Accelerating new physics searches with XSEC

Jeriek Van den Abeele

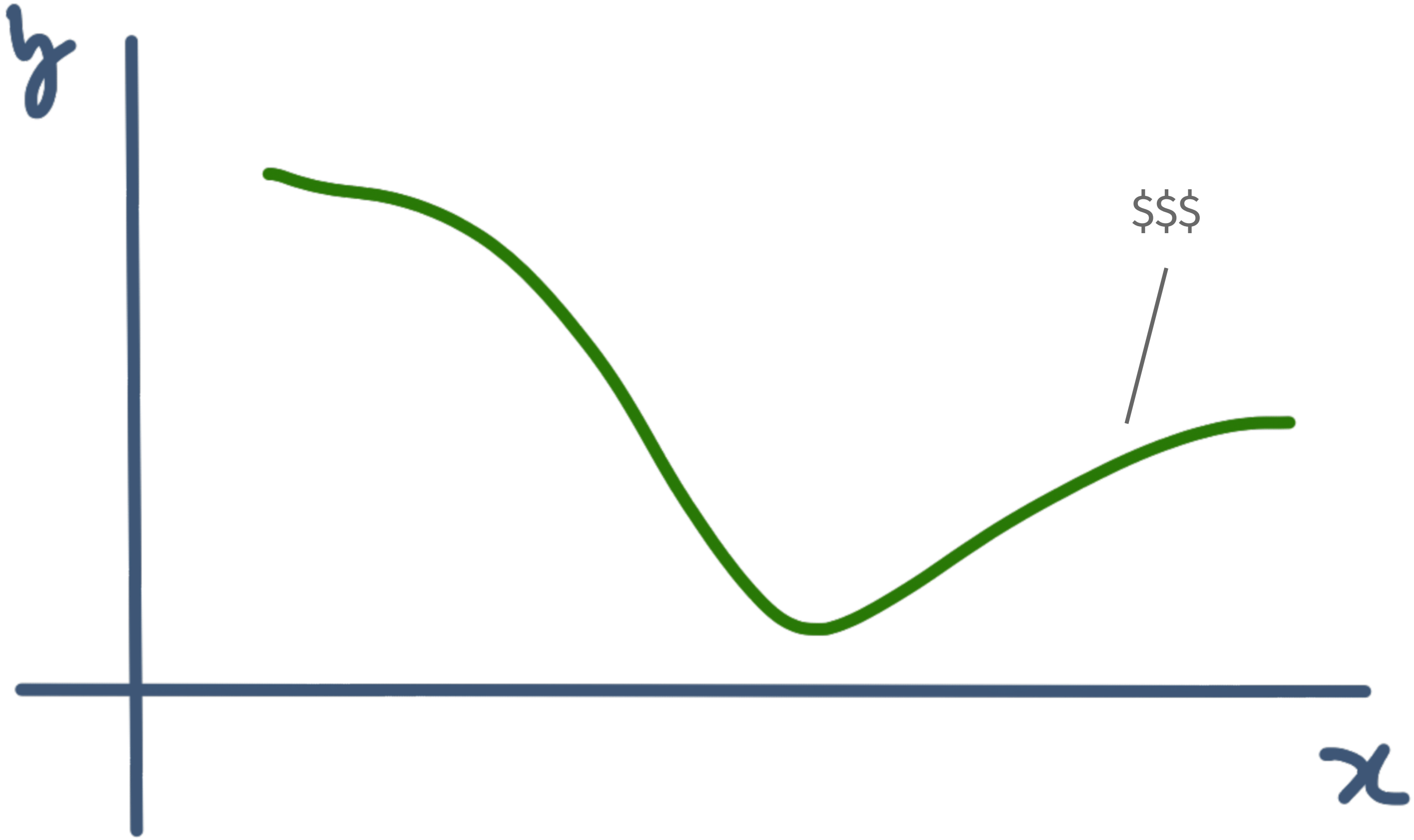
Spåtind 2020
Skeikampen – January 6

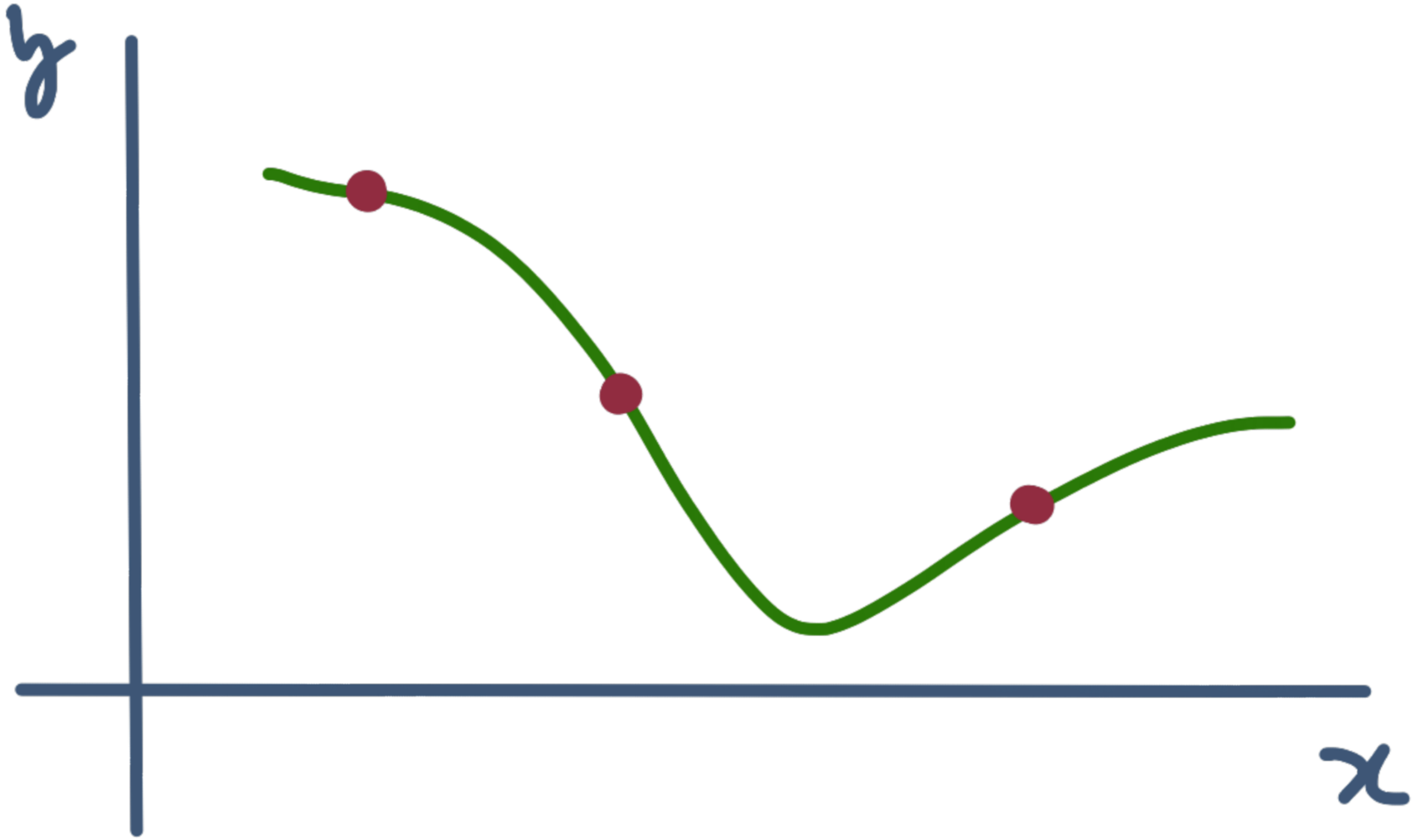
Based on work with A. Buckley, I.A.V. Holm, A. Kvellestad, A. Raklev, P. Scott, J.V. Sparre

