Universiteit
Leiden
The Netherlands

Scalability and uncertainty of Gaussian processes

Hadji, M.A.

Citation
Hadji, M. A. (2023, January 25). *Scalability and uncertainty of Gaussian processes*. Retrieved from https://hdl.handle.net/1887/3513272

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3513272

Note: To cite this publication please use the final published version (if applicable).

Bibliography

Arbel, J., Gayraud, G., and Rousseau, J. (2013). Bayesian optimal adaptive estimation using a sieve prior. *Scandinavian Journal of Statistics*, 40(3):549–570.

Balder, E., Gilliland, D., and van Houwelingen, J. (1983). On the essential completeness of Bayes empirical Bayes decision rules. *Statistics & Risk Modeling*, 1(4-5):503–510.

Belitser, E. (2017). On coverage and local radial rates of credible sets. *Ann. Statist.*, 45(3):1124–1151.

Belitser, E. and Enikeeva, F. (2008). Empirical Bayesian test of the smoothness. *Math. Methods Statist.*, 17(1):1–18.

Bényi, A. and Oh, T. (2013). The sobolev inequality on the torus revisited. *Publicationes Mathematicae Debrecen*, 83:359–374.

Berlinet, A. and C. Thomas-Agnan, C. (2004). *RKHS and Stochastic Processes*, pages 55–108. Springer US.

Bernardo, J. and Smith, A. (1994). *Bayesian Theory*. John Wiley, New York.

Bhattacharya, A. and Pati, D. (2015). Adaptive Bayesian inference in the Gaussian sequence model using exponential-variance priors. *Statist. Prob. Letters*, 103:100–104.

Bhattacharya, A., Pati, D., and Yang, Y. (2017). Frequentist coverage and sup-norm convergence rate in gaussian process regression. *arXiv e-prints*.

Bickel, P. J. (1982). On Adaptive Estimation. *The Annals of Statistics*, 10(3):647 – 671.

Bickel, P. J. and Doksum, K. A. (1977). *Mathematical Statistics: Basic Ideas and Selected Topics*. Holden-Day Company, Oakland, California.

Box, G. (1976). Science and statistics. *Journal of the American Statistical Association*, 71(356):791–799.

Brown, L. and Low, M. (1996). A constrained risk inequality with applications to nonparametric function estimation. *Ann. Statist.*, 24:2524–2535.

Cai, T. and Low, M. (2004). An adaptation theory for nonparametric confidence intervals. *Ann. Statist.*, 32:1805–1840.

Cao, Y. and Fleet, D. (2014). Generalized product of experts for automaticand principled fusion of gaussian process predictions. *arXiv e-prints*.

Castillo, I. and Nickl, R. (2013). Nonparametric Bernsteinñ-von Mises theorems in gaussian white noise. *Ann. Statist.*, 41(4):1999–2028.

Castillo, I. and Nickl, R. (2014). On the Bernsteinñ-von Mises phenomenon for nonparametric bayes procedures. *Ann. Statist.*, 42(5):1941–1969.

Choudhuri, N., Ghosal, S., and Roy, A. (2007). Nonparametric binary regression using a Gaussian process prior. *Statistical Methodology*, 4:227–243.

Cobos, F., Kühn, T., and Sickel, W. (2015). Optimal approximation of multivariate periodic Sobolev functions in the sup-norm. *Journal of Functional Analysis*, 270.

Cox, D. D. (1993). An analysis of bayesian inference for nonparametric regression. *Ann. Statist.*, 21(2):903–923.

de Jonge, R. and van Zanten, J. H. (2009). Adaptive nonparametric bayesian inference using location-scale mixture priors. Technical report.

Diaconis, P. and Freedman, D. (1986). On the consistency of Bayes estimates. *Ann. Statist.*, 14:1–26.

Donoho, D. L. (1994). Statistical estimation and optimal recovery. *Ann. Statist.*, 22(1):238–270.

Doob, J. L. (1949). Application of the theory of martingales. In *Le calcul des probabilités et ses applications*, number 13 in CNRS International Colloquia, pages 23–27. CNRS, Paris.

Ferrari-Trecate, G., Williams, C., and Opper, M. (1998). Finite-dimensional approximation of gaussian processes. In Kearns, M., Solla, S., and Cohn, D., editors, *Advances in Neural Information Processing Systems*, volume 11. MIT Press.

