Addendum for
Maximum likelihood estimation in Gaussian process regression is ill-posed*

Regretfully, our literature review missed an important piece of work by Ben Salem et al. [1]. Consider a product kernel of the form

$$K(x,y) = \prod_{i=1}^{d} \Phi\left(\frac{x_i - y_i}{\lambda_i}\right)$$

with d-dimensional inputs $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$ and positive lengthscales $\lambda_1, \ldots, \lambda_d$. Let $\ell(\theta \mid Y)$ be the modified log-likelihood function in our Equation (2.8). Ben Salem et al. [1, Proposition 4.3] have proved, under assumptions on K and the data that are similar to those in our Theorem 5.3, that

$$\lim_{\lambda_p \to \infty} \ell(\lambda_1, \ldots, \lambda_d \mid Y) = -\infty,$$

a result for which they give the interpretation that "[...] ML can detect inactive input variables and assign them large correlation lengths" (p. 1379). Also compare their Proposition 4.4 on cross-validation to our Remark 3.2. Importantly, their proof (in Appendix A.1) of Proposition 4.3 is, more or less, identical the proof (in Section 7.4) of Equation (2.11) in our article. Note that our Theorem 5.3 is a rather straightforward consequence of Equation (2.11).

References

[1] Ben Salem, Bachoc, Roustant, Gamboa & Tomaso (2019). Gaussian process-based dimension reduction for goal-oriented sequential design. *SIAM/ASA Journal on Uncertainty Quantification*, 7(4):1369–1397.

*Karvonen & Oates (2023). Maximum likelihood estimation in Gaussian process regression is ill-posed. *Journal of Machine Learning Research*, 24(120):1–47. https://jmlr.org/papers/v24/22-1153.html