

AISTATS 2021

[HTTPS://AVT.IM/](https://AVT.IM/) •  @AVT_IM

Matérn Gaussian Processes on Graphs

Viacheslav Borovitskiy,* Iskander Azangulov,* Alexander Terenin,*
Peter Mostowsky, Marc Peter Deisenroth, and Nicolas Durrande



*equal contribution



St Petersburg
University

Imperial College
London

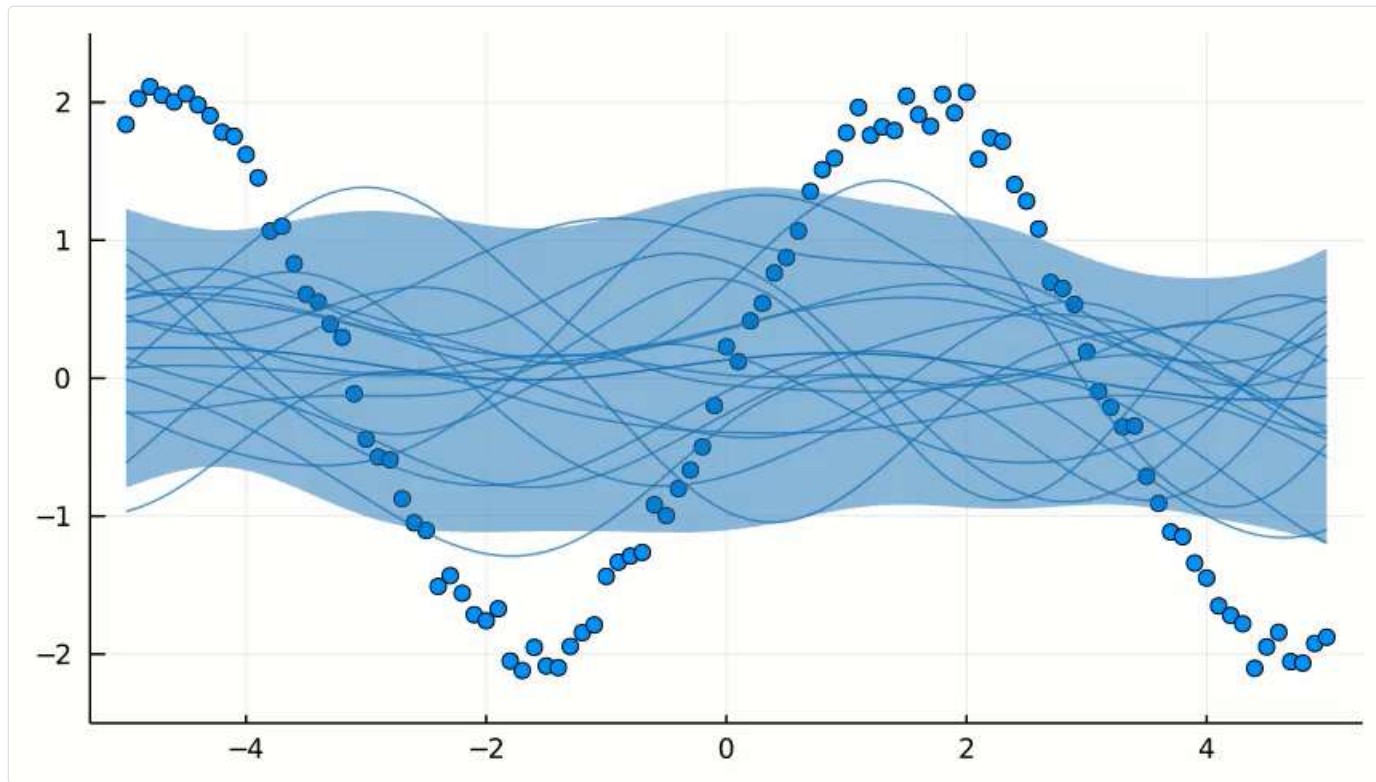


UCL



Secondmind

Gaussian Processes



Gaussian Processes

Definition. A *Gaussian process* is random function $f : X \rightarrow \mathbb{R}$ such that for any $x_1, \dots, x_n$, the vector $f(x_1), \dots, f(x_n)$ is multivariate Gaussian.

