RPV Single Stop Search Update

Non-parametric 2D Background Estimation Using Gaussian Processes

Charlie Kapsiak

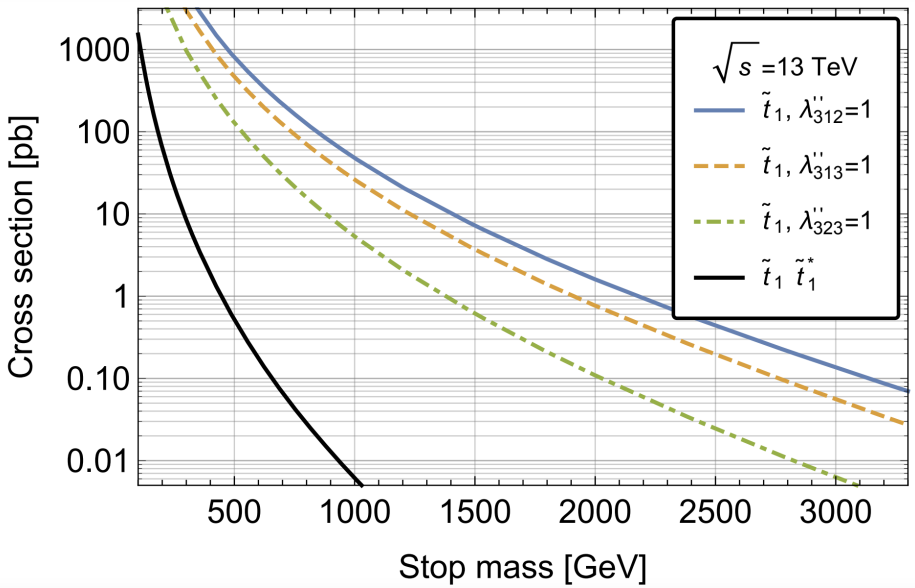
2024-05-01

Table Of Contents

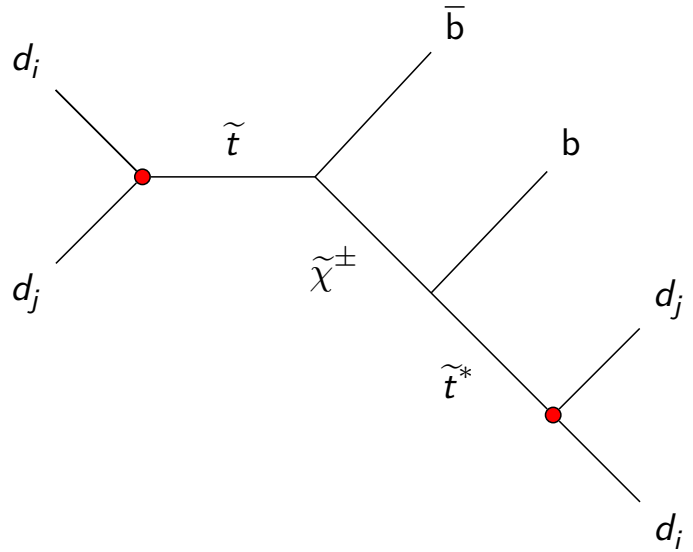
Intro

Analysis Target Model

- Searching for then production and decay of a single $\tilde{t}$ to 4 standard model quarks through an RPV coupling.
- Well motivated channel to look for SUSY:
  - Unexplored region of RPV parameter space
  - Large cross section allows us to probe higher masses



Source: [3]



Analysis Status

The past months have seen substantial progress on several fronts. We summarize below the current major areas of work:

Key Analysis Elements

- Control/Signal region definitions.
- Mass reconstruction.
- Background estimation
- Statistical analysis procedure.
- Trigger studies.
- Central MC production.
- Early Run3 data.