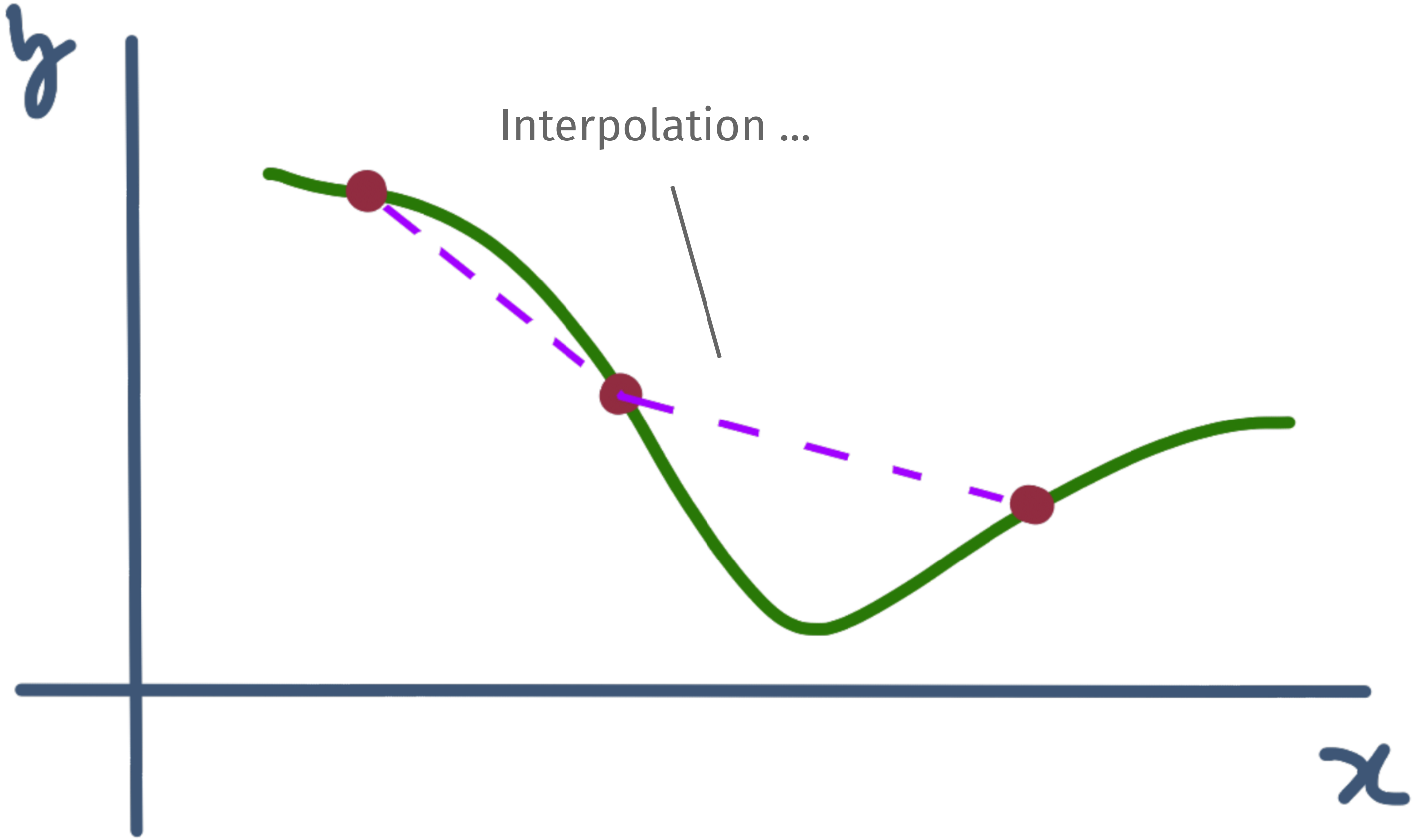
@JeriekVda

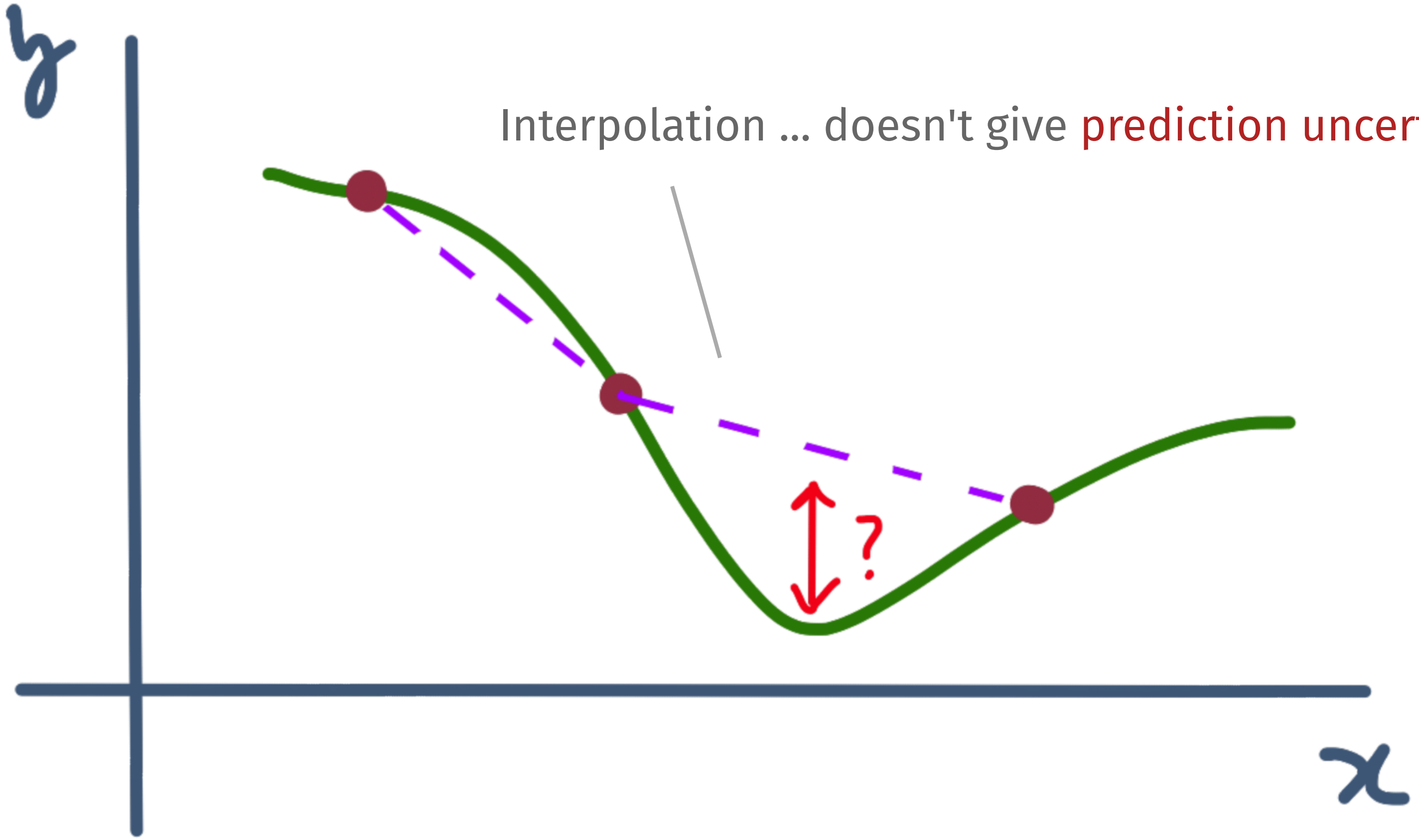
UiO : University of Oslo

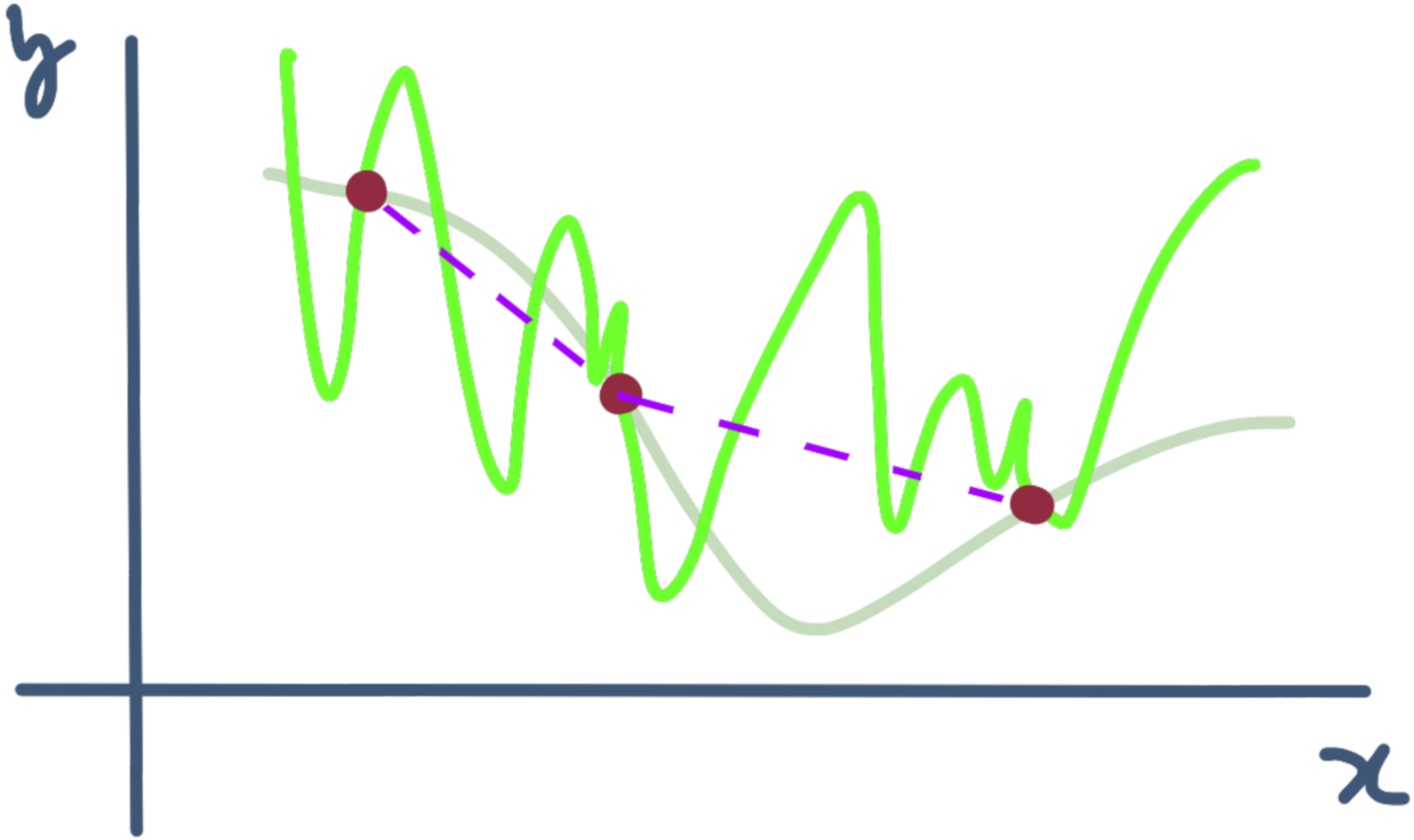
Time is money

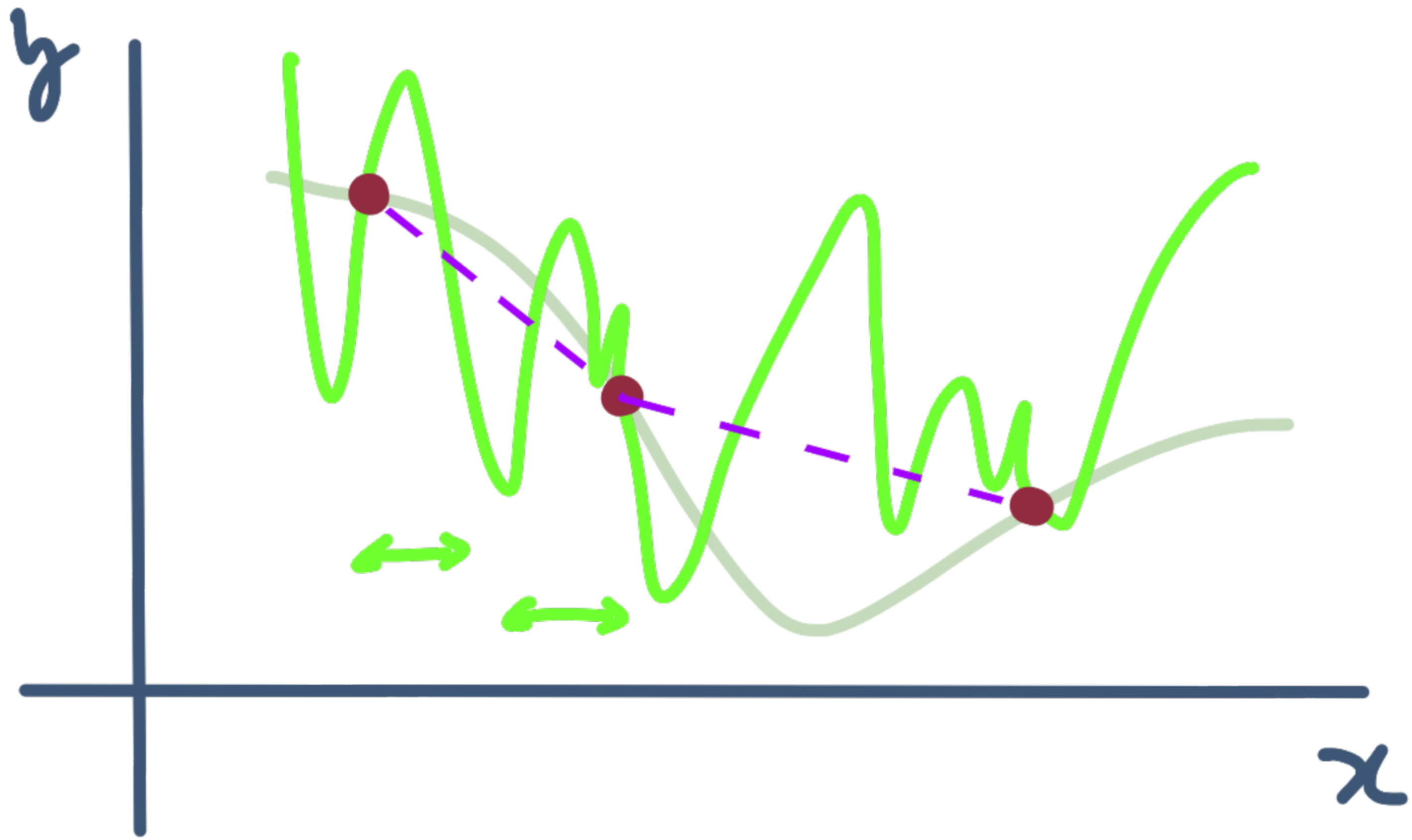


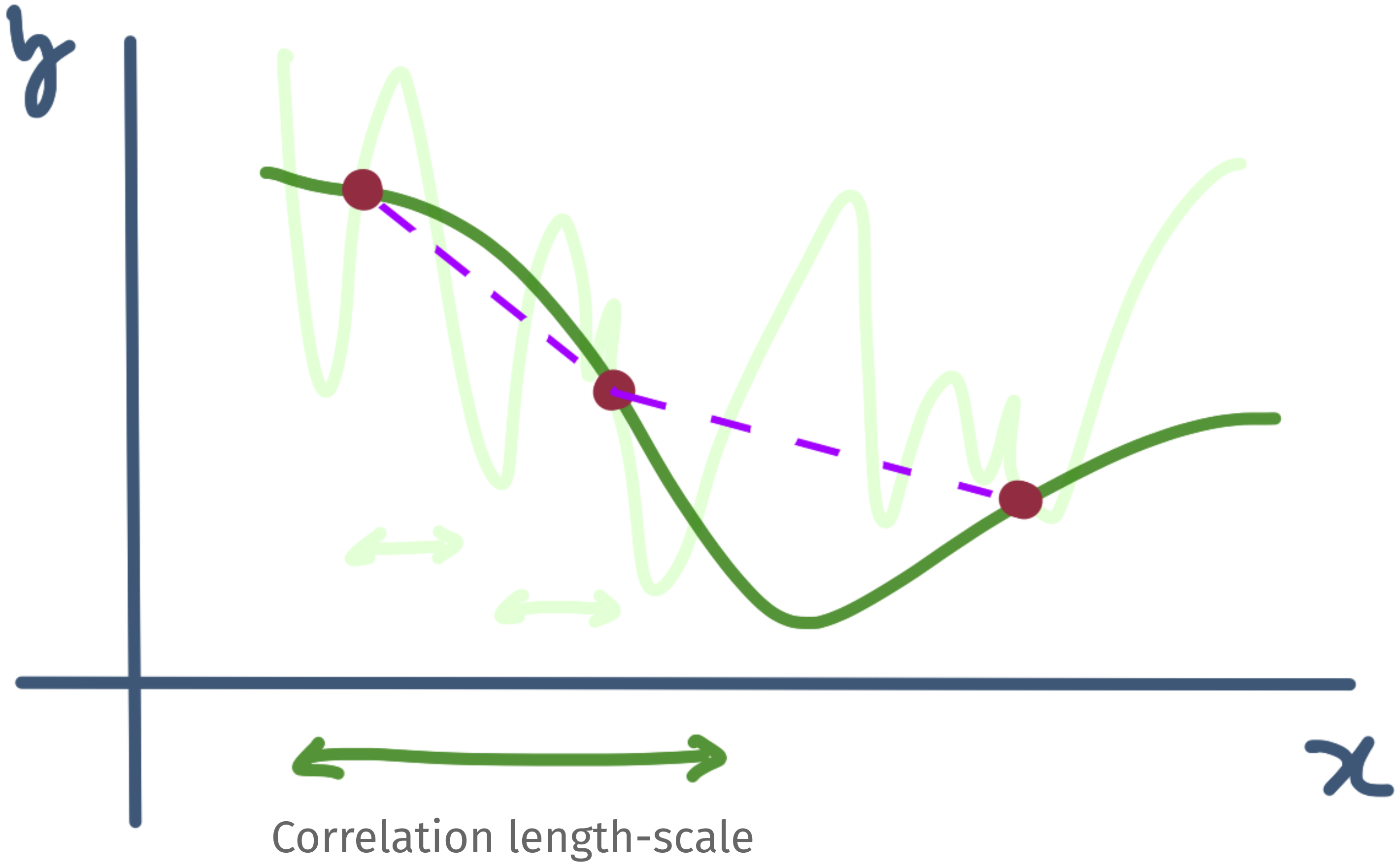

y
x


y
Interpolation ...
x

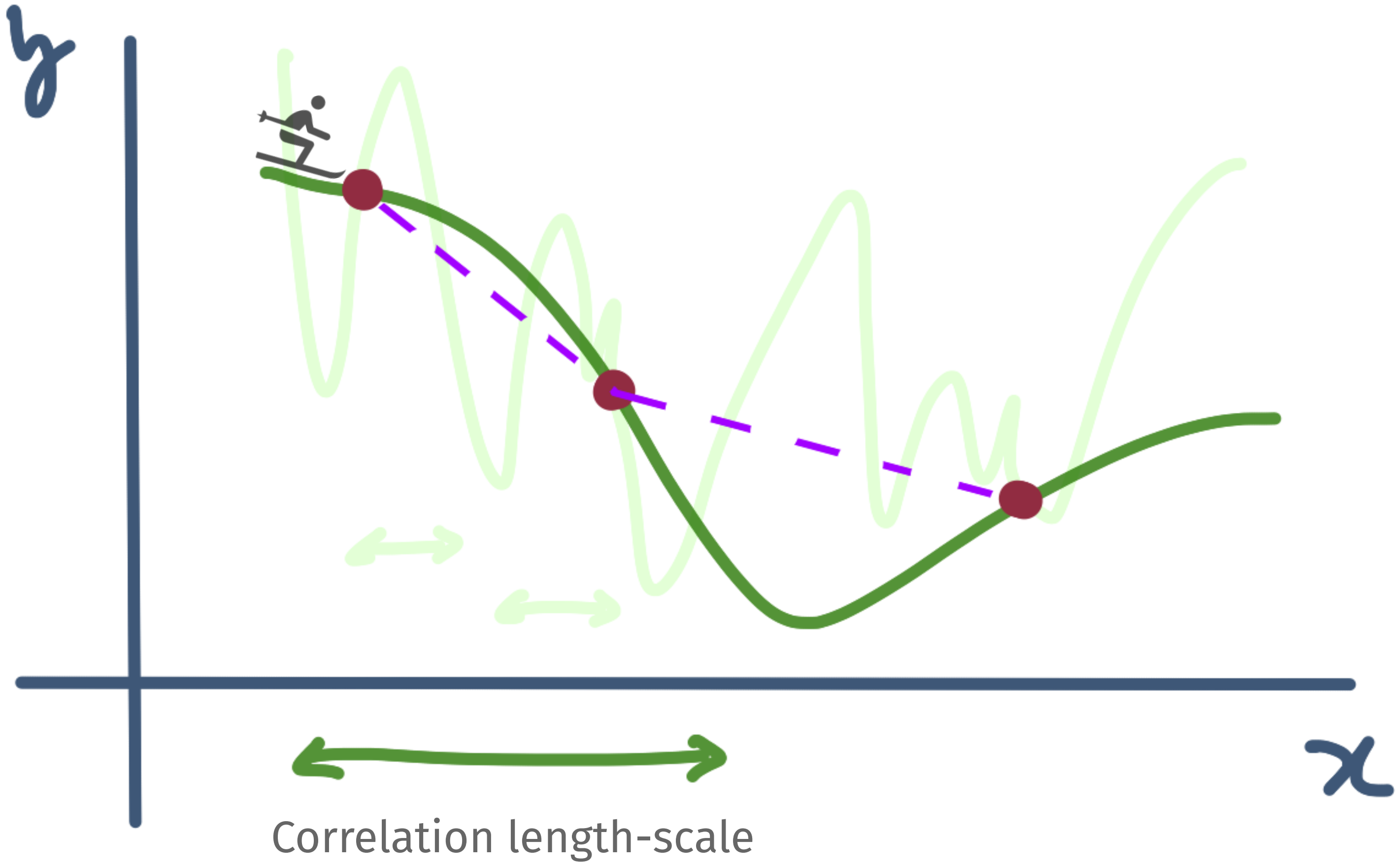

y
Interpolation ... doesn't give prediction uncertainty!