Freedman, D. (1963). On the asymptotic behavior of Bayes' estimates in the discrete case. *The Annals of Mathematical Statistics*, 34(4):1386 – 1403.

Gal, Y., van der Wilk, M., and Rasmussen, C. E. (2014). Distributed variational inference in sparse Gaussian process regression and latent variable models. *arXiv:1402.1389*.

Ghosal, S., Ghosh, J. K., and van der Vaart, A. (2000). Convergence rates of posterior distributions. *Ann. Statist.*, 28:500–531.

Ghosal, S. and van der Vaart, A. (2007). Convergence rates of posterior distributions for non iid observations. *Ann. Statist.*, 35(1):192–223.

Ghosal, S. and Van der Vaart, A. (2017). *Fundamentals of nonparametric Bayesian inference*, volume 44. Cambridge University Press.

Gibbs, N., Jr, W. P., and Stockmeyer, P. (1976). An algorithm for reducing the bandwidth and profile of a sparse matrix. *SIAM J. Numer. Anal.*, 13(2):236–250.

Giné, E. and Nickl, R. (2010). Confidence bands in density estimation. *Ann. Statist.*, 38(2):1122–1170.

Giné, E. and Nickl, R. (2016). *Mathematical foundations of infinite-dimensional statistical models*. Cambridge series in statistical and probabilistic mathematics.

Goldenshluger, A. and Nemirovski, A. (1997). On spatial adaptive estimation of nonparametric regression. *Math. Meth. Statistics*, 6:135–170.

Guhaniyogi, R., Li, C., Savitsky, T. D., and Srivastava, S. (2017). A divide-and-conquer bayesian approach to large-scale kriging. *arXiv preprint arXiv:1712.09767*.

Hadji, A. and Szabo, B. (2021). Can we trust bayesian uncertainty quantification from gaussian process priors with squared exponential covariance kernel? *SIAM/ASA Journal on Uncertainty Quantification*, 9(1):185–230.

Hensman, J., Fusi, N., and Lawrence, N. D. (2013). Gaussian processes for big data. UAI'13, page 282–290, Arlington, Virginia, USA. AUAI Press.

Huang, T.-M. (2004). Convergence rates for posterior distributions and adaptive estimation. *The Annals of Statistics*, 32(4):1556 – 1593.

Hunter, J. (2013). Distributions and sobolev spaces. Lecture Notes: Analysis Prelim Workshop. Department of Mathematics of the University of California Davis.

Isserlis, L. (1916). On certain probable errors and correlation coefficients of multiple frequency distributions with skew regression. *Biometrika*, 11(3):185–190.

Jiang, W. and Zhang, C.-H. (2009). General maximum likelihood empirical bayes estimation of normal means. *Ann. Statist.*

Johnstone, I. and Silverman, B. (2005). Empirical bayes selection of wavelet thresholds. *Ann. Statist.*, 33:1700–1752.

Jordan, M. and Jacobs, R. (1994). Hierarchical mixtures of experts and the em algorithm. *Neural Comput.*, 6(2):181–214.

Karhunen, K. (1947). über lineare Methoden in der Wahrscheinlichkeitsrechnung. *Annales Academiae Scientiarum Fennicae Series A. I. Mathematica*, (37):1–79.

Kim, H.-M., Mallick, B. K., and Holmes, C. C. (2005). Analyzing nonstationary spatial data using piecewise gaussian processes. *Journal of the American Statistical Association*, 100(470):653–668.

Kim, Y. and Lee, J. (2001). On posterior consistency of survival models. *The Annals of Statistics*, 29(3):666 – 686.

Knapik, B., van der Vaart, A. W., and van Zanten, J. H. (2011). Bayesian inverse problems with Gaussian priors. *Ann. Statist.*, 39(5):2626–2657.

Knapik, B. T., Szabó, B. T., van der Vaart, A. W., and van Zanten, J. H. (2016). Bayes procedures for adaptive inference in inverse problems for the white noise model. *Probability Theory and Related Fields*, 164(3):771–813.

Lehmann, E. and Casella, G. (1998). *Theory of Point Estimation (revised edition)*. Springer-Verlag, New York.