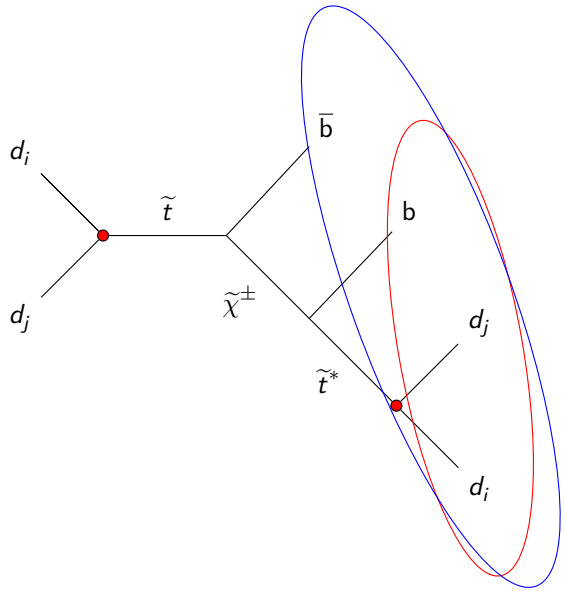
Past Presentations

- 2023-10-06: Introduction Talk
- Background Estimation

● No Work ● Preliminary ● Working Version ● Analysis Ready

General Search Strategy

- General search strategy is to perform a one or two dimensional bump hunt for both the $\widetilde{t}$ and the $\widetilde{\chi}^{\pm}$ resonances.
- For many mass splittings, the resonances are well separated both in $m_{\widetilde{t},reco}$ and $m_{\widetilde{\chi}^{\pm},reco}$ space, providing additional discriminating power.
- Key point is to effectively estimate the background.
- However, a simple cut strategy on one mass axis can result in sculpting of the background, making estimation difficult.



Estimation Strategies

- For all bump hunts, key technique is estimation of the background shape.
- Traditional bump hunts have used ad-hoc functions, chosen because they approximate the observed shape.
- However, this can introduce bias from the choice of function, and it has been shown that they scale poorly with increasing luminosity [1].
- For multidimensional searches, the problem can also be compounded by selecting a function for a potentially nontrivial 2D shape.

Current Strategy

- We have implemented our background estimation using Gaussian process regression [3].
- This is non-parametric technique that reduces bias from the choice of parametric function.
- It has been shown to be robust against increasing luminosity[1].
- It is naturally extensible to multiple dimensions.
- Very well studied in statistics literature, and has a large number of well established implementations [2].

GP Regression

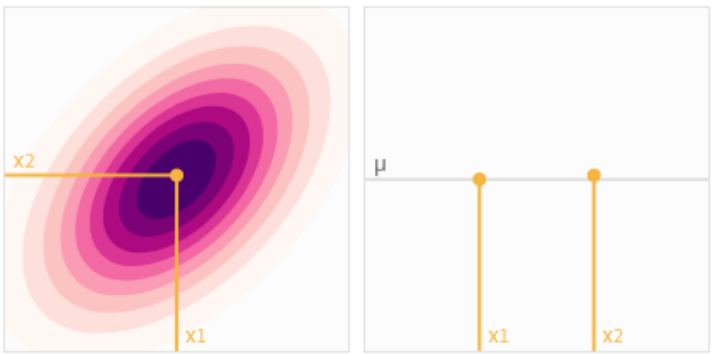
What is a Gaussian Process?

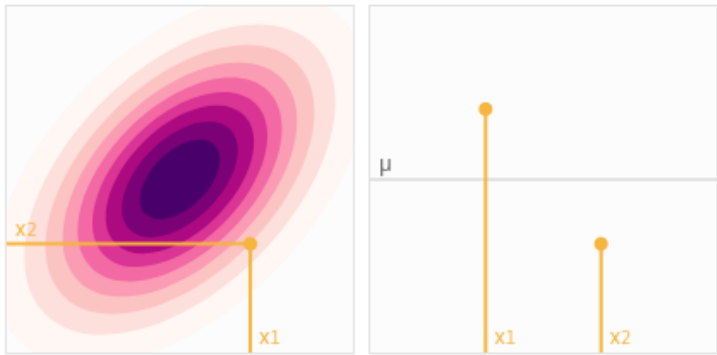
Definition

A gaussian process is a possibly infinite series of random variables, any finite subset of which is jointly gaussian.