?
x

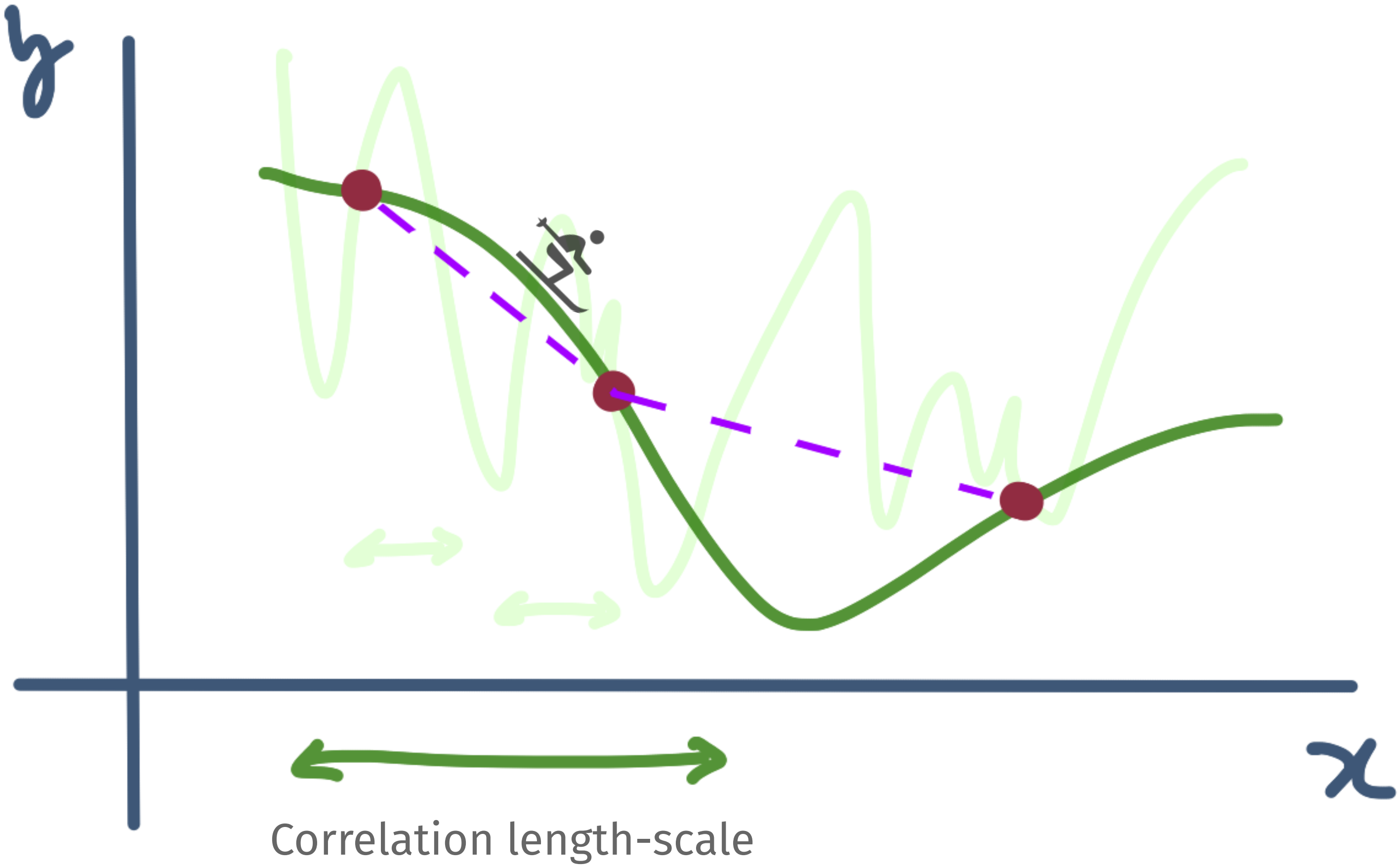

y
x

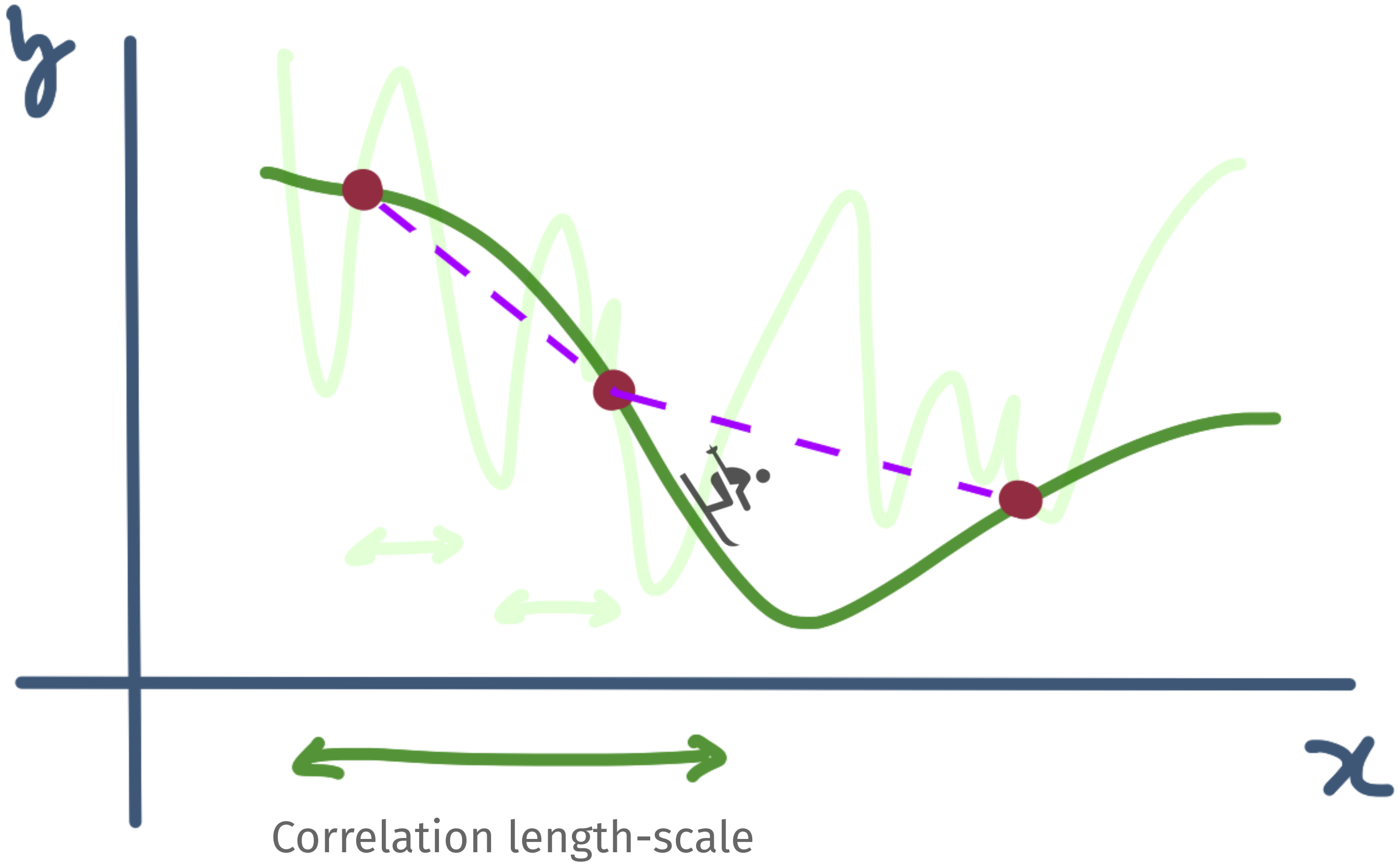

y
x



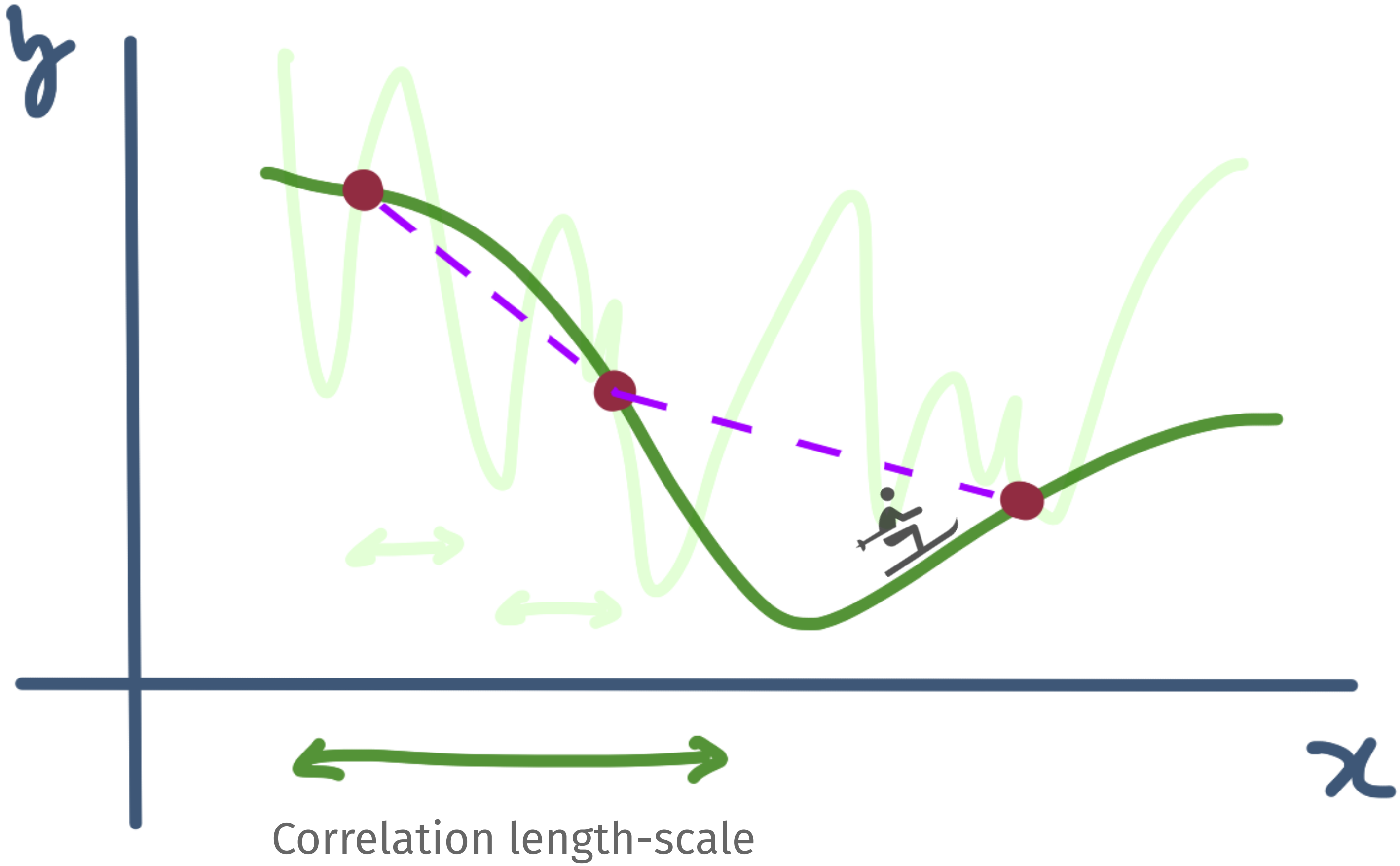
y
x
Correlation length-scale

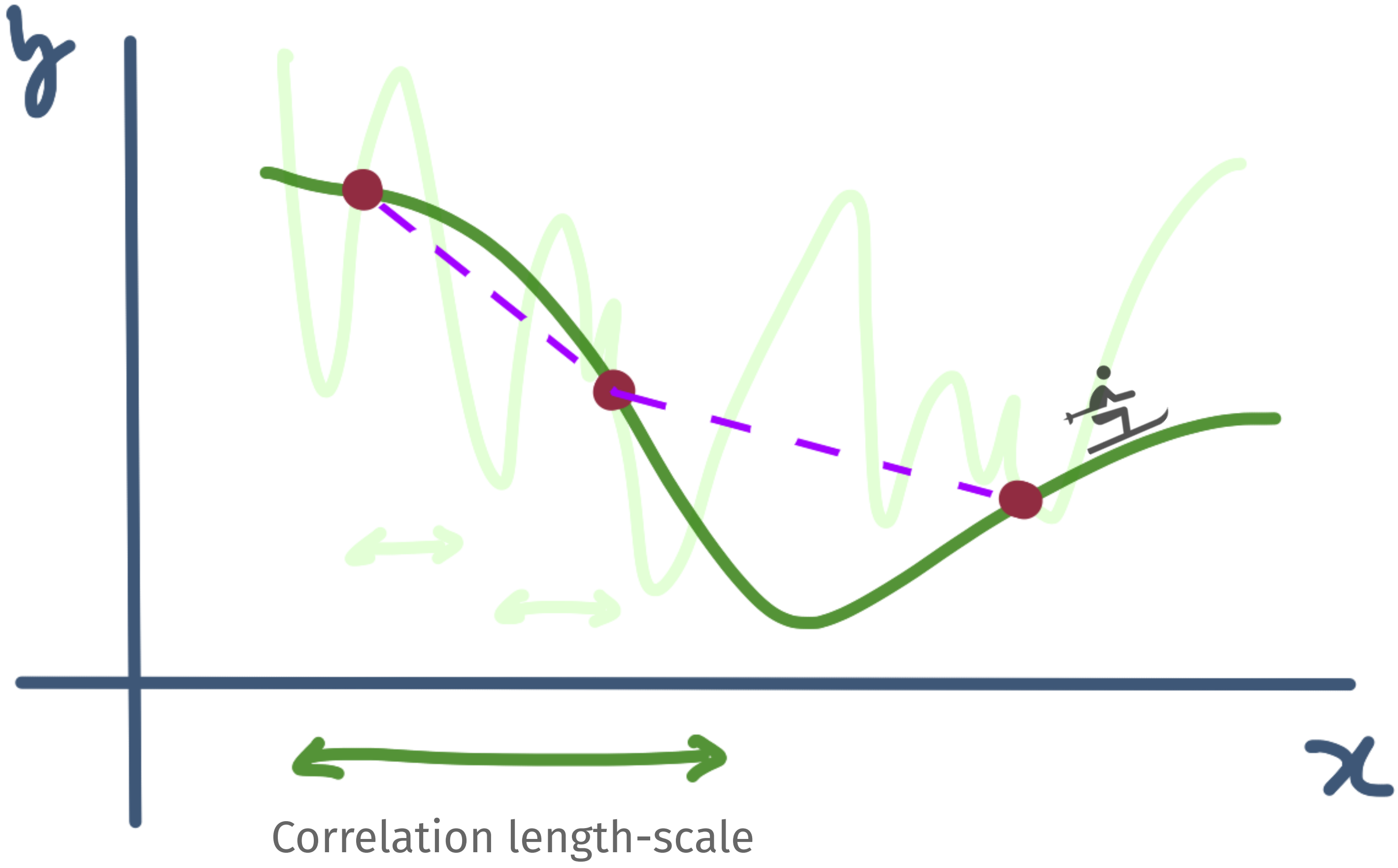

y
x
Correlation length-scale

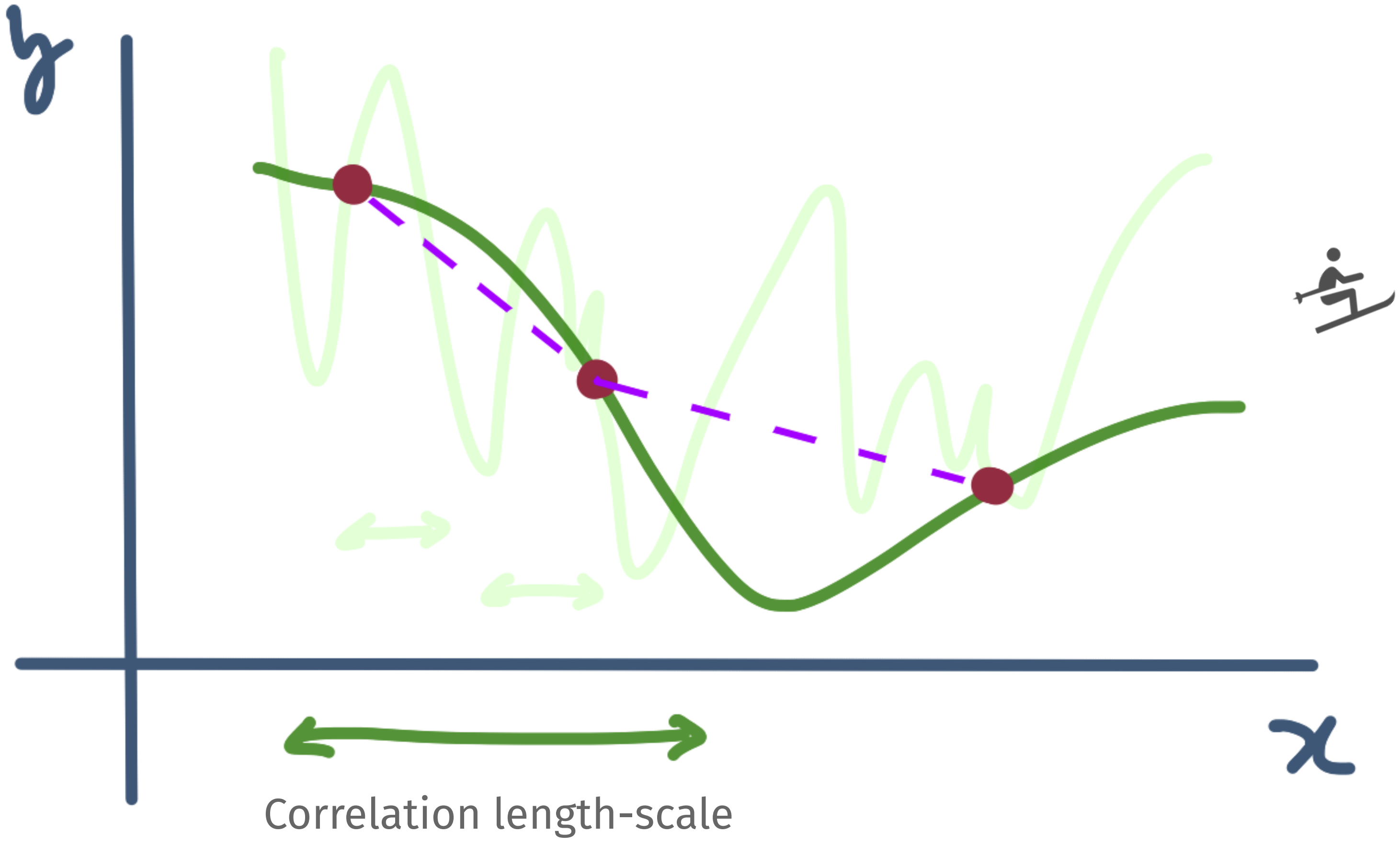

y
x
Correlation length-scale



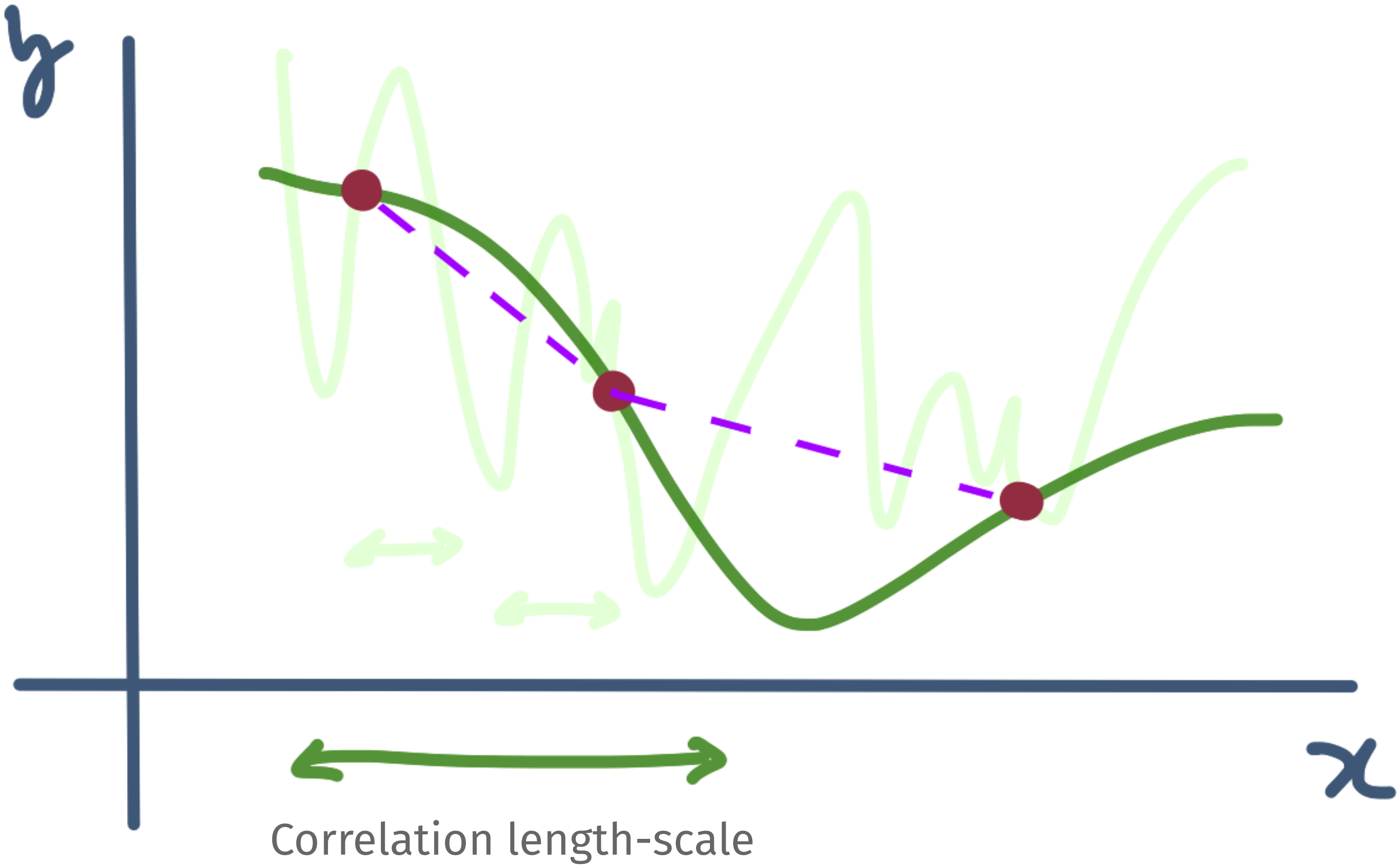
y
x
