This paper presents the PARNI MCMC algorithm for variable selection in the big-data-and-many-covariates regime under GLMs and survival models. The design of the algorithm is the same to one presented in an earlier article by the same authors for high-dimensional variable selection under linear models. The novelty is extending the algorithm to the two new model families.

The paper includes a thorough discussion on existing and competing approaches, and presents several variants of the new PARNI algorithm. I'm not an expert on the previous paper, but I don't think there were any particular technical difficulties to be overcome in extending the scope of the PARNI method. The majority of the novelty was subsumed into approximating a marginal likelihood under the new model classes, for which there are standard methods. If I missed something, then it would be worth emphasizing. That being said, the GLMs and survival models are central statistical tools, and generalising to apply to to them can be a worthwhile endeavor.

The efficiency of several PARNI-variants is compared to a more standard benchmark in a number of numerical experiments. The PARNI-methods in question are adaptive or based on sophisticated, informed proposals. In contrast, it looks to me like the benchmark Add-Delete-Swap methods proposed changes to only one covariate at a time, despite an adaptive version changing multiple variables at once being available. A comment about data augmentation schemes being implemented in Rcpp also makes me suspect that the implementations of the methods are not always comparable. Hence, I'm not confident that the comparisons (normalised to equal wallclock time) are always fair, and that at least some of the advantage of the PARNI schemes could be artificial. Indeed, there are several examples where the ADS schemes perform better, though the authors explain that those tend to have flatter posteriors or fewer covariates, perhaps rendering them less interesting.

The simulated example comparing three PARNI schemes examines mixing, but none of the real data examples do. Hence, it could be that the lower MSE exhibited by the PARNI method is to some extent a consequence of it being less effective at exploring the state space. I don't see any positive evidence of that being the case, but I think it should be checked. An $\hat{R}$-statistic or a multivariate ESS are usually fairly good ways to assess convergence and mixing, in my experience.

I don't think the paper contains a description of code availability, and I didn't see a supplemental file to that effect either. I think open code is absolutely essential for a publication like this. It would not only render the work reproducible (which it currently isn't), but also make it possible for practitioners to use the methodology in their own work, at least provided that the implementation was reasonably straightforward to customize (which it should be). I think an open software resource for practitioners would more than make up for the somewhat incremental methodological contribution in this paper.