Lember, J. and van der Vaart, A. (2007). On universal bayesian adaptation. *Statistics & Decisions*, 25(2):127–152.

Lepski, O. V. and Spokoiny, V. G. (1997). Optimal pointwise adaptive methods in nonparametric estimation. *The Annals of Statistics*, 25(6):2512 – 2546.

Loève, M. (1978). *Probability theory Vol. II*, volume 46 of *Graduate Texts in Mathematics*. Springer-Verlag.

Mallasto, A. and Feragen, A. (2017). Learning from uncertain curves: The 2-Wasserstein metric for Gaussian processes. In Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R., editors, *Advances in Neural Information Processing Systems 30*, pages 5660–5670. Curran Associates, Inc.

Meeds, E. and Osindero, S. (2006). An alternative infinite mixture of gaussian process experts. In Weiss, Y., Schölkopf, B., and Platt, J., editors, *Advances in Neural Information Processing Systems*, volume 18. MIT Press.

Mercer, J. (1909). Functions of positive and negative type, and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society, London*, 209:415–446.

Minsker, S., Srivastava, S., Lin, L., and Dunson, D. (2014). Robus and scalabale Bayes via a media of subset posterior measures. *arXiv preprint*.

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability and its Applications*, 9:141–142.

Ng, J. and Deisenroth, M. (2014). Hierarchical mixture-of-experts modelfor large-scale gaussian process regression. *arXiv e-prints*.

Ng, J. and Deisenroth, M. (2015). In *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 1481–1490.

Nussbaum, M. (1996). Asymptotic equivalence of density estimation and gaussian white noise. *Ann. Statist.*, 24(6):2399–2430.

Opper, M. and Vivarelli, F. (1999). General bounds on bayes errors for regression with gaussian processes. In Kearns, M., Solla, S., and Cohn, D., editors, *Advances in Neural Information Processing Systems II*, pages 302–308. MIT Press.

Park, C. and Apley, D. W. (2018). Patchwork kriging for large-scale gaussian process regression. *J. Mach. Learn. Res.*, 19:7:1–7:43.

Park, C. and Huang, J. Z. (2016). Efficient computation of gaussian process regression for large spatial data sets by patching local gaussian processes. *Journal of Machine Learning Research*, 17(174):1–29.

Quiñonero-Candela, J. and Rasmussen, C. E. (2005). A unifying view of sparse approximate gaussian process regression. *J. Machine Learning Research*, 6:1939–1959.

Rasmussen, C. and Ghahramani, Z. (2002). Infinite mixtures of gaussian process experts. In Dietterich, T., Becker, S., and Ghahramani, Z., editors, *Advances in Neural Information Processing Systems*, volume 14. MIT Press.

Rasmussen, C. and Williams, C. (2006). *Gaussian processes for machine learning*. MIT Press, Boston.

Rasmussen, C. E. (2004). Gaussian processes in machine learning. In *Advanced lectures on machine learning*, pages 63–71. Springer.

Ray, K. (2017). Adaptive bernstein-von mises theorems in gaussian white noise. *Ann. Statist.*, 45(6):2511–2536.

Robert, C. (2001). *The Bayesian Choice*. Springer-Verlag, New York, second edition.

Robins, J. and van der Vaart, A. (2006). Adaptive nonparametric confidence sets. *Ann. Statist.*, 34(1):229–253.

Rousseau, J. and Szabo, B. (2017). Asymptotic behaviour of the empirical bayes posteriors associated to maximum marginal likelihood estimator. *Ann. Statist.*, 45:833–865.

Rousseau, J. and Szabo, B. (2020). Asymptotic frequentist coverage properties of bayesian credible sets for sieve priors. *Annals of Statistics*, 48(4):2155–2179.

Saad, Y. (1990). Sparskit: a basic tool kit for sparse matrix computations.

Samo, Y.-L. K. and Roberts, S. J. (2016). String and membrane gaussian processes. *Journal of Machine Learning Research*, 17(131):1–87.

Schwartz, L. (1965). On Bayes procedures. *Z. Warsch. Verw. Gebiete*, 4:10–26.

Scott, S., Blocker, A., Bonassi, F., Chipman, H., George, E., and McCulloch, R. (2016). Bayes and big data: The consensus monte carlo algorithm. *International Journal of Management Science and Engineering Management*, 11(2):78–88.