Correlation length-scale

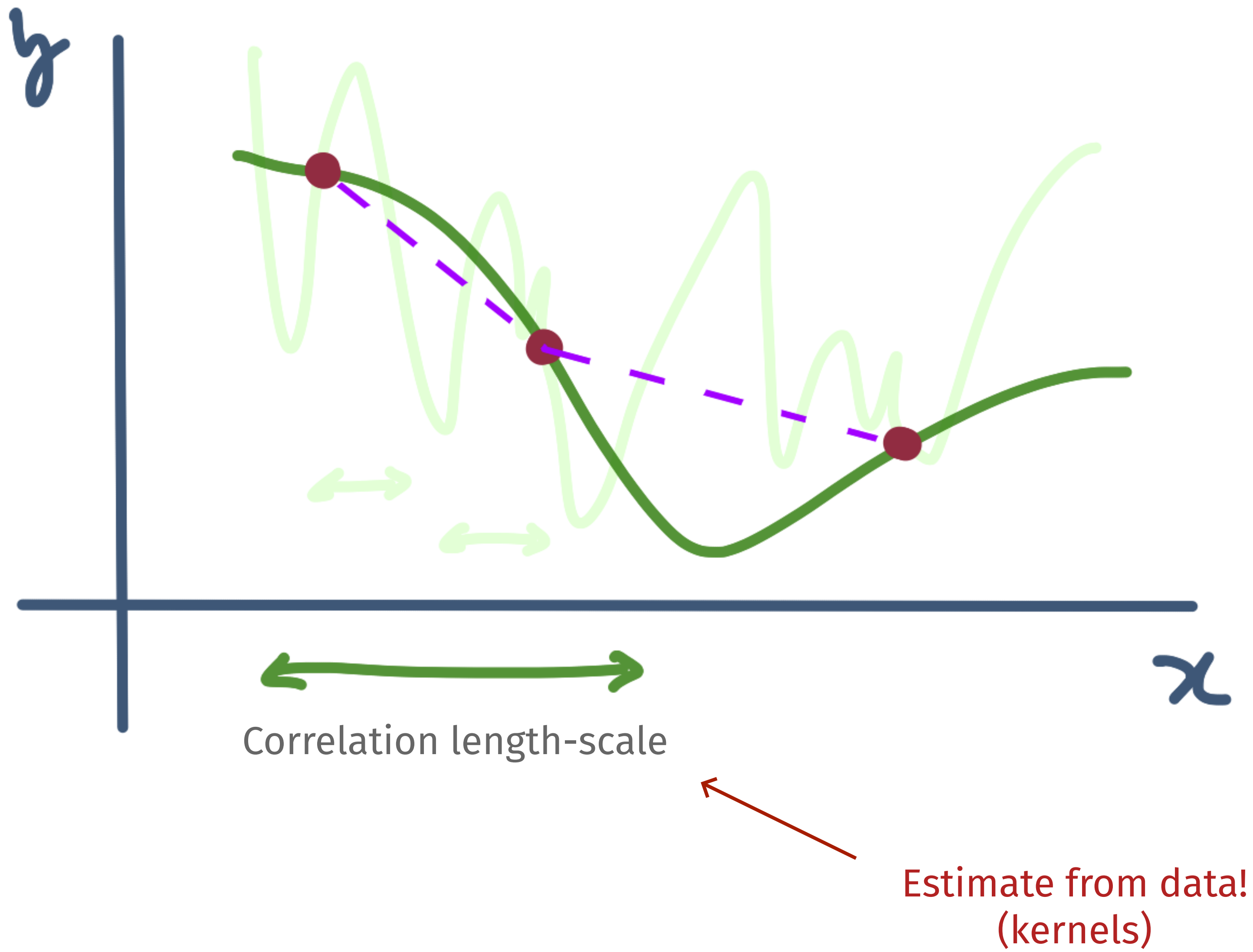

y
x
Correlation length-scale


y
x
Correlation length-scale

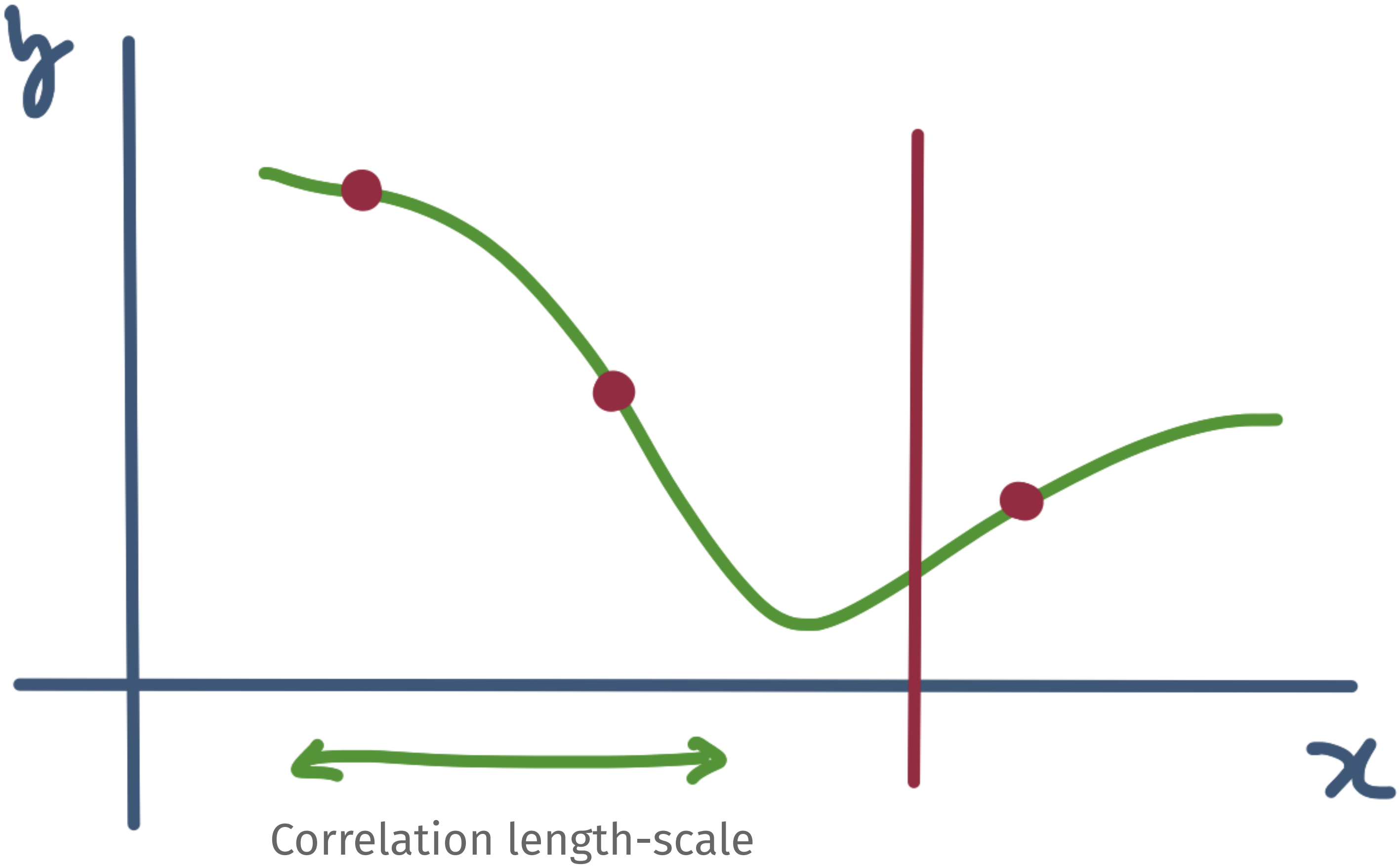


y
x
Correlation length-scale

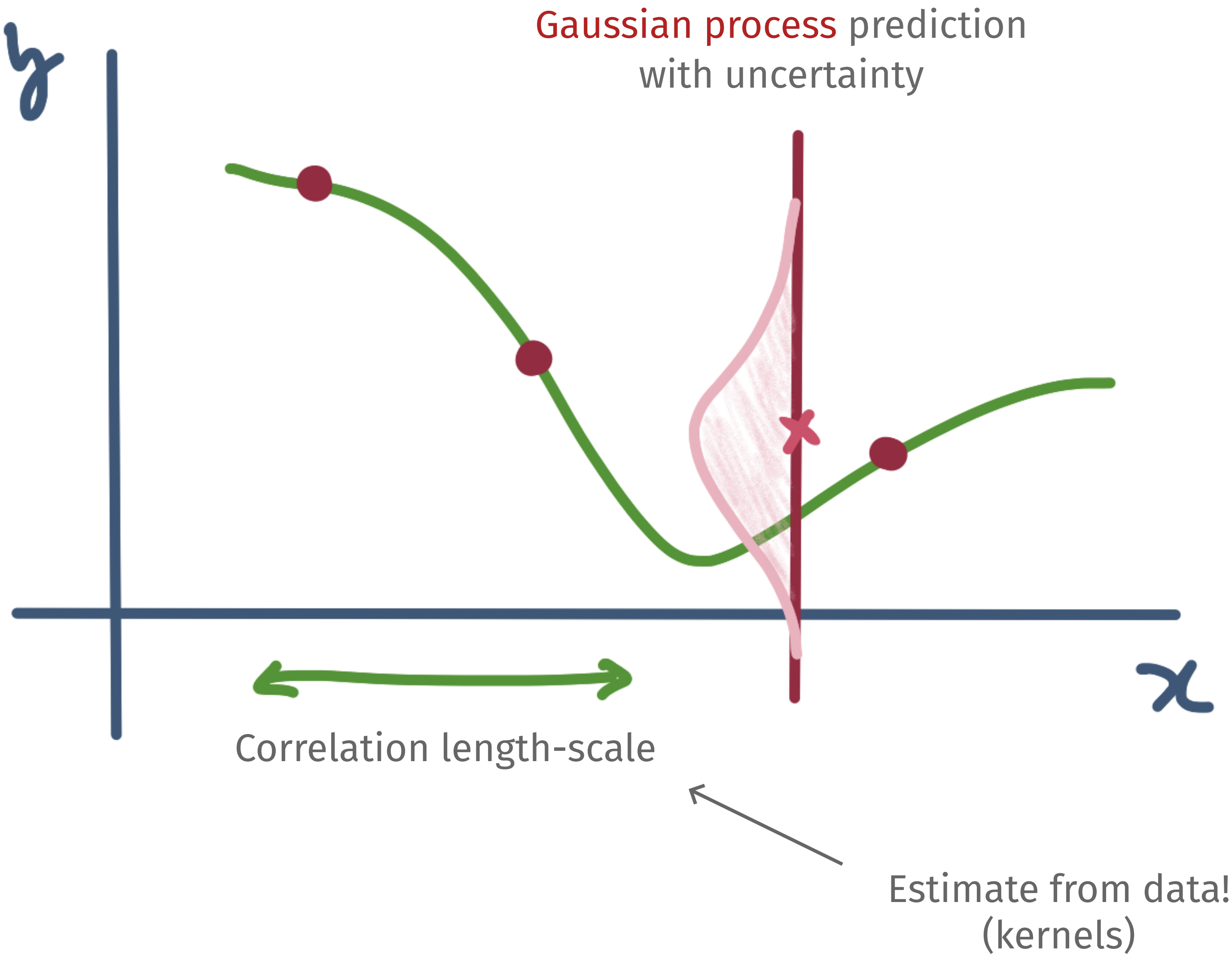

y
x
Correlation length-scale



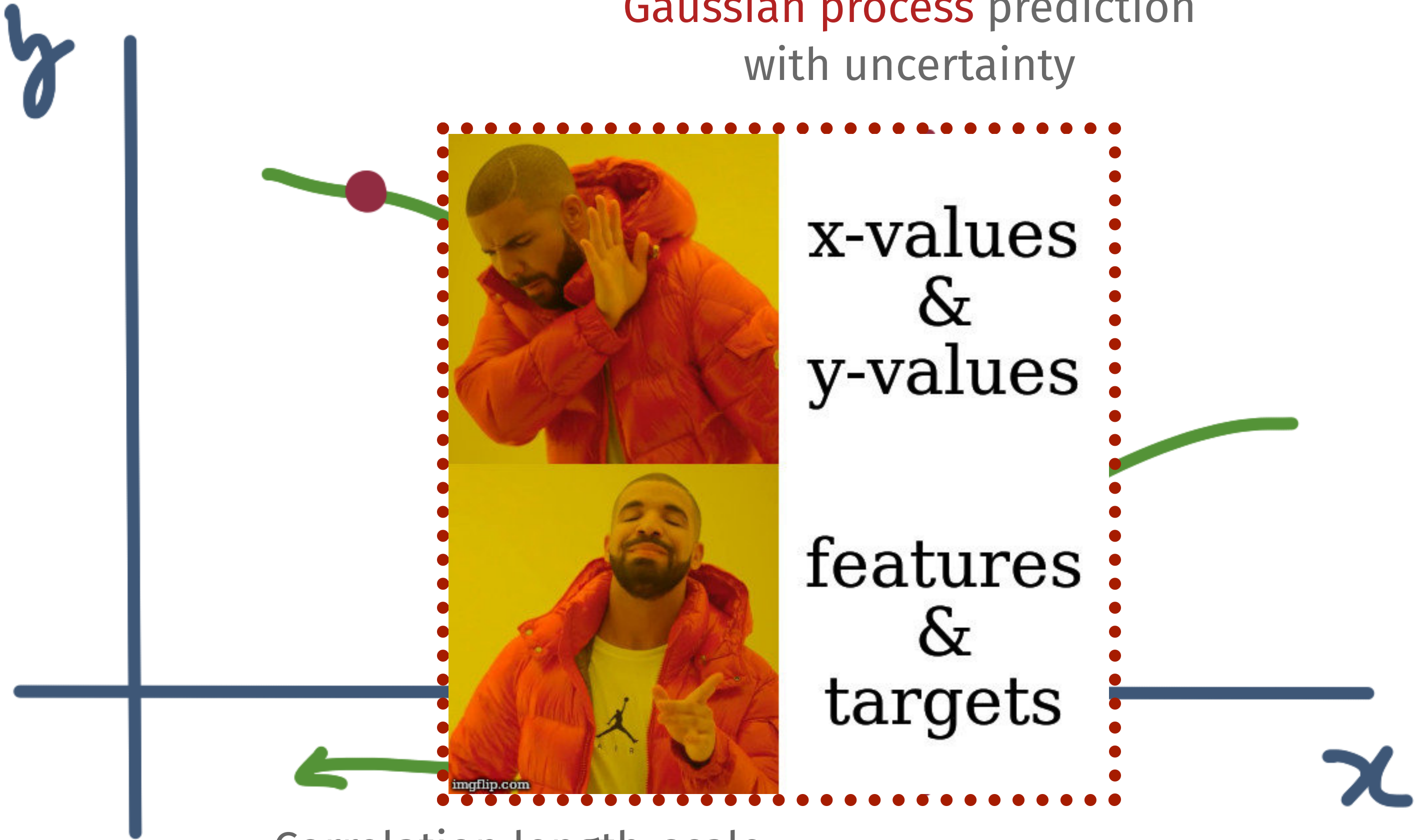
y
x
Correlation length-scale
Estimate from data!
(kernels)