Generall, the random variables are indexed by real values x, since we are generally considering regression over $\mathbb{R}^n$.
A gaussian process $f(x)$ is is completely defined by its mean and covariance

$$\begin{aligned} m(x) &= \mathbb{E}\left[f(x)\right] \\ k(x, x') &= \mathbb{E}\left[\left(f(x) - m(x)\right)\left(f(x') - m(x')\right)\right] \end{aligned} \tag{1}$$





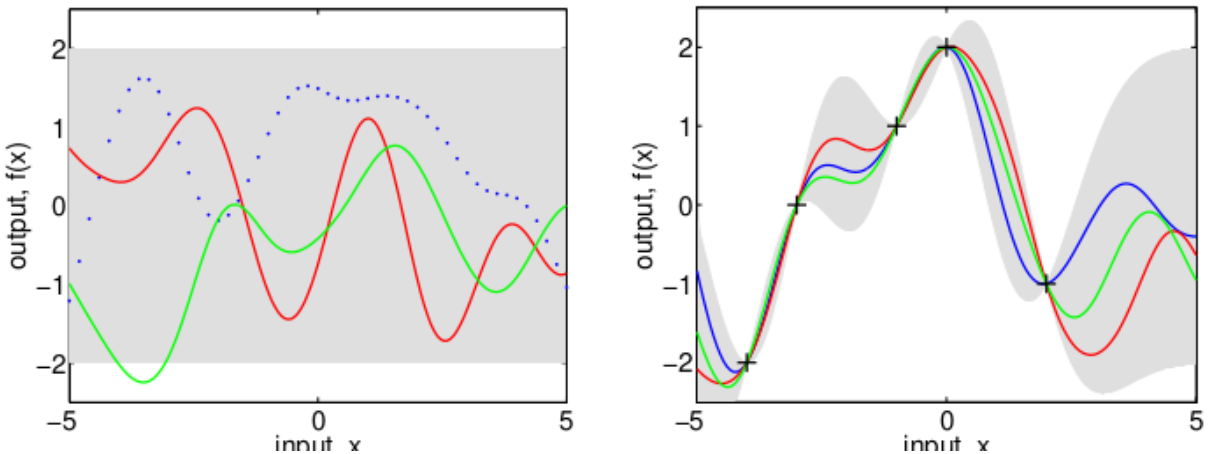
Function Distributions

Gaussian process allow us to define distributions over the space of functions. Given a gaussian process $\mathcal{GP}\left(m(x), k(x, x')\right)$, and some function h, then

$$\mathbf{h} \sim N(m(X), k(X, X)) \tag{2}$$

Given n points in $\mathbb{R}^k$, the gaussian process defined a n dimensional multivariate gaussian $\mathcal{N}$. If a function $h(x)$ has has values $h_1, h_2, ..., h_n$ at those points, then

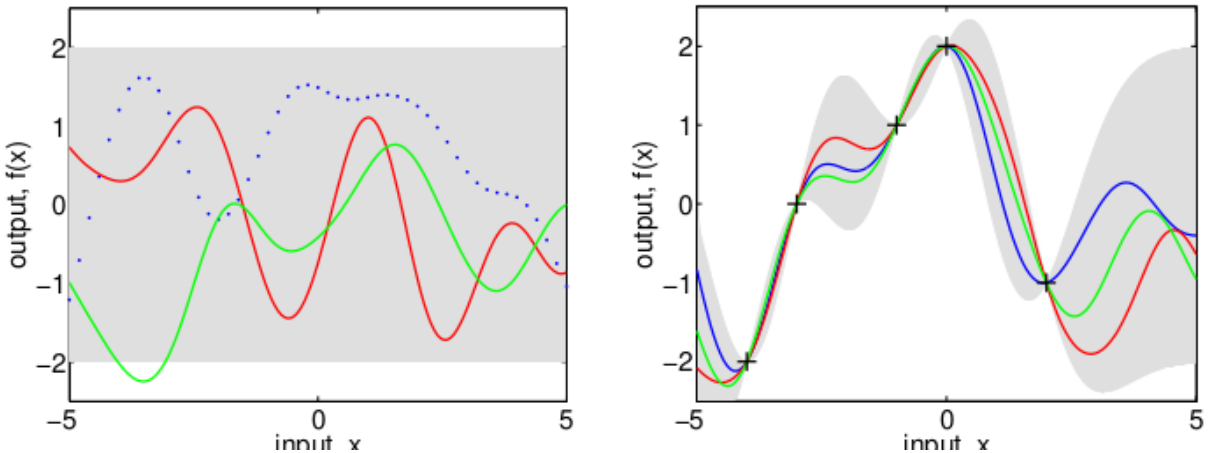
$$p\left(h\right) \sim \mathcal{N}(h_1, ..., h_n) \tag{3}$$



Source: [3]