Serra, P. and Krivobokova, T. (2017). Adaptive empirical bayesian smoothing splines. *Bayesian Anal.*, 12(1):219–238.

Sniekers, S. and van der Vaart, A. (2015a). Adaptive Bayesian credible sets in regression with a Gaussian process prior. *Electron. J. Stat.*, 9(2):2475–2527.

Sniekers, S. and van der Vaart, A. (2015b). Credible sets in the fixed design model with Brownian motion prior. *Journal of Statistical Planning and Inference*, 166:78–86.

Srivastava, S., Cevher, V., Dinh, Q., and Dunson, D. (2015). WASP: Scalable Bayes via barycenters of subset posteriors. In Lebanon, G. and Vishwanathan, S., editors, *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics*, volume 38 of *Proceedings of Machine Learning Research*, pages 912–920, San Diego, California, USA. PMLR.

Szabó, B. and van Zanten, H. (2019). An asymptotic analysis of distributed nonparametric methods. *Journal of Machine Learning Research*, 20(87):1–30.

Szabo, B. T., van der Pas, S., and van der Vaart, A. W. (2017). Uncertainty quantification for the horseshoe (with discussion). *Bayesian Analysis*, 12(4):1221 – 1274.

Szabo, B. T., van der Vaart, A. W., and van Zanten, J. H. (2013). Empirical bayes scaling of gaussian priors in the white noise model. *Electron. J. Statist.*, 7:991–1018.

Szabo, B. T., van der Vaart, A. W., and van Zanten, J. H. (2015). Frequentist coverage of adaptive nonparametric Bayesian credible sets. *Annals of Statistics*, 43(4):1391–1428.

Titsias, M. (2009). Variational learning of inducing variables in sparse Gaussian Processes. In *Artificial Intelligence and Statistics*, pages 567–574.

Tokdar, S. and Ghosal, S. (2005). Posterior consistency of Gaussian process priors in density estimation. *J. Statist. Plann. Inference*, 137:34–42.

Tresp, V. (2000). The generalized bayesian committee machine. In *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, page 130–139, New York, NY, USA. Association for Computing Machinery.

Tresp, V. (2001). Mixtures of gaussian processes. In Leen, T., Dietterich, T., and Tresp, V., editors, *Advances in Neural Information Processing Systems*, volume 13. MIT Press.

Tsybakov, A. B. (2009). Introduction to nonparametric estimation. revised and extended from the 2004 french original. translated by vladimir zaiats.

van der Pas, S., Szabo, B., and van der Vaart, A. (2017). Adaptive posterior contraction rates for the horseshoe. *Electron. J. Statist.*, 11(2):3196–3225.

van der Vaart, A. and van Zanten, J. H. (2007). Bayesian inference with rescaled Gaussian process priors. *Electron. J. Statist.*, 1:433–448.

van der Vaart, A. and van Zanten, J. H. (2009a). Adaptive Bayesian estimation using a Gaussian random field with inverse Gamma bandwidth. *Ann. Statist.*, 37:2655–2675.

van der Vaart, A. and van Zanten, J. H. (2011). Information rates of nonparametric Gaussian process methods. *Journal of Machine Learning Research*, 12:2095–2119.

van der Vaart, A. W. and van Zanten, J. H. (2008). Rates of contraction of posterior distributions based on Gaussian process priors. *Ann. Statist.*, 36(3):1435–1463.

van der Vaart, A. W. and van Zanten, J. H. (2009b). Adaptive bayesian estimation using a gaussian random field with inverse gamma bandwidth. *Ann. Statist.*, 37(5B):2655–2675.

Van Der Vaart, A. W. and Wellner, J. A. (1996). Weak convergence. In *Weak convergence and empirical processes*, pages 16–28. Springer.

Wasserman, L. (2006). *All of Nonparametric Statistics*. Springer-Verlag, Berlin, Heidelberg.

Watson, G. S. (1964). Smooth regression analysis. *Sankhyā Ser.*, 26:359–372.

Yoo, W. W. and Ghosal, S. (2016). Supremum norm posterior contraction and credible sets for nonparametric multivariate regression. *Ann. Statist.*, 44(3):1069–1102.