Every GP is characterized by a *mean* $\mu(\cdot)$ and a *kernel* $k(\cdot, \cdot)$. We have

$$f(\mathbf{x}) \sim \mathcal{N}(\boldsymbol{\mu}_x, \mathbf{K}_{xx})$$

where $\boldsymbol{\mu}_x = \mu(\mathbf{x})$ and $\mathbf{K}_{xx'} = k(\mathbf{x}, \mathbf{x}')$.

Bayesian Learning

$$y_i = f(x_i) \quad f \sim \text{GP}(0, k)$$

$$(f \mid \mathbf{y})(\cdot) = f(\cdot) + \mathbf{K}_{(\cdot)\mathbf{x}} \mathbf{K}_{\mathbf{x}\mathbf{x}}^{-1} (\mathbf{y} - f(\mathbf{x}))$$

J. T. Wilson, V. Borovitskiy, A. Terenin, P. Mostowsky, M. P. Deisenroth. Efficiently Sampling Functions from Gaussian Process Posteriors. ICML 2020.

Matérn Kernel

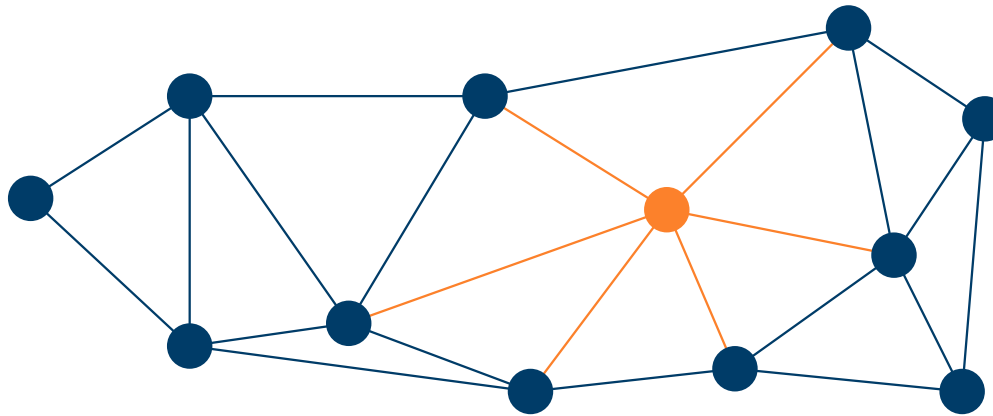
$$k(x, x') = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu} \frac{\|x - x'\|}{\kappa} \right)^\nu K_\nu \left(\sqrt{2\nu} \frac{\|x - x'\|}{\kappa} \right)$$

σ^2 : variance κ : length scale ν : smoothness

$\nu \rightarrow \infty$: recovers squared exponential kernel

Defines GPs $f : \mathbb{R}^d \rightarrow \mathbb{R}$

Weighted Undirected Graphs



$$f\left(\text{star graph}\right) \rightarrow \mathbb{R}$$

Stochastic Partial Differential Equations

$$\underbrace{\left(\frac{2\nu}{\kappa^2} - \Delta\right)^{\frac{\nu}{2} + \frac{d}{4}} f = \mathcal{W}}_{\text{Matérn}} \qquad \underbrace{e^{-\frac{\kappa^2}{4}\Delta} f = \mathcal{W}}_{\text{squared exponential}}$$

Δ : Laplacian $\mathcal{W}$: (rescaled) white noise

$e^{-\frac{\kappa^2}{4}\Delta}$: (rescaled) heat semigroup

The Graph Laplacian

$$(\Delta \mathbf{f})(x) = \sum_{x' \sim x} w_{xx'} (f(x) - f(x'))$$

$$\Delta = \mathbf{D} - \mathbf{W}$$

D: degree matrix **W**: (weighted) adjacency matrix

Note: different sign convention, analogous to Euclidean $-\Delta$

Graph Matérn Gaussian Processes

$$\underbrace{\left(\frac{2\nu}{\kappa^2} + \Delta\right)^{\frac{\nu}{2}} \mathbf{f} = \mathbf{w}}_{\text{Matérn}}$$

$$\underbrace{e^{\frac{\kappa^2}{4}} \Delta \mathbf{f} = \mathbf{w}}_{\text{squared exponential}}$$