Gaussian Process Regression

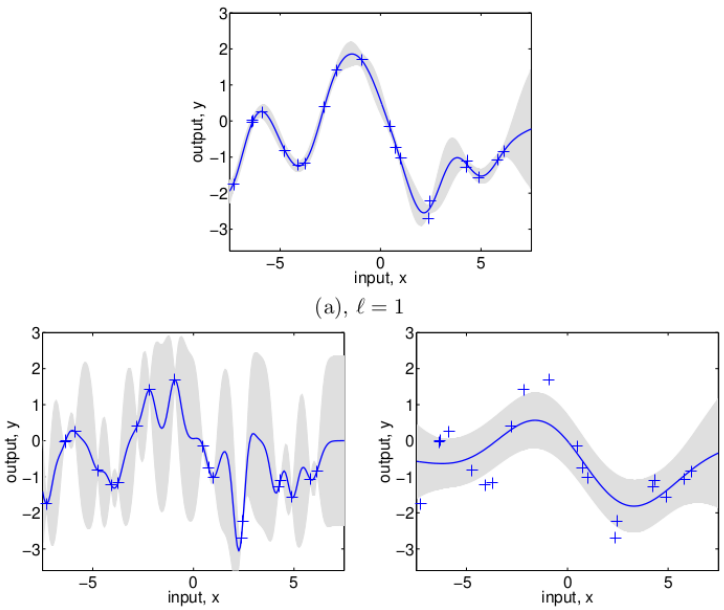
- The ability to define distributions over functions allows us to do inference using Baye's theorem.
- Specifically, given N training points and a Gaussian process prior, we can produce a posterior Gaussian process that provides a means to do regression.



Source: [3]

Kernels

- The choice of kernel is the most important aspect of Gaussian processes.
- Different kernels reflect different understandings of how the points should be correlated, how smooth the process the should be, etc.
- There is a great deal of literature on kernels and kernel selection. Some common categories:
  - Simple functions, like the square exponential or Mátern kernel.
  - Spectral mixtures
  - Deep learning kernels

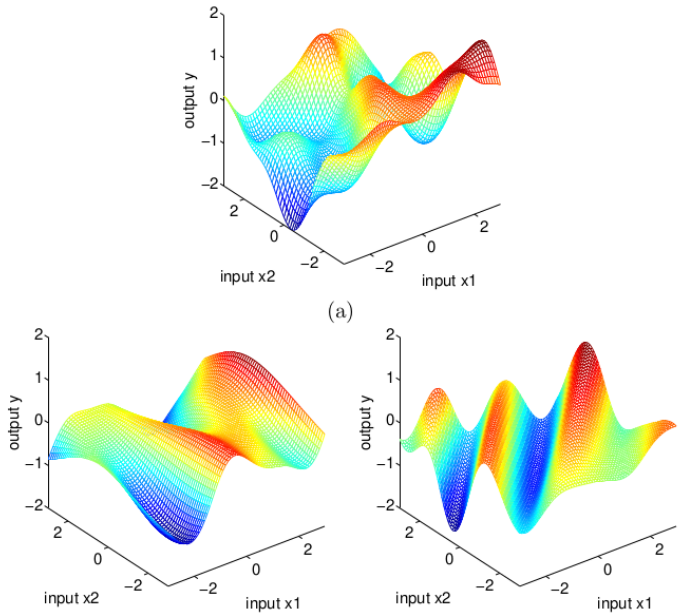


Source: [3]

2D Regression

Kernels

- The machinery for Gaussian processes extends naturally to multi-dimensional problems. In our case, we examine two dimensions.
- In two dimenions, the kernel can have an even richer structure.



Source: [3]

