in practice how to choose the tuning parameters a and b is very important.