Bayesian Analysis (2026) **TBA**, Number TBA, pp. 1

Contributed Discussion on "Robust Probabilistic Inference via a Constrained Transport Metric" by A. Chakraborty, A. Bhattacharya, and D. Pati

Francesco Gaffi* and Beatrice Franzolini†

We congratulate the authors for an original and stimulating contribution. The proposed D-BETEL framework offers an elegant and effective compromise between robustness and interpretability, where inference remains focused on a finite-dimensional parameter θ, indexing an interpretable parametric model F_θ, while still avoiding a full specification of the data-generating distribution. The procedure is particularly appealing in that inference results from a prior, specified only on the parameter of interest θ, and a modified likelihood contribution, obtained by finding the maximum-entropy reweighting of the empirical distribution of the observations that lies within a neighborhood of F_θ. Hence, values of θ are favored when this distance-based compatibility with F_θ can be achieved with little departure from the uniform empirical weights. In this sense, the modified likelihood rewards models F_θ that explain the bulk of the data while still allowing atypical observations to be downweighted. We view the nonparametric Bayesian interpretation in Section 5, which is shown to arise in an appropriate asymptotic regime, as a distinctive strength of this approach. Unlike many alternative pseudo-likelihood or generalized Bayes approaches, D-BETEL is shown to arise as a limiting marginal posterior under a suitable nonparametric hierarchical construction, thereby providing a genuine probabilistic motivation for its use.

A feature we find especially interesting is that the proposed approach allows the prior distribution to be defined directly on the finite-dimensional parameter of interest θ, while still encompassing data-generating processes more general than the parametric family F_θ. This aspect brings the proposal into close contact with recent developments in Bayesian nonparametrics, where prior information is incorporated by specifying the distribution of selected functionals of a random probability measure. In Gaffi et al. (2025), the emphasis is on prior elicitation within a given class: once this class has been chosen, the base measure is selected so that a linear functional, typically the mean, has a prescribed distribution, while preserving the structure that allows one to rely on standard inferential techniques for that class. In Kessler et al. (2015), a complementary strategy is adopted: a canonical nonparametric prior is modified so that the induced marginal distribution of a functional matches a desired one, resulting in a prior that generally lies outside the original class and requires adapted posterior sampling methods. In both cases, the underlying idea is that prior information is more naturally expressed in terms of functionals than in terms of the full infinite-dimensional object.

*University of Bergamo, Italy francesco.gaffi@unibg.it
†University of Milano-Bicocca, Italy beatrice.franzolini@unimib.it

D-BETEL differs from these approaches in how this idea is implemented, but appears to be driven by a closely related principle. Rather than specifying or modifying a prior so as to match the distribution of a functional, it introduces a constraint through a discrepancy functional,

$$Q \mapsto D(F_\theta, Q),$$

and selects admissible distributions accordingly. This functional is nonlinear and defined through an optimization over couplings, and therefore lies outside the class of functionals typically considered in the literature. The distinction is not only formal. Even in relatively simple settings, moving from linear to nonlinear functionals entails a substantial increase in complexity: for instance, quadratic functionals of random objects driven by completely random measures already require nontrivial tools based on Wiener–Itô decompositions and asymptotic analysis, as shown in Peccati and Prünster (2008), and explicit distributional control is generally difficult to obtain. In this sense, D-BETEL can be interpreted as extending the functional-based perspective toward functionals that capture global features of the distribution, such as overall shape and tail behavior.

At the same time, the analogy remains structurally strong. As in the approaches mentioned above, the infinite-dimensional problem is effectively controlled through a lower-dimensional quantity, and prior information is incorporated indirectly through the behavior of a functional. The nonparametric Bayesian representation provided in the paper further clarifies this point: the procedure can be embedded in a hierarchical model involving a random mixing distribution, although this equivalence holds in an asymptotic regime where nuisance parameters are marginalized out. In practice, the method avoids working with the infinite-dimensional object altogether, and the induced prior on it is only implicit, but the overall mechanism remains closely aligned with a functional specification viewpoint.

This perspective suggests a natural extension. The constraint $D(F_\theta, Q) \le \varepsilon$ imposes a sharp boundary on the value of a nonlinear functional. One may instead consider specifying a prior distribution on $D(F_\theta, \tilde{P})$ for a random probability measure $\tilde{P}$, thereby replacing the hard constraint with a probabilistic one. Such a formulation would be closer in spirit to approaches based on fixing the distribution of a functional, while potentially avoiding boundary effects induced by the constraint. It might also provide a route to an exact nonparametric Bayesian representation, rather than one obtained through asymptotic equivalence, by explicitly incorporating prior beliefs on the extent of departure from the centering model.

From this point of view, the contribution of the paper is twofold. On the one hand, it provides a practically effective method for robust inference that bypasses the direct handling of infinite-dimensional parameters. On the other hand, it points toward a broader perspective in which nonlinear functionals, such as transport-based discrepancies, play a central role in linking parametric structure and nonparametric flexibility. While the literature on linear functionals offers useful guidance, extending these ideas to more complex functionals remains largely open. The present work shows that this direction is not only conceptually meaningful but also may admit concrete implementation, and therefore deserves further investigation.

References

Gaffi, F., Lijoi, A., and Prünster, I. (2025). "Random probability measures with fixed mean distributions." *The Annals of Applied Probability*, 35(4): 2239 – 2261.
URL https://doi.org/10.1214/24-AAP2130 1

Kessler, D. C., Hoff, P. D., and Dunson, D. B. (2015). "Marginally Specified Priors for Non-Parametric Bayesian Estimation." *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 77(1): 35–58.
URL https://doi.org/10.1111/rssb.12059 1

Peccati, G. and Prünster, I. (2008). "Linear and quadratic functionals of random hazard rates: An asymptotic analysis." *The Annals of Applied Probability*, 18(5): 1910 – 1943.
URL https://doi.org/10.1214/07-AAP509 2