Estimate from data!
(kernels)


Gaussian process prediction
with uncertainty
y
x
Correlation length-scale
Estimate from data!
(kernels)

Gaussian process prediction
with uncertainty



Correlation length-scale

Estimate from data!
(kernels)

Gaussian Processes 101

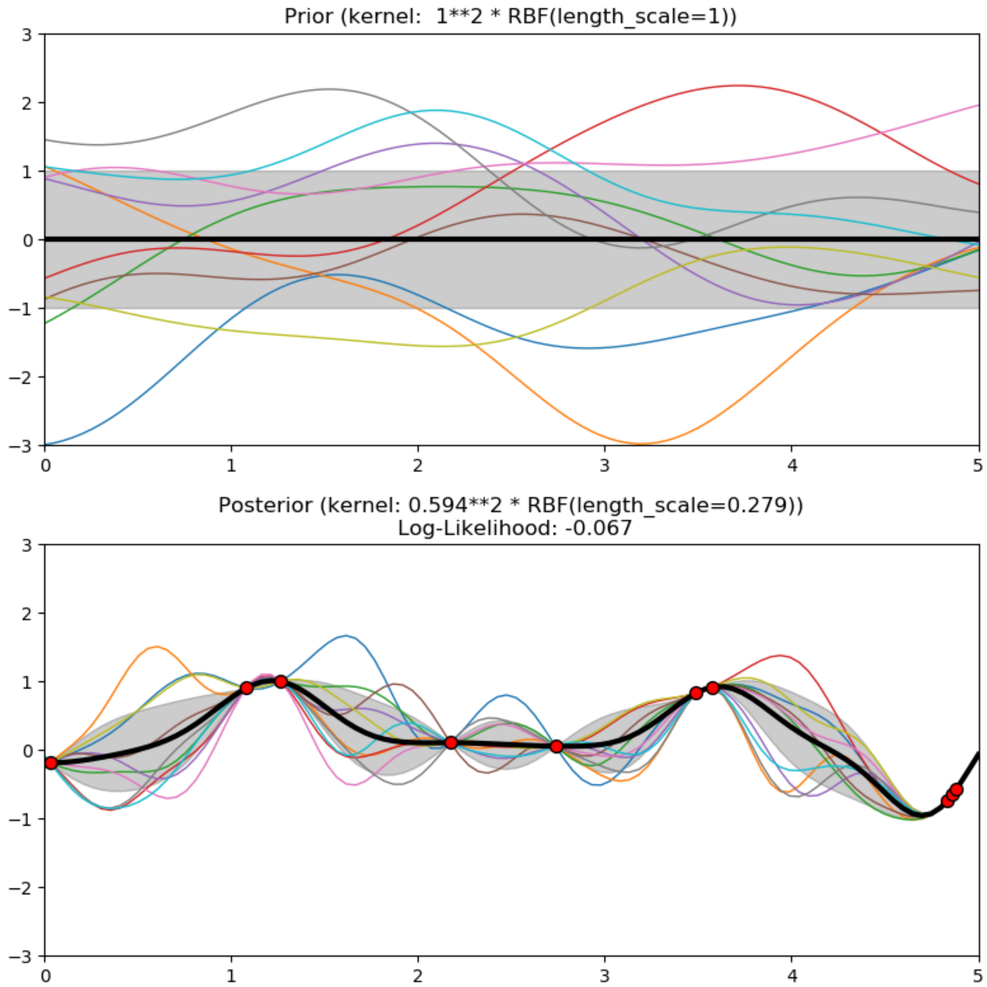
prior over all functions
with the estimated smoothness

$\mathcal{N}(\vec{0}, \Sigma^2)$

data

$\mathcal{N}(\tilde{\mu}, \tilde{\Sigma}^2)$

posterior over functions



Radial Basis Function kernel

Gaussian Processes 101

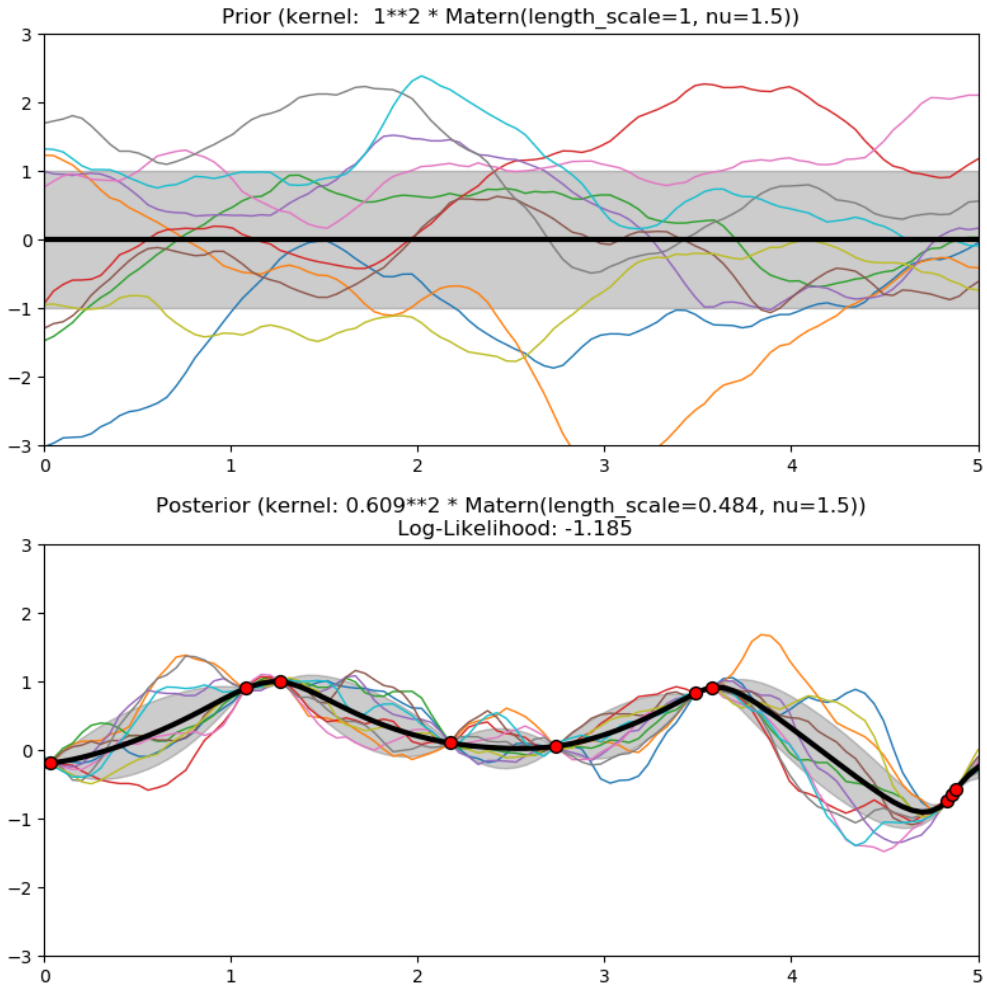
prior over all functions
with the estimated smoothness

$$\mathcal{N}(\vec{0}, \Sigma^2)$$

data

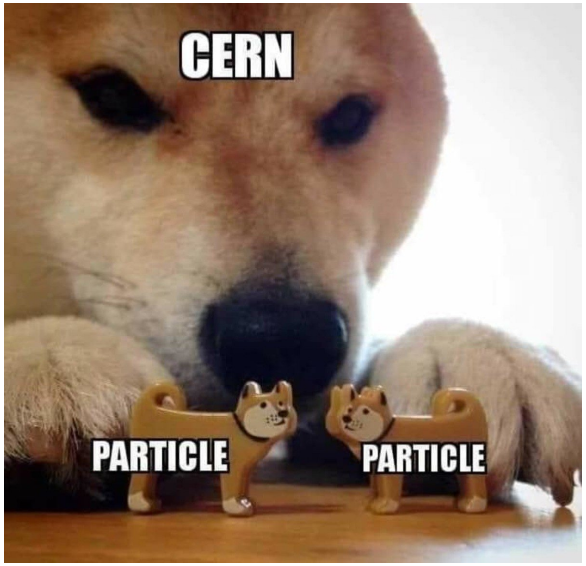
$$\mathcal{N}(\tilde{\mu}, \tilde{\Sigma}^2)$$

posterior over functions



Matérn kernel

New physics? Yes, please!



Global fits and the need for speed

Idea: consistent comparison of theories to all available data

$$\mathcal{L} = \mathcal{L}_{\text{collider}} \times \mathcal{L}_{\text{Higgs}} \times \mathcal{L}_{\text{DM}} \times \mathcal{L}_{\text{EWPO}} \times \mathcal{L}_{\text{flavour}} \times \ldots$$

- High-dimensional parameter space with varying phenomenology

  (e.g. MSSM-24)
- Quick prediction of **next-to-leading order cross sections** is crucial!
- Existing tools have drawbacks

  (Prospino: slow, NLL-fast: limited validity)

Global fits and the need for speed

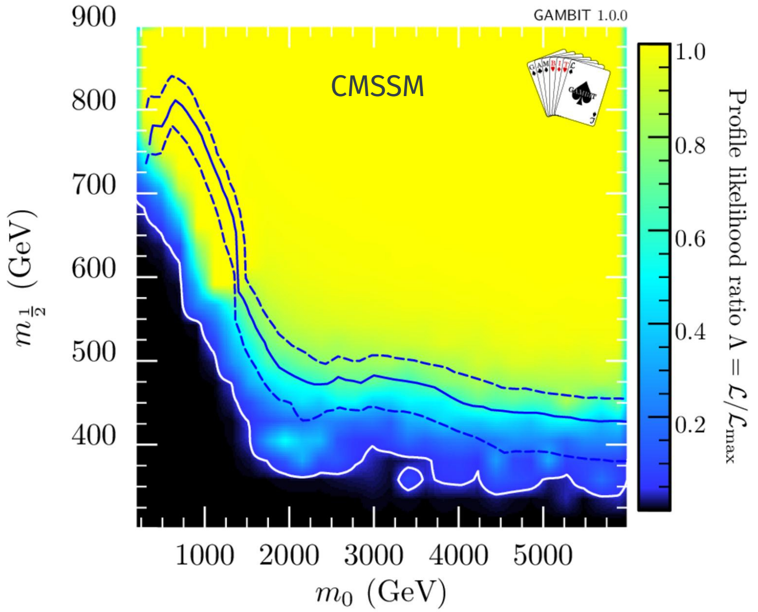
Idea: consistent comparison of theories to all available data

$$\mathcal{L} = \mathcal{L}_{\text{collider}} \times \mathcal{L}_{\text{Higgs}} \times \mathcal{L}_{\text{DM}} \times \mathcal{L}_{\text{EWPO}} \times \mathcal{L}_{\text{flavour}} \times \dots$$

- High-dimensional parameter space with varying phenomenology

  (e.g. MSSM-24)
- Quick prediction of **next-to-leading order cross sections** is crucial!
- Existing tools have drawbacks

  (Prospino: slow, NLL-fast: limited validity)



[GAMBIT, 1705.07919]

GAMBIT: The Global And Modular BSM Inference Tool

gambit.hepforge.org EPJC **77** (2017) 784 arXiv:1705.07908

- Extensive model database – not just SUSY
- Extensive observable/data libraries
- Many statistical and scanning options (Bayesian & frequentist)
- *Fast* LHC likelihood calculator
- Massively parallel
- Fully open-source
- Fast definition of new datasets and theories
- Plug and play scanning, physics and likelihood packages

Members of:
ATLAS, Belle-II, CLiC, CMS, CTA, *Fermi*-LAT, DARWIN, IceCube, LHCb, SHiP, XENON
Authors of:
DarkSUSY, DDCalc, Diver, FlexibleSUSY, gamlike, GM2Calc, IsaTols, nulike, PolyChord, Rivet, SoftSUSY, SuperISO, SUSY-AI, WIMPSim

Recent collaborators:
Peter Athron, Csaba Balázs, Ankit Beniwal, Sanjay Bloor, Torsten Bringmann, Andy Buckley, José Eliel Camargo-Molina, Marcin Chrząszcz, Jonathan Cornell, Matthias Danninger, Joakim Edsjö, Ben Farmer, Andrew Fowlie, Tomás E. Gonzalo, Will Handley, Sebastian Hoof, Selim Hotinli, Felix Kahlhoefer, Anders Kvellestad, Julia Harz, Paul Jackson, Farvah Mahmoudi, Greg Martinez, Are Raklev, Janina Renk, Chris Rogan, Roberto Ruiz de Austri, Pat Scott, Patrick Stöcker, Aaron Vincent, Christoph Weniger, Martin White, Yang Zhang

40+ participants in 11 experiments and 14 major theory codes

Method

Pre-trained, distributed Gaussian processes

Goal

Fast estimate of SUSY (strong) production cross sections at NLO, and uncertainties from

- regression itself
- renormalisation scale
- PDF variation
- α_s variation

XSEC

the cross-section evaluation code

A. Buckley, I. A. V. Holm, A. Kvellestad, A. Raklev, P. Scott, J. V. Sparre, *J. Van den Abeele*

Interface

```
* Requested processes, in order:
   (1000021, 1000021)
* Input features:
   (1000021, 1000021) :
      [m1000021 = 1000.0, m2000004 = 500.0, m2000003 = 500.0,
m2000002 = 500.0, m2000001 = 500.0, m1000004 = 500.0, m1000003
 = 500.0, m1000002 = 500.0, m1000001 = 500.0, mean = 500.0]
* xsection_central (fb): [224.8033]
* regdown_rel:    [0.9177]
* regup_rel:      [1.0823]
* scaledown_rel:  [0.8751]
* scaleup_rel:    [1.0934]
* pdfdown_rel:    [0.9005]
* pdfup_rel:      [1.0995]
* alphasdown_rel: [0.8908]
* alphasup_rel:   [1.1092]
***************************************************************
```

Stand-alone Python code, also implemented in GAMBIT

Soon public on GitHub!