Δ : graph Laplacian $\mathbf{w}$: standard Gaussian

Graph Matérn Gaussian Processes

$$\underbrace{\boldsymbol{f} \sim \text{N} \left(\mathbf{0}, \left(\frac{2\nu}{\kappa^2} + \mathbf{\Delta} \right)^{-\nu} \right)}_{\text{Matérn}}$$

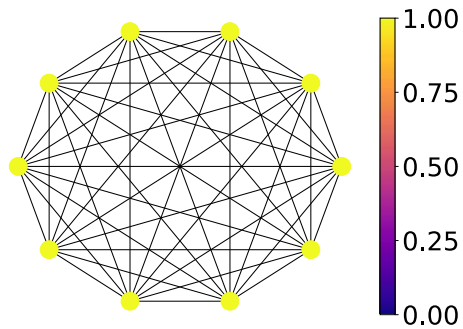
$$\underbrace{\boldsymbol{f} \sim \text{N} \left(\mathbf{0}, e^{-\frac{\kappa^2}{2}} \mathbf{\Delta} \right)}_{\text{squared exponential}}$$

Graph Fourier Features

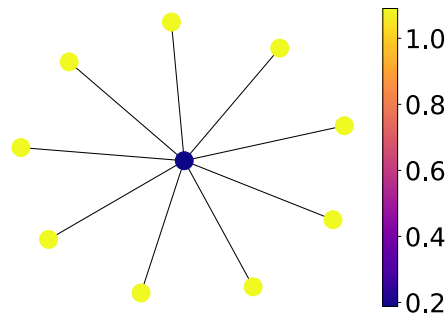
$$k_\nu(x, x') = \frac{\sigma^2}{C_\nu} \sum_{n=1}^{|G|} \left(\frac{2\nu}{\kappa^2} + \lambda_n \right)^{-\nu} \mathbf{f}_n(x) \mathbf{f}_n(x')$$

