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).

Scalability and uncertainty of Gaussian processes

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op woensdag 25 januari 2023
klokke 11.15 uur

door

Mohamed Amine Hadji

geboren te Algiers
in 1993

Promotor:
Prof. dr. A.W. van der Vaart (Technische Universiteit Delft)

Co-promotor:
Dr. B.T. Szabó (Universitá Bocconi)

Promotiecommissie:
Prof. dr. F.A.J de Haas
Prof. dr. P.D. Grünwald
Prof. dr. I. Castillo (Sorbonne Université)
Dr. P.J. de Andrade Serra (Vrije Universiteit Amsterdam)
Dr. H.N. Kekkonen (Technische Universiteit Delft)

Contents

1 Introduction **7**
§1.1 Statistical science . . . 7
§1.2 Frequentist and Bayesian inference . . . 8
§1.2.1 Frequentist inference . . . 8
§1.2.2 Bayesian inference . . . 8
§1.2.3 Frequentist Bayes . . . 9
§1.3 Non-parametric asymptotic statistics . . . 10
§1.3.1 Consistency . . . 11
§1.3.2 Contraction rates . . . 12
§1.3.3 Bayesian uncertainty quantification . . . 13
§1.4 Gaussian processes . . . 14
§1.5 Models . . . 16
§1.5.1 Signal-in-white-noise model . . . 16
§1.5.2 Non-parametric regression . . . 17
§1.6 Adaptation . . . 17
§1.6.1 Empirical Bayes . . . 18
§1.6.2 Hierarchical Bayes . . . 19
§1.7 Distributed computation . . . 19
§1.7.1 Uniformly random partitioning . . . 20
§1.7.2 Spatial partitioning . . . 21
§1.8 Overview . . . 21
§1.9 Notations . . . 21

2 Adaptive credible sets with a squared-exponential GP prior **23**
§2.1 Main results . . . 23
§2.1.1 Model description . . . 23
§2.1.2 Uncertainty quantification . . . 25
§2.2 Numerical analysis . . . 29
§2.2.1 Gaussian white noise model . . . 30
§2.2.2 Non-parametric regression and classification . . . 30
§2.3 Discussion . . . 34
§2.4 Some properties of the MMLE . . . 36
§2.4.1 Deterministic bounds . . . 36
§2.4.2 Contraction rates . . . 39
§2.4.3 Proof of Theorem 2.4.4 . . . 40
§2.5 Proof of the empirical Bayes part of Theorem 2.1.5 . . . 44

§2.6 Proof of Theorem 2.1.4 . . . 48
§2.7 Proof of Theorem 2.1.3 and the empirical Bayes part of Corollary 2.1.2 49
§2.8 Proof of Theorem 2.4.1 . . . 50
§2.8.1 $\mathbb{M}_n(a)$ on $[1, \underline{a}_n)$ . . . 50
§2.8.2 $\mathbb{M}_n(a)$ on $[\overline{a}_n, A_n]$ . . . 53
§2.9 Proof of Theorem 2.1.6 . . . 56
§2.10 Proofs for the Hierarchical Bayes procedure . . . 58
§2.10.1 Proof of Theorem 2.4.5 . . . 58
§2.10.2 Proof of Theorem 2.1.3 - Hierarchical Bayes part . . . 59
§2.10.3 Proof of Theorem 2.1.5 - Hierarchical Bayes part . . . 61
§2.10.4 Proof of Corollary 2.1.2 . . . 62
§2.10.5 Proof of Lemma 2.10.1 . . . 63
§2.11 Technical Lemmas . . . 63
§2.12 Extra simulation study . . . 64

3 Optimal recovery and coverage for distributed Bayesian non-parametric regression 69
§3.1 GP regression framework . . . 69
§3.2 Distributed GP regression . . . 71
§3.2.1 Optimal Distributed Methods . . . 73
§3.2.2 Posterior contraction rate . . . 74
§3.3 Distributed uncertainty quantification . . . 76
§3.4 Discussion . . . 78
§3.5 Proofs of the main results . . . 78
§3.5.1 Kernel Ridge Regression in non-distributed setting . . . 78
§3.5.2 Kernel Ridge Regression in distributed setting . . . 80
§3.5.3 Proof of Theorem 3.2.2 . . . 80
§3.5.4 Proof of Theorem 3.2.1 . . . 85
§3.5.5 Proof of Theorem 3.3.1 . . . 87
§3.6 Proof of the Corollaries . . . 96
§3.6.1 Proof of Corollary 3.2.3 . . . 96
§3.6.2 Proof of Corollary 3.2.4 . . . 97
§3.6.3 Proof of Corollary 3.3.2 . . . 98
§3.6.4 Proof of Corollary 3.3.3 . . . 99
§3.7 Technical lemmas . . . 99

4 Optimal recovery for spatially distributed Gaussian process regression 109
§4.1 Spatially distributed GP regression . . . 109
§4.1.1 Posterior contraction for distributed GP . . . 110
§4.1.2 Adaptation . . . 111
§4.2 Application to the Integrated Brownian Motion . . . 112
§4.3 Proofs of the general results . . . 113
§4.3.1 Proof of Theorem 4.1.1 . . . 114
§4.3.2 Proof of Theorem 4.1.2 . . . 114
§4.3.3 Proof of Corollary 4.2.1 . . . 115
§4.3.4 Proof of Corollary 4.2.2 . . . 118

§4.4 Auxiliary Lemmas 125

5 Simulation study **127**
§5.1 Distributed GP regression 127
§5.1.1 Uniformly random 128
§5.1.2 Spatial 128
§5.1.3 Weighted-average model 128
§5.2 Numerical study 129
§5.2.1 Simulated Data 129
§5.2.2 Airline Delays (USA Flight) 136

Bibliography **139**

Summary **147**

Samenvatting **149**

Acknowledgements **151**

Curriculum Vitae **153**