Processes

$$pp \to \tilde{g}\tilde{g},\ \tilde{g}\tilde{q}_i,\ \tilde{q}_i\tilde{q}_j,\ \tilde{q}_i\tilde{q}_j^*,\ \tilde{b}_i\tilde{b}_i^*,\ \tilde{t}_i\tilde{t}_i^*$$

at $\sqrt{s} = 7/8/13/14$ TeV

Workflow

Generating data

Random sampling
▼
SUSY spectrum
▼
Cross sections

▶

Training GPs

Optimise kernel hyperparameters
▼
Compute covariances

▶

GP predictions

Input parameters
▼
Linear algebra
▼
Cross section estimates

Workflow

Generating data

Random sampling
▼
SUSY spectrum
▼
Cross sections

Training GPs

Optimise kernel hyperparameters
▼
Compute covariances

GP predictions

Input parameters
▼
Linear algebra
▼
Cross section estimates

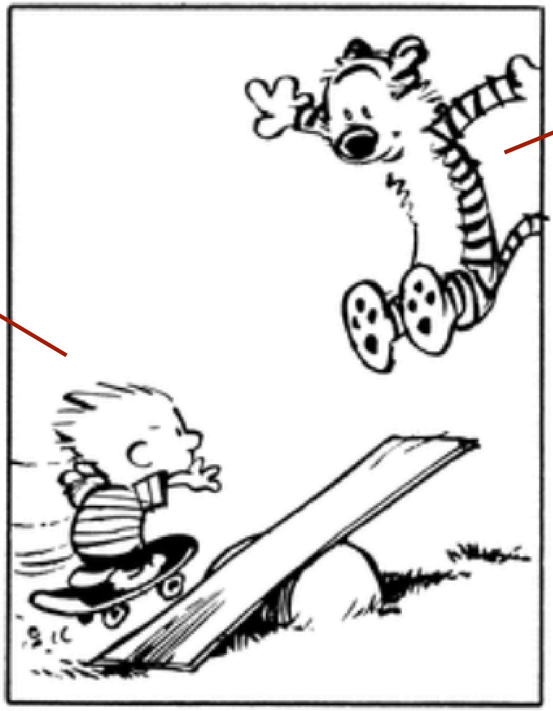
XSEC

A balancing act

Evaluation speed

Training scales as $\mathcal{O}(n^3)$, prediction as $\mathcal{O}(n^2)$

Distributed Gaussian processes



Sample coverage

Need to cover a large parameter space

Mix of random samples with different priors in mass space

A balancing act

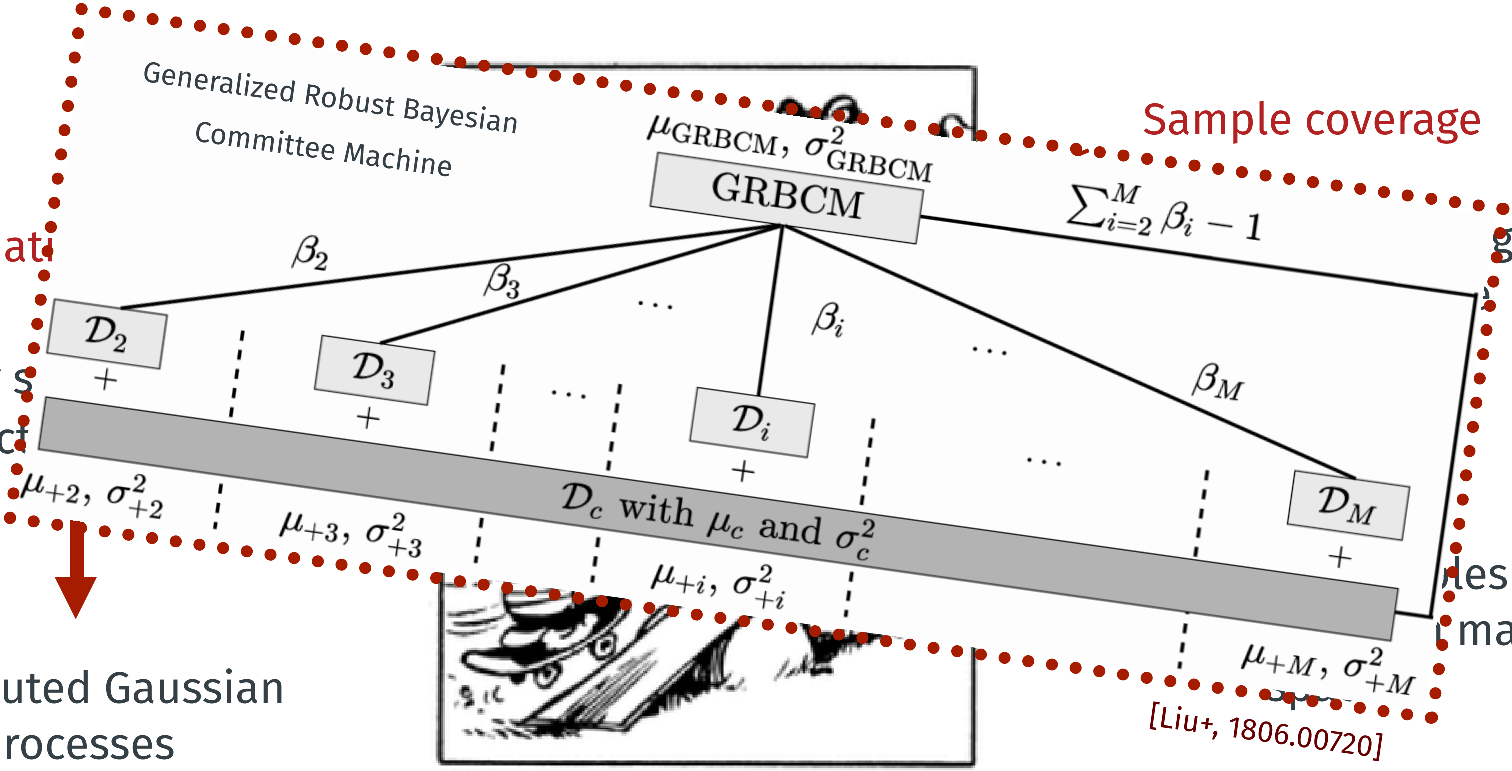
Sample coverage

Evaluat

ge

Training s
predict

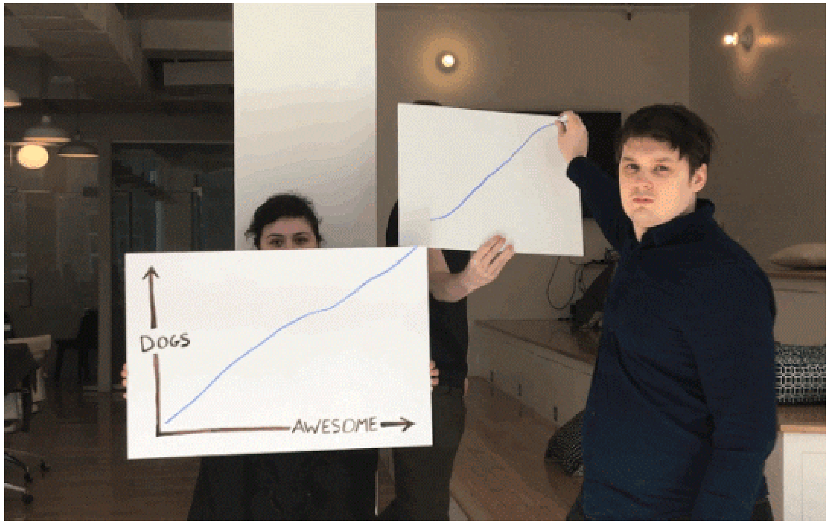
Distributed Gaussian processes



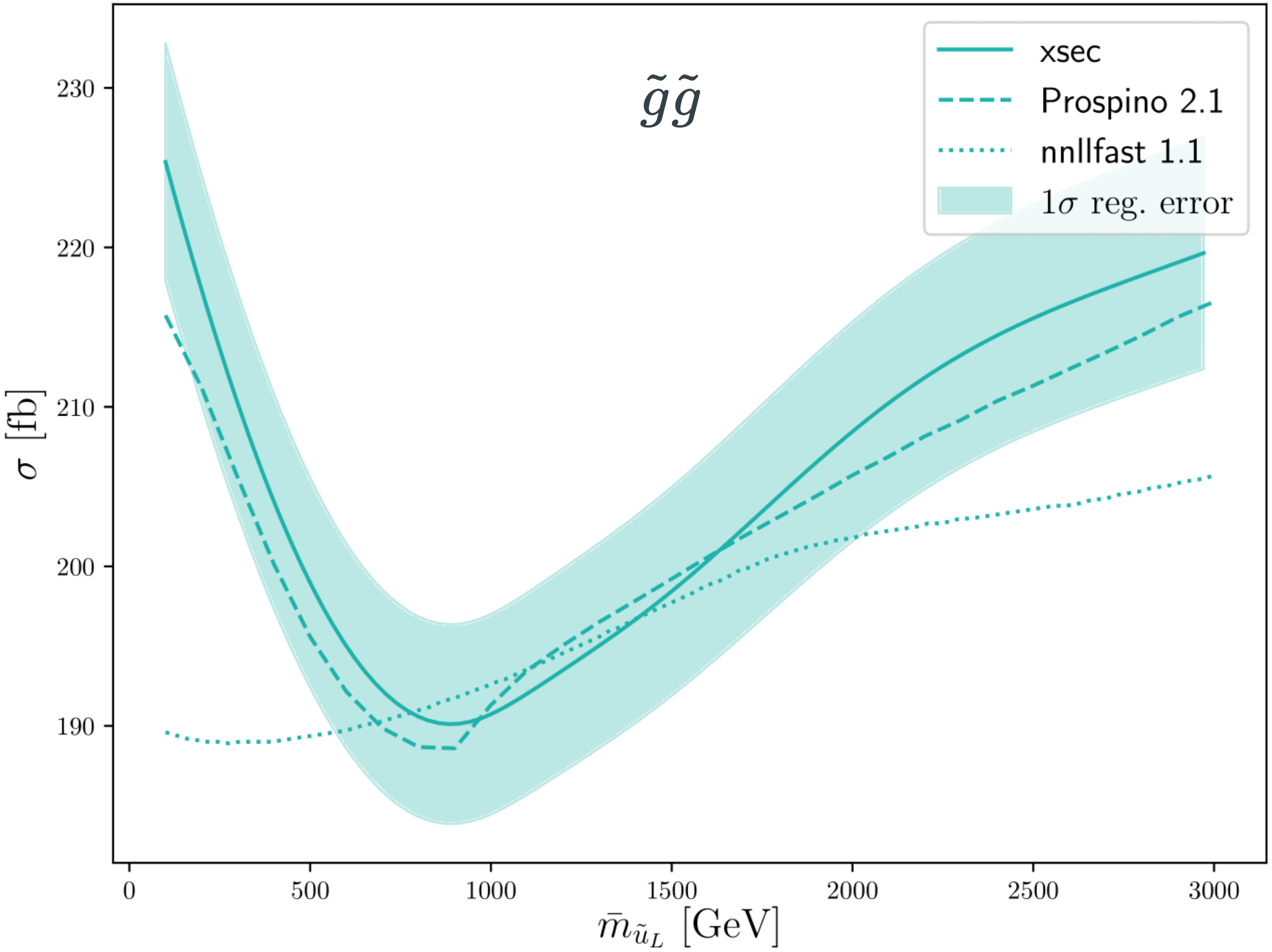
les with
mass

[Liu+, 1806.00720]