$\lambda_n, \mathbf{f}_n$: eigenvalues and eigenvectors of graph Laplacian

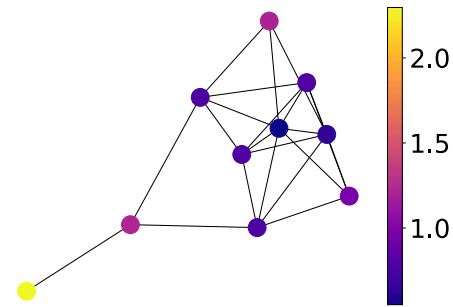
Prior Variance



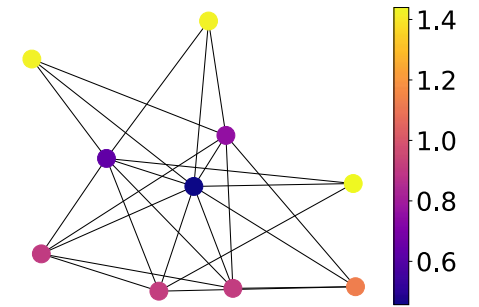
(a) Complete graph



(b) Star graph



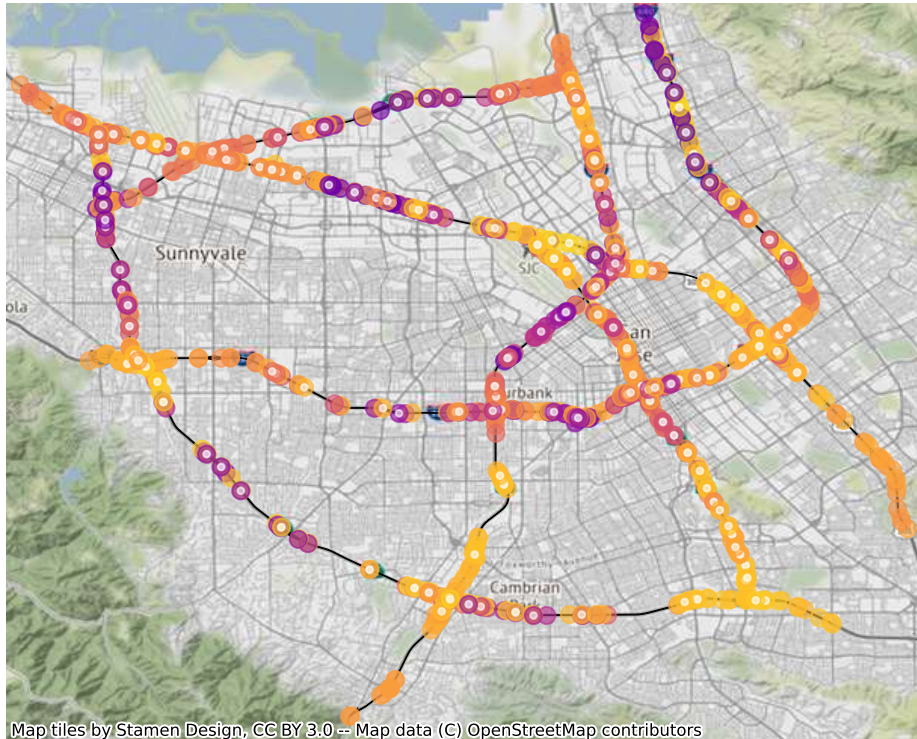
(c) Random graph



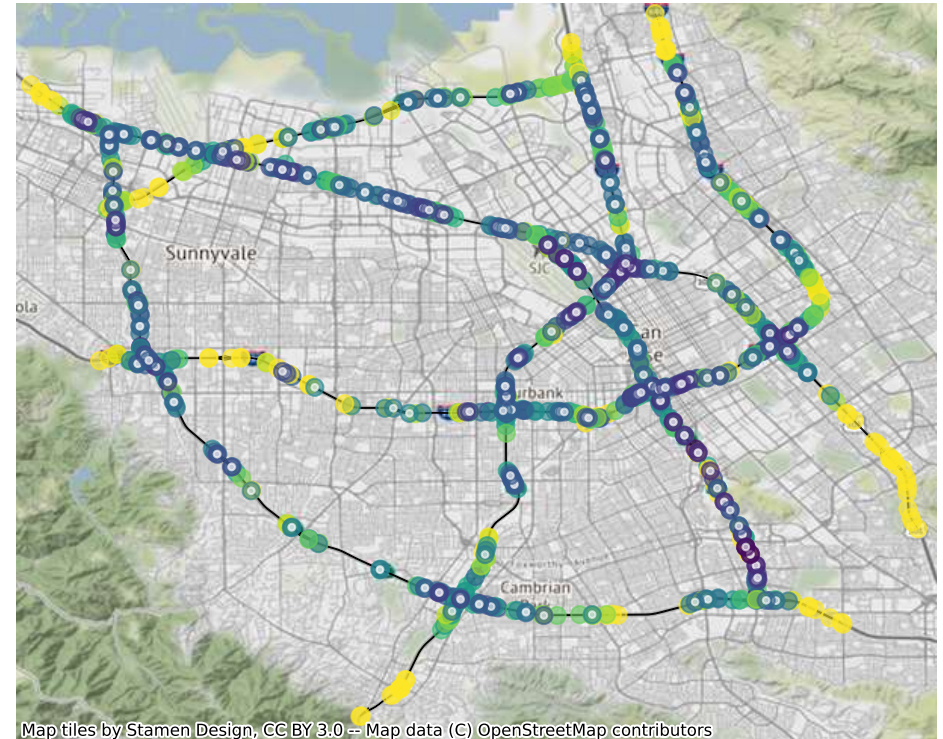
(d) Random graph

Prior variance depends on geometry

Example: Graph Interpolation of Traffic



(a) Predictive mean



(b) Standard deviation

Example: Graph Interpolation of Traffic

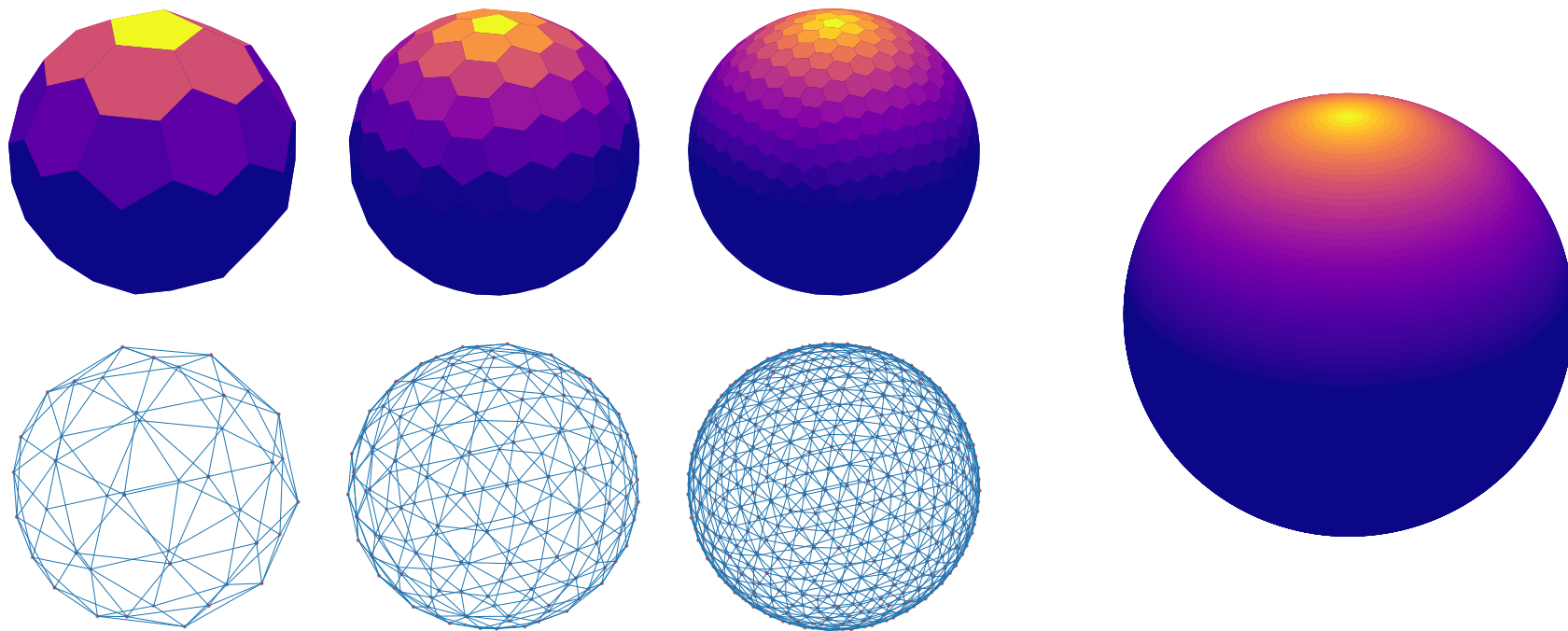


(a) Predictive mean



(b) Standard deviation

Connection with Matérn Gaussian Processes on Riemannian Manifolds



V. Borovitskiy, A. Terenin, P. Mostowsky, M. P. Deisenroth. Matérn Gaussian Processes on Riemannian Manifolds. NeurIPS 2020.

Thank you!

[HTTPS://AVT.IM/](https://AVT.IM/) •  @AVT_IM

V. Borovitskiy, I. Azangulov, A. Terenin, P. Mostowsky, M. P. Deisenroth. Matérn Gaussian Processes on Graphs. *Artificial Intelligence and Statistics*, 2021.

V. Borovitskiy, A. Terenin, P. Mostowsky, M. P. Deisenroth. Matérn Gaussian Processes on Riemannian Manifolds. *Advances in Neural Information Processing Systems*, 2020.

J. T. Wilson, V. Borovitskiy, A. Terenin, P. Mostowsky, M. P. Deisenroth. Pathwise Conditioning of Gaussian Processes. *Journal of Machine Learning Research*, 2021.

J. T. Wilson, V. Borovitskiy, A. Terenin, P. Mostowsky, M. P. Deisenroth. Efficiently Sampling Functions from Gaussian Process Posteriors. *International Conference on Machine Learning*, 2020.