Kernels in 2D

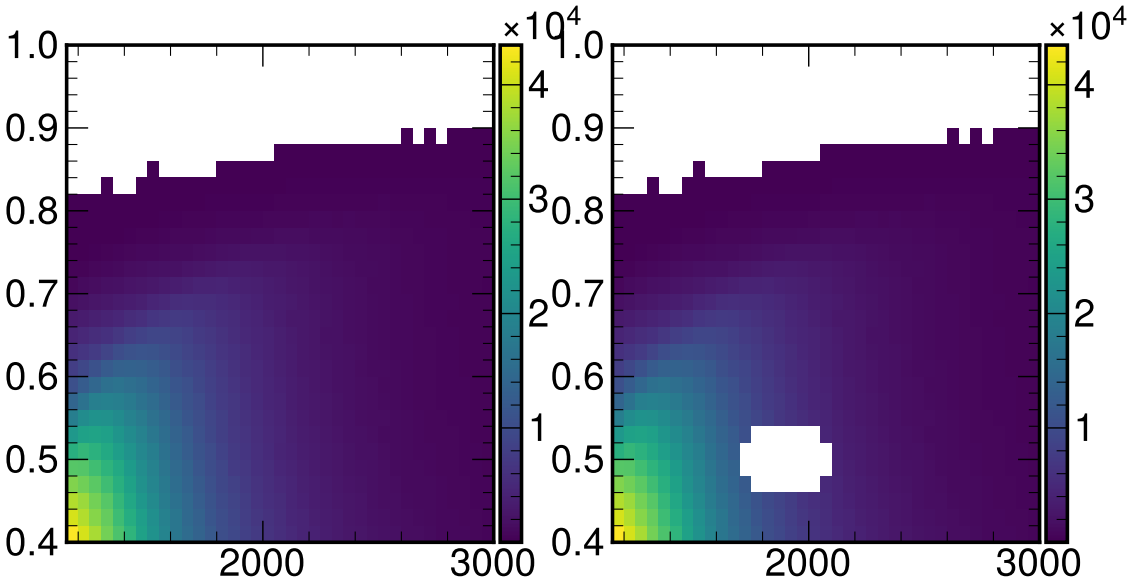
- Show simplest RBF Kernel
- Show GRBF kernel
- Show NN kernel

SHOW IMAGES HERE OF WHAT THE DIFFERENT KERNELS LOOK LIKE

Results

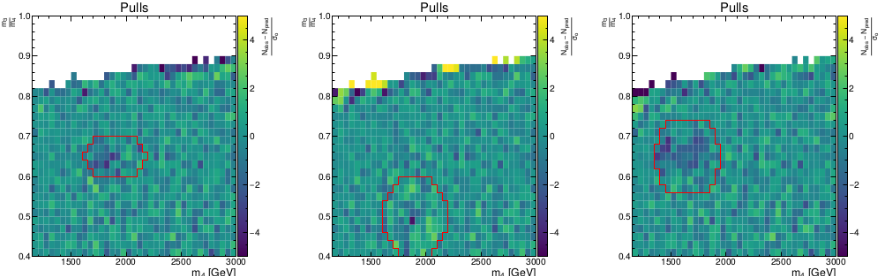
Overview

- We use GP regression to estimate backgrounds in both one and two dimensional searches.
- In both cases, we mask the region of space where the signal is expected, then use regression to estimate the background in the masked region.
- Use both simulation and control region data to study efficacy of different methods.
- Majority of studies have been focused on kernel selection and approximation techniques.