Plots, plots, plots

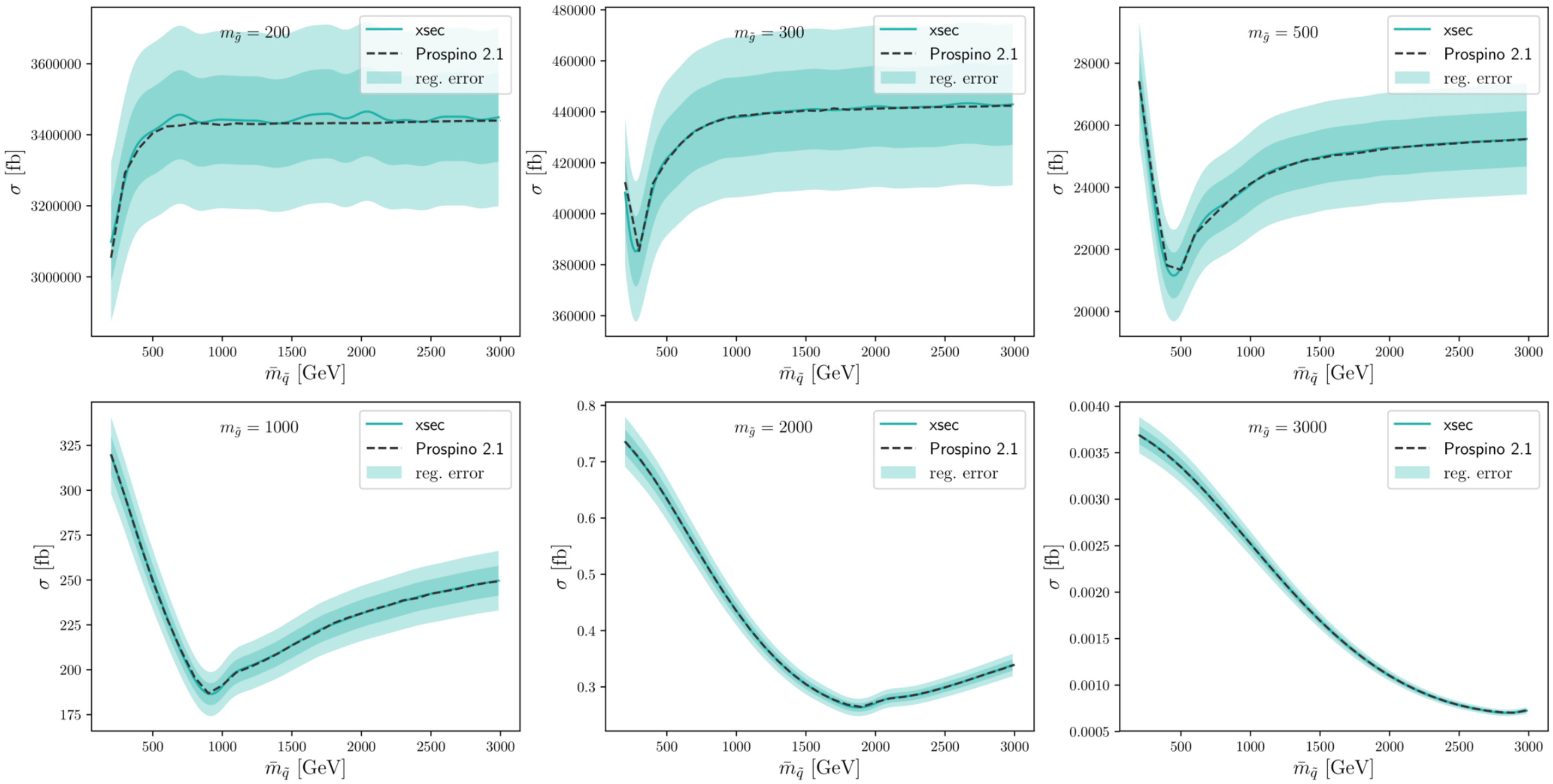


XSEC: the cross-section evaluation code



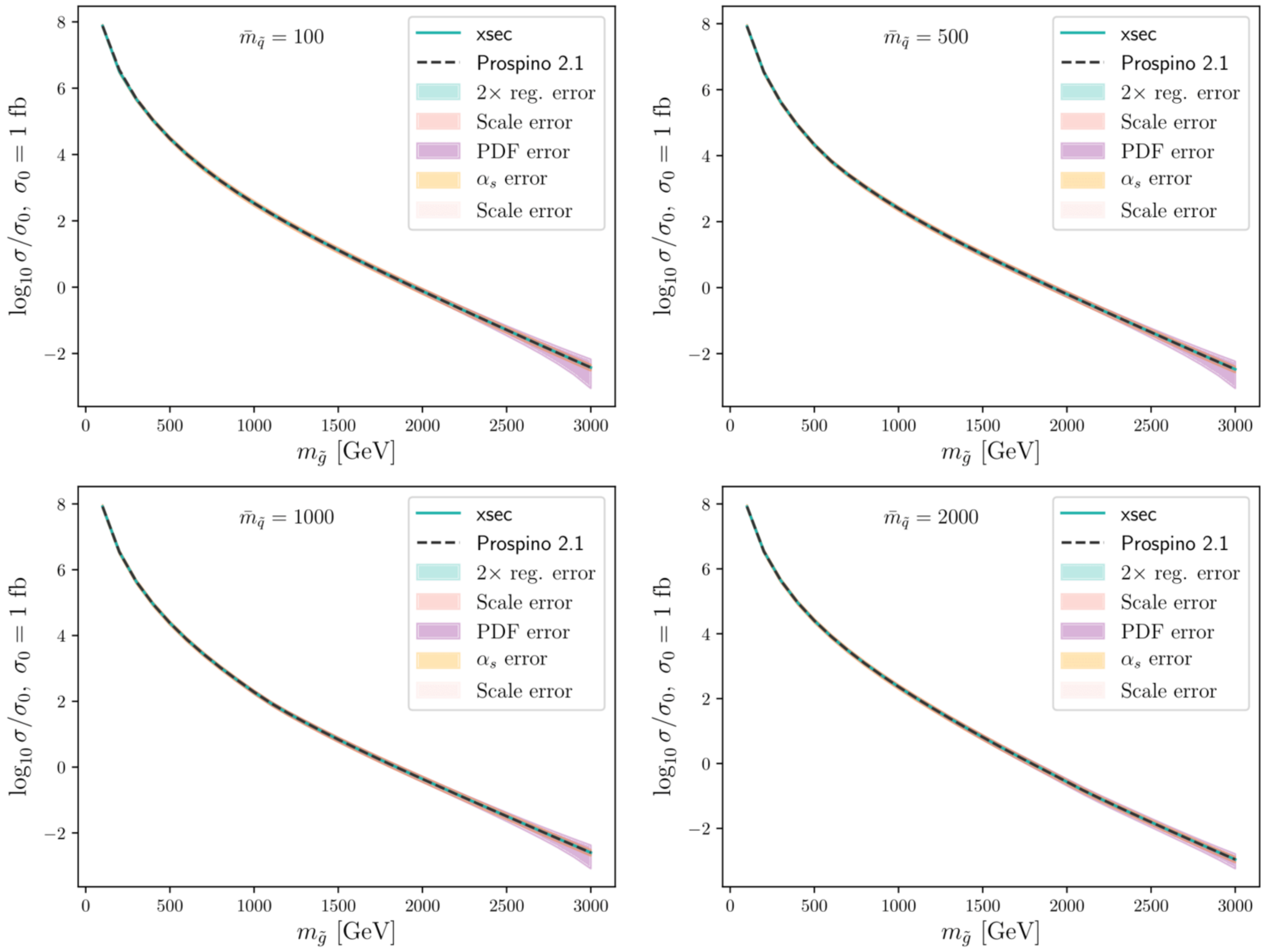
XSEC: the cross-section evaluation code

$\tilde{g}\tilde{g}$

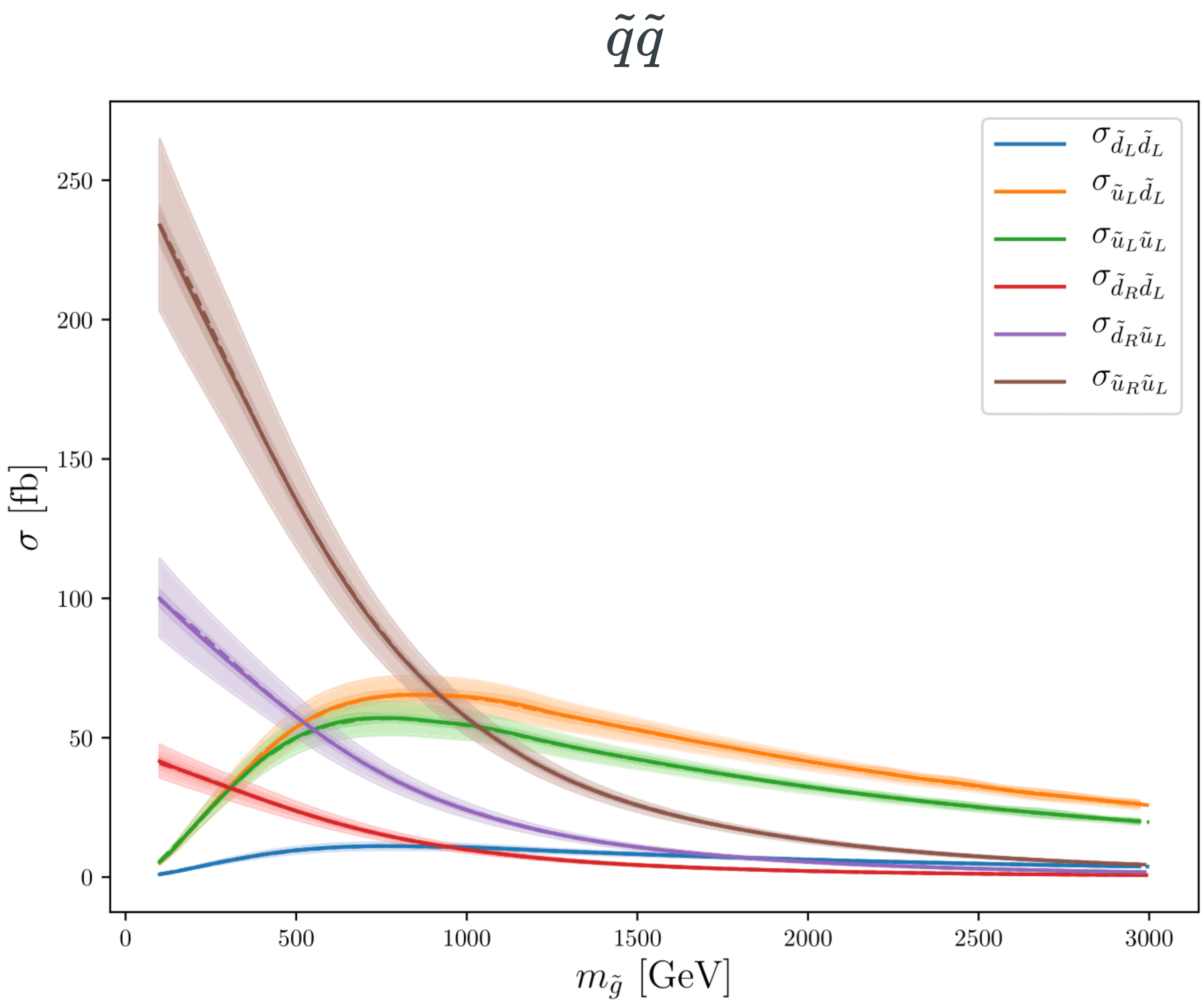


XSEC: the cross-section evaluation code

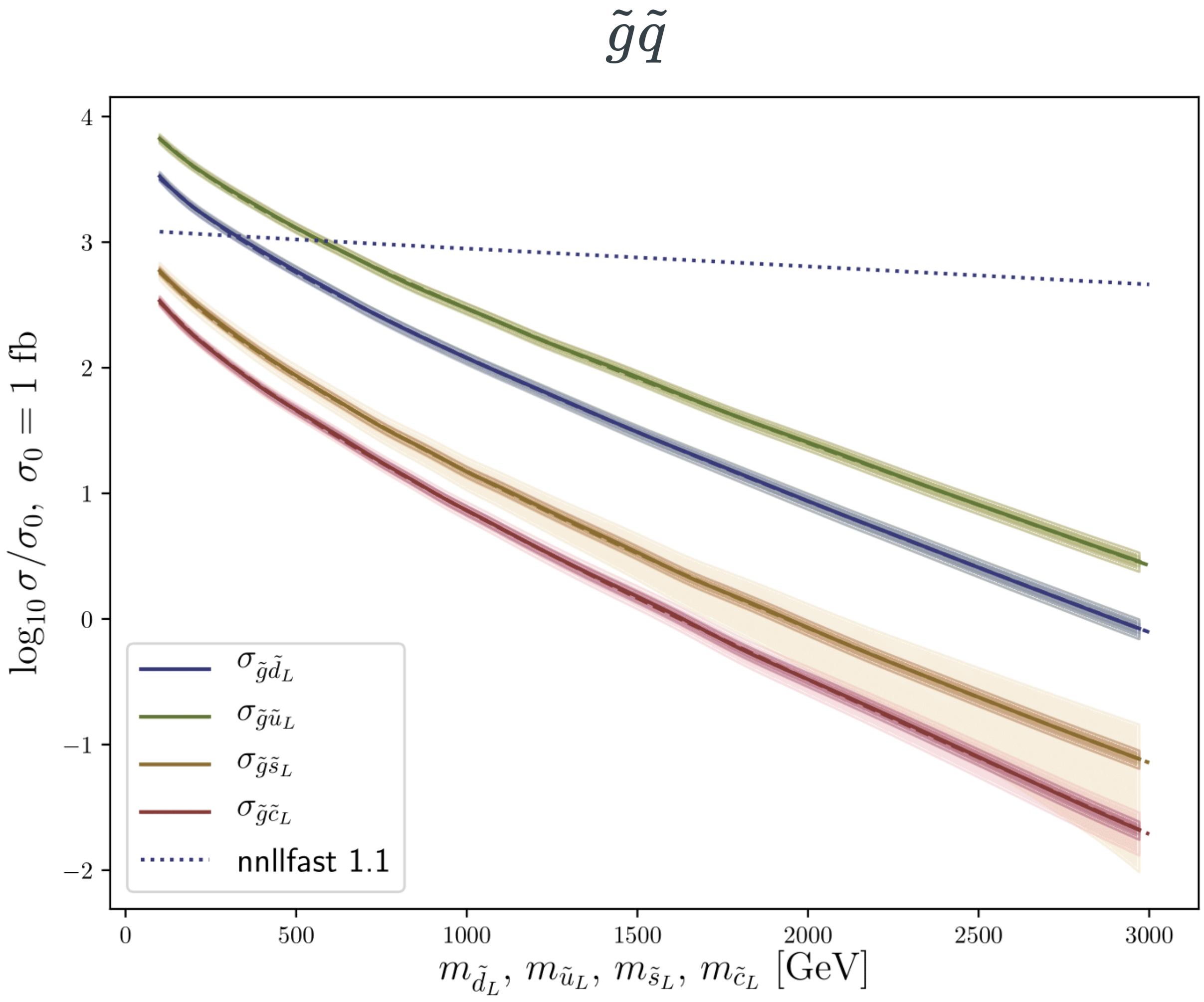
$\tilde{g}\tilde{g}$



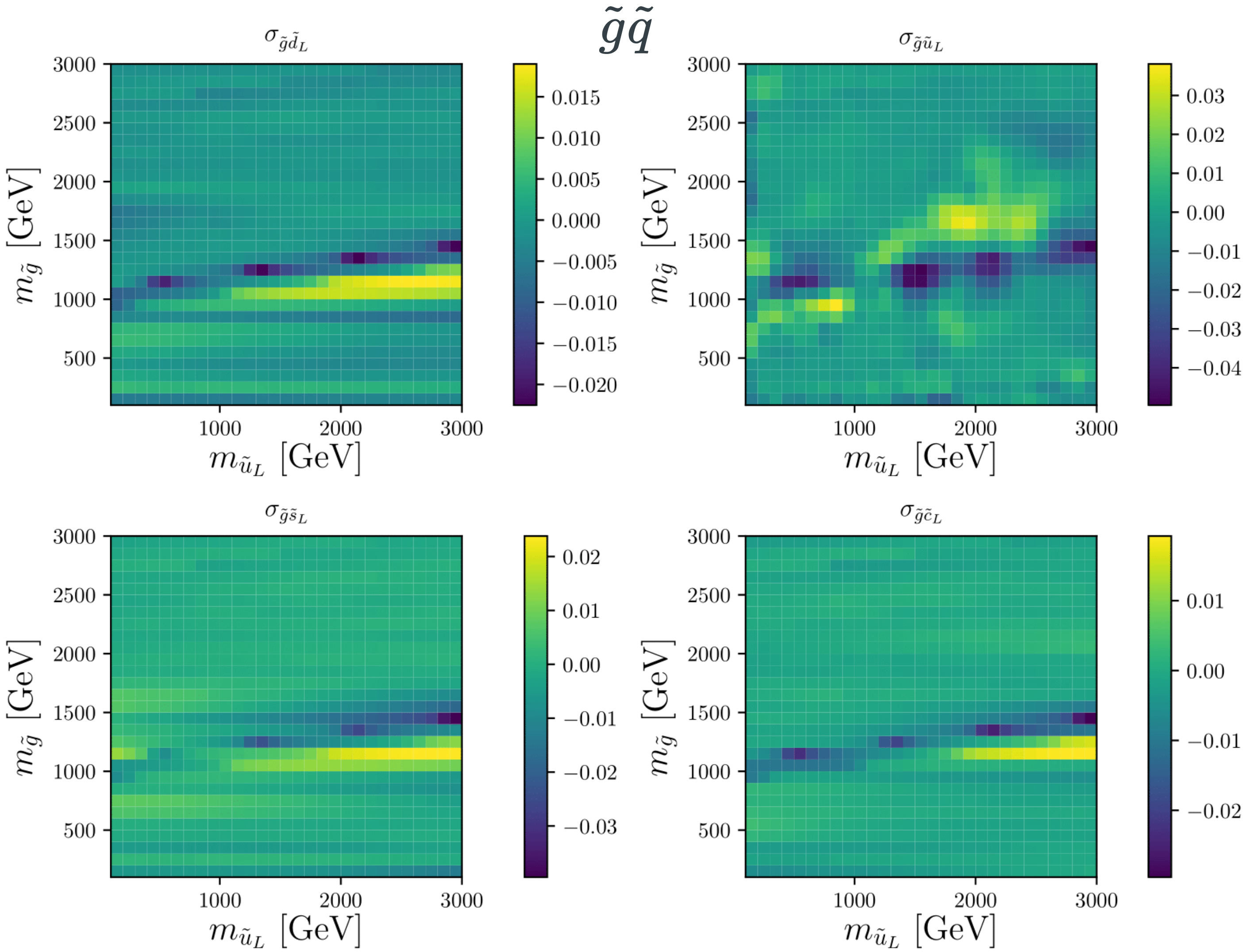
XSEC: the cross-section evaluation code



XSEC: the cross-section evaluation code



XSEC: the cross-section evaluation code



Thank you!

Backup slides

Gaussian Processes 101 (non-parametric regression)

Consider the data as a sample from a multivariate **Gaussian distribution**.

The covariance matrix controls smoothness.
Assume it is given by a **kernel function**, like

$$k(x_1, x_2) = \exp\left(-\frac{|x_2 - x_1|^2}{2l^2}\right).$$

Bayesian approach to estimate $y_* = f(x_*)$:

prior over functions

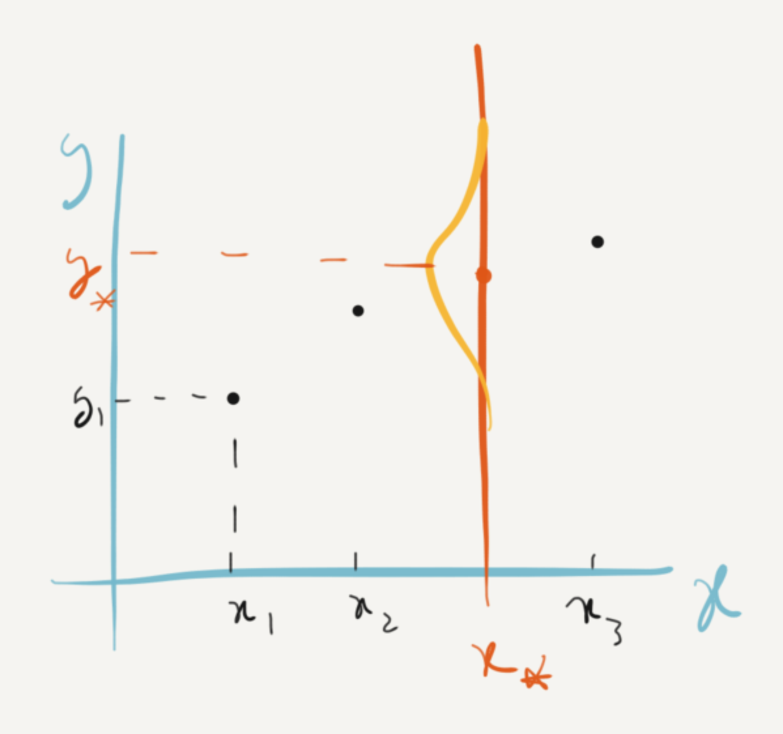
$$\begin{pmatrix} \mathbf{y} \\ y_* \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mathbf{0} \\ 0 \end{pmatrix}, \begin{bmatrix} K & K_* \\ K_*^T & K_{**} \end{bmatrix}\right)$$

$$K_{ij} = k(x_i, x_j)$$
$$K_{*i} = k(x_*, x_i)$$
$$K_{**} = k(x_*, x_*)$$

data

posterior over functions

$$y_* | \mathbf{y} \sim \mathcal{N}(\underbrace{K_* K^{-1} \mathbf{y}}_{\text{mean}},\ \underbrace{K_{**} - K_* K^{-1} K_*^T}_{\text{covariance}})$$



Gaussian Processes 101 (non-parametric regression)

Consider the data as a sample from a multivariate **Gaussian distribution**.

The covariance matrix controls smoothness.

Assume it is given by a **kernel function**, like

$$k(x_1, x_2) = \exp\left(-\frac{|x_2 - x_1|^2}{2l^2}\right).$$

Bayesian approach to estimate $y_* = f(x_*)$:

prior over functions

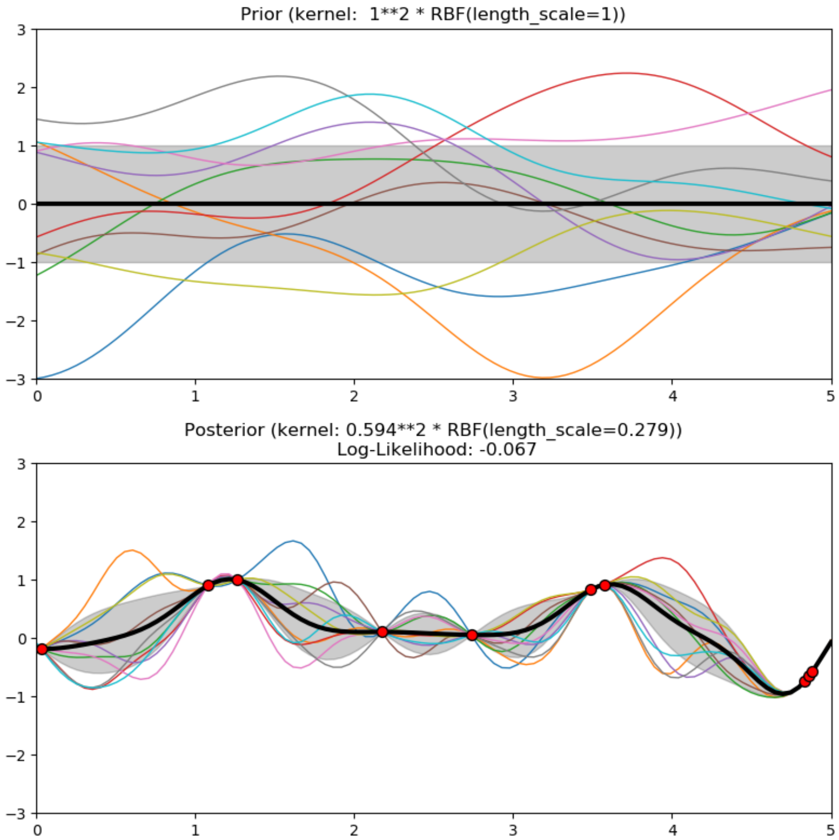
$$\begin{pmatrix} \mathbf{y} \\ y_* \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mathbf{0} \\ 0 \end{pmatrix}, \begin{bmatrix} K & K_* \\ K_*^T & K_{**} \end{bmatrix}\right)$$

$$K_{ij} = k(x_i, x_j)$$
$$K_{*i} = k(x_*, x_i)$$
$$K_{**} = k(x_*, x_*)$$

data

posterior over functions

$$y_* | \mathbf{y} \sim \mathcal{N}(\underbrace{K_* K^{-1} \mathbf{y}}_{\text{mean}},\ \underbrace{K_{**} - K_* K^{-1} K_*^T}_{\text{covariance}})$$



Radial Basis Function kernel

Gaussian Processes 101 (non-parametric regression)

Consider the data as a sample from a multivariate **Gaussian distribution**.

The covariance matrix controls smoothness.

Assume it is given by a **kernel function**, like

$$k(x_1, x_2) = \exp\left(-\frac{|x_2 - x_1|^2}{2l^2}\right).$$

Bayesian approach to estimate $y_* = f(x_*)$:

prior over functions

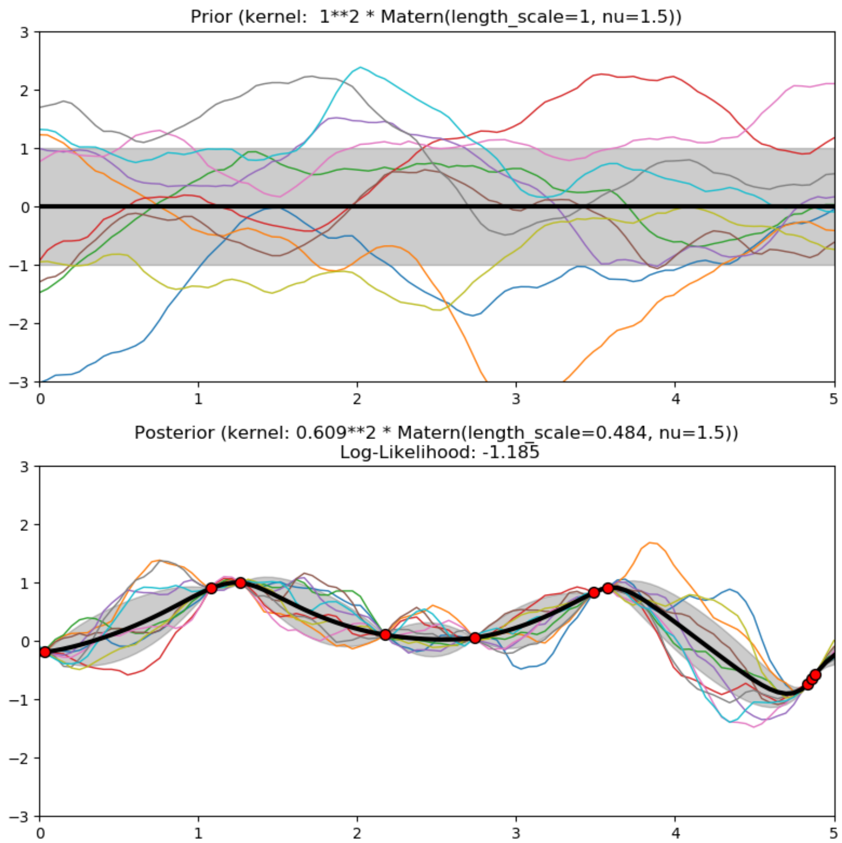
$$\begin{pmatrix} \mathbf{y} \\ y_* \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mathbf{0} \\ 0 \end{pmatrix}, \begin{bmatrix} K & K_* \\ K_*^T & K_{**} \end{bmatrix}\right)$$

$$K_{ij} = k(x_i, x_j)$$
$$K_{*i} = k(x_*, x_i)$$
$$K_{**} = k(x_*, x_*)$$

data

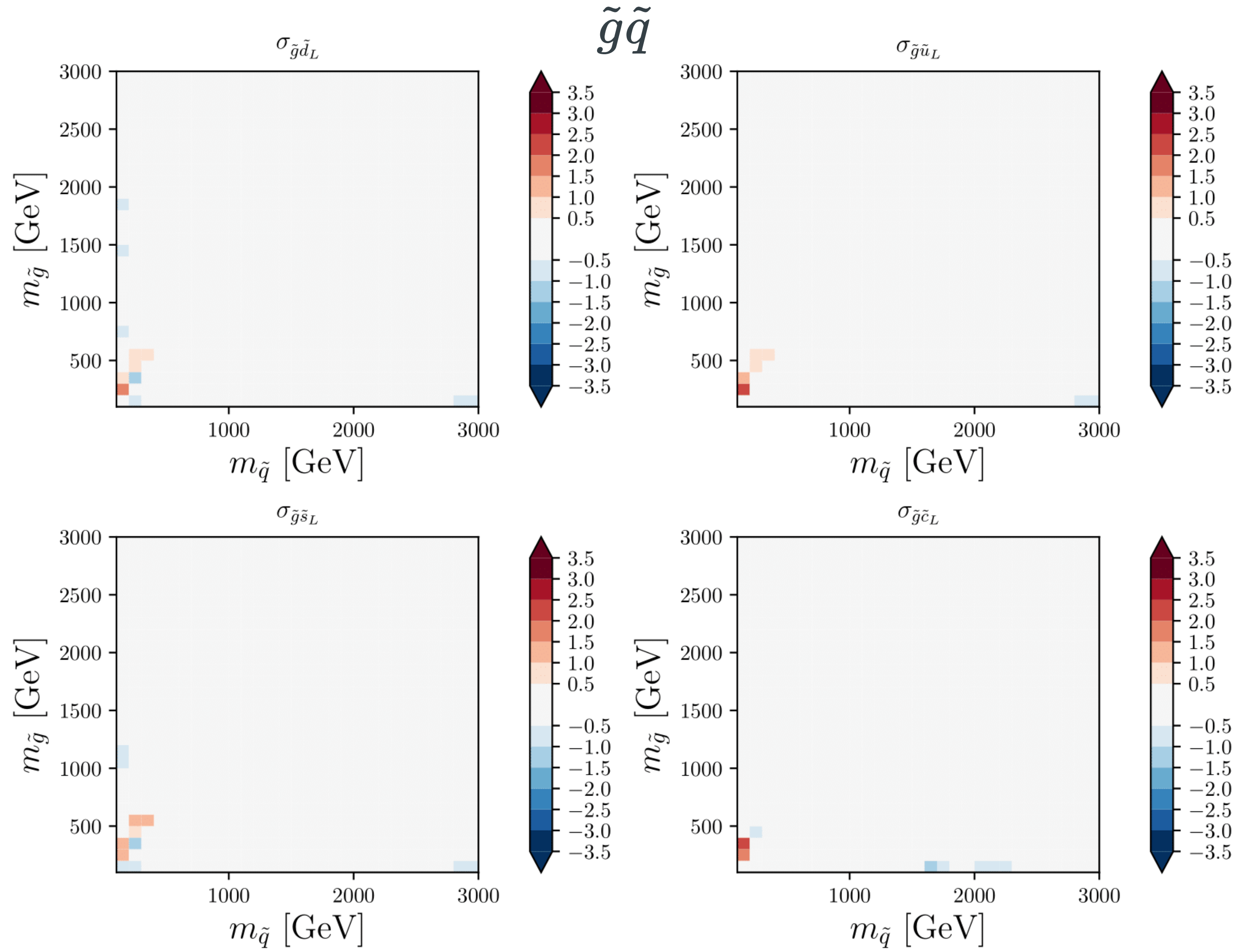
posterior over functions

$$y_* | \mathbf{y} \sim \mathcal{N}(\underbrace{K_* K^{-1} \mathbf{y}}_{\text{mean}},\ \underbrace{K_{**} - K_* K^{-1} K_*^T}_{\text{covariance}})$$



Matérn kernel

XSEC: the cross section evaluation code



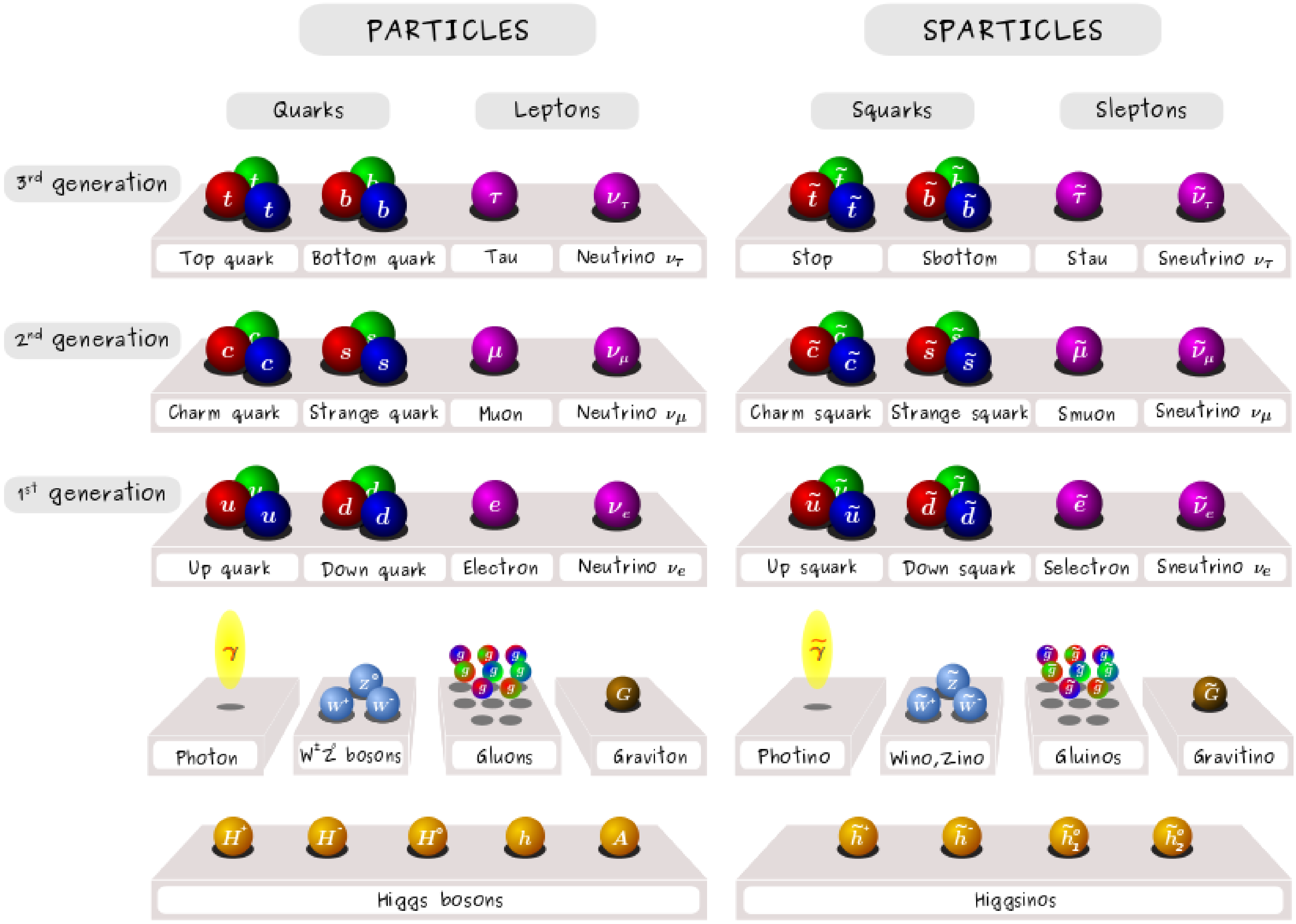


Image: Claire David (CERN)

Scale-dependence of LO/NLO

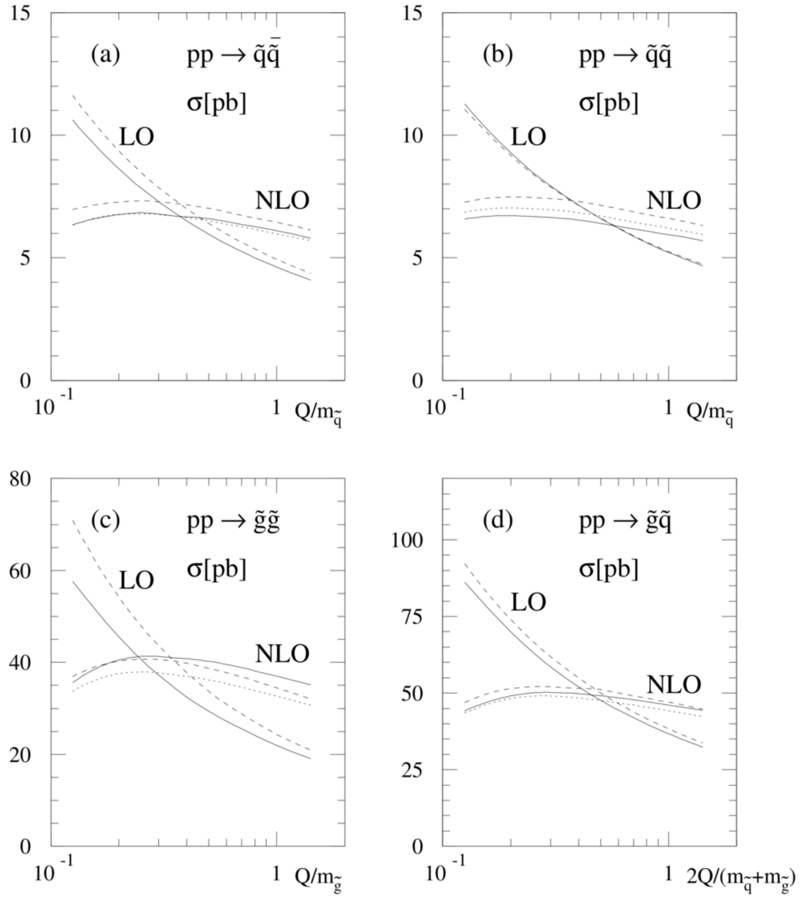


Figure 14: The dependence on the renormalization/factorization scale Q of the LO and NLO cross-sections for (a) squark–antisquark, (b) squark–squark, (c) gluino–gluino, and (d) squark–gluino production at the LHC ($\sqrt{S}$ = 14 TeV). Parton densities: GRV94 (solid), CTEQ3 (dashed), and MRS(A') (dotted). Mass parameters: $m_{\tilde{q}}$ = 600 GeV, $m_{\tilde{g}}$ = 500 GeV, and m_t = 175 GeV.

[Beenakker+, hep-ph/9610490]