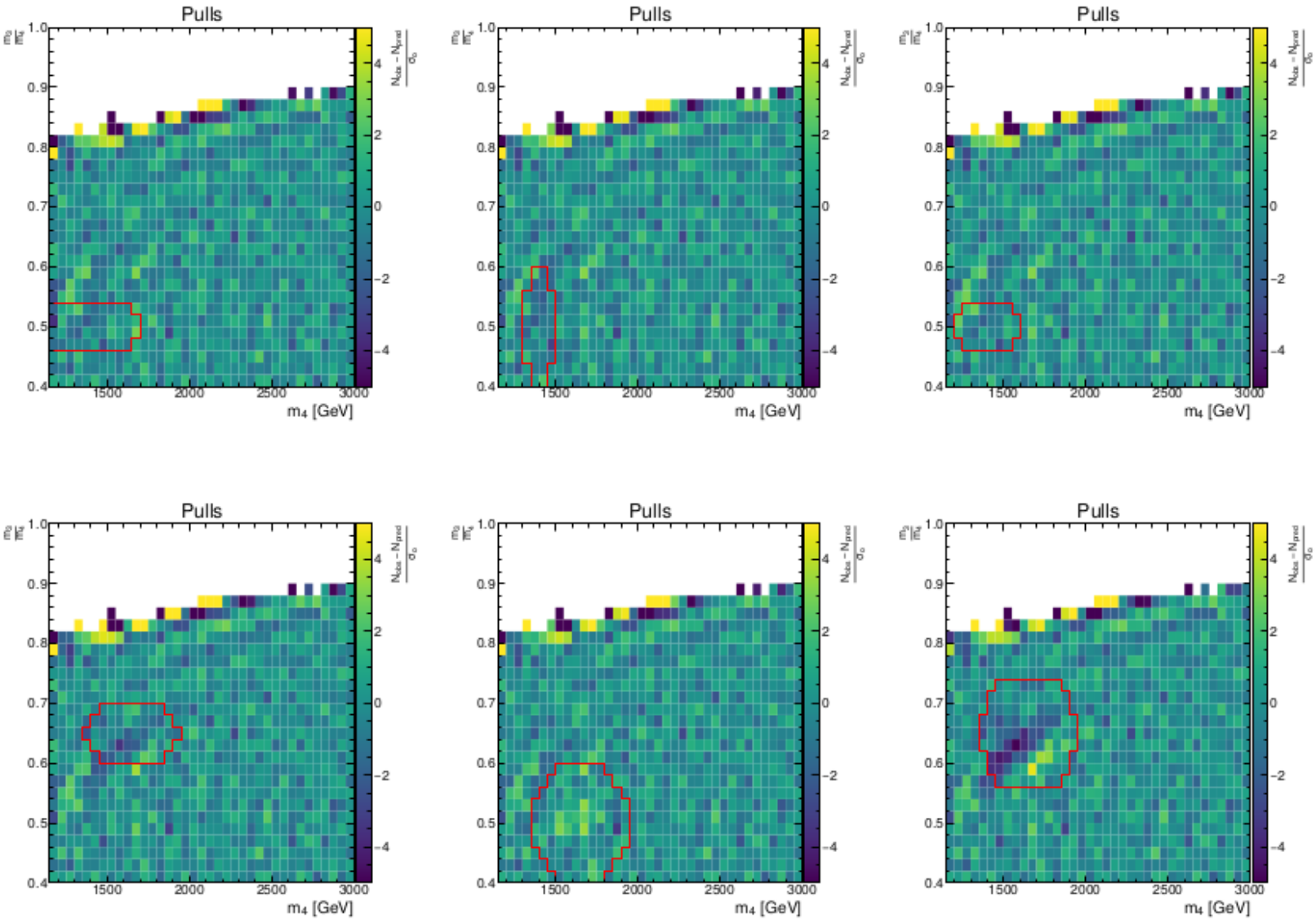


2D Results

- See promising results for estimation in windows of varying sizes over the 2D plane.
- Depending on region, either generalized RBF kernels or and RBF kernel supplemented with the deep network have shown promising and robust estimative abilities.



General RBF Fit Results



Statistical Considerations

Notes on Statistical Procedure

- Gaussian process regression provides a complete posterior distribution describing the background.
- Therefore, a proper statistical treatment requires considering not just the posterior mean, but the complete distribution.
  - We are working on an implementation in combine, using the eigenvectors of the posterior covariance as nuisance parameters templates.
  - MCMC/SVI can be externally using packages like pyro or emcee to build the statistical model directly in python.

MAYBE IMAGE FROM COMBINE

Systematic Uncertainties and Validation

- We envision a validation strategy where the kernel hyperparameters are trained on SR data, then the regression is validated by examining the fit quality in simulation and CR data.
- More work needed to determine exactly what systematics should be considered, and how they should be implemented.

SHOW COMPARISON OF CR AND MC

Conclusion

Related Ongoing Work

- Unify kernels for different mass points. We believe this will be relatively straightforward since the GRBF kernel lies in the space of deep kernels.
- Finalize Combine implementation and perform validation.
- Work on exact definitions of systematic uncertainties.
-

Conclusion

- The past months have seen substantial progress on the background estimation and statistical analysis procedure.
- Framework can now produce good estimates for 2D backgrounds over a range of locations and masking windows.
- We hope to hear any feedback from experts regarding the methodology, or any comments or suggestions!

Thank you!

Bibliography


Meghan Frate and Et al.
Modeling Smooth Backgrounds and Generic Localized Signals with Gaussian Processes, September 2017.


Jacob R. Gardner and Et al.
GPyTorch: Blackbox Matrix-Matrix Gaussian Process Inference with GPU Acceleration, June 2021.


Carl Edward Rasmussen and Christopher K. I. Williams.
Gaussian Processes for Machine Learning.
Adaptive Computation and Machine Learning. MIT Press, Cambridge, Mass, 2006.

Appendix

Luminosity Issues

Table of Common Kernels