Search for Triphoton Resonances in Boosted Final States with NNs and GPs

Austin Townsend[1], Marc Osherson[1], Mike Hildreth[1], Stephen Clark[2]
University of Notre Dame[1], Rutgers University[2]

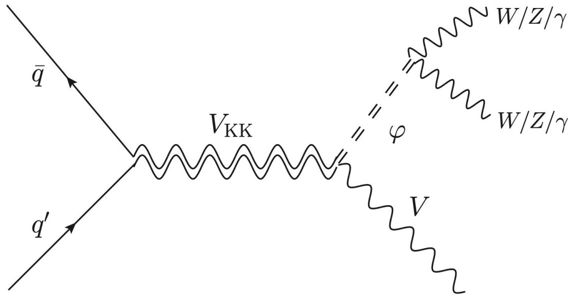
atownse2@nd.edu

Outline

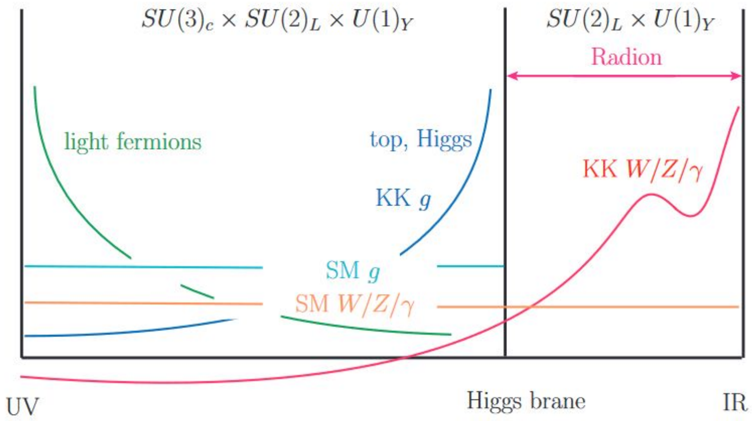
- Signal Model
- Experimental Motivation
- Diphoton Tagger
- Event Selection
- Signal/Background Modeling
    - Parametric
    - Semi-parametric GP
    - Open questions
- Next steps

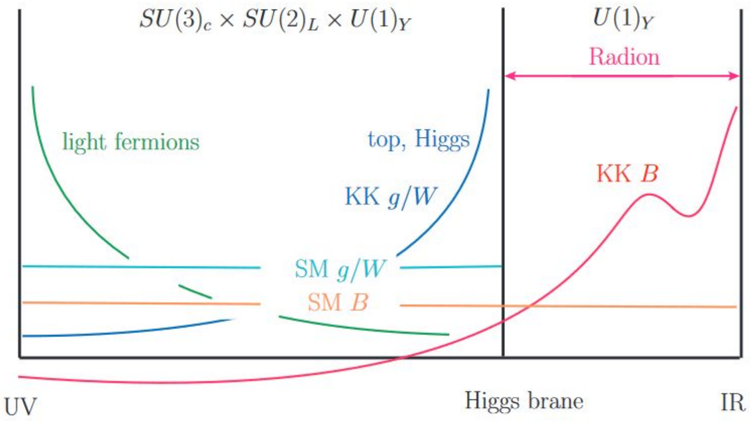
Signal Model

- Warped Extra Dimensions
  - Hierarchy problem
  - Dark matter
- Cascade decay is dominant



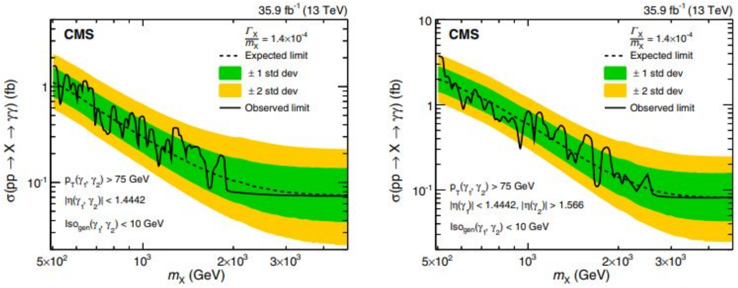
[Agashe et. al. 2017 Dedicated Strategies for Triboson Signals from Cascade Decays of Vector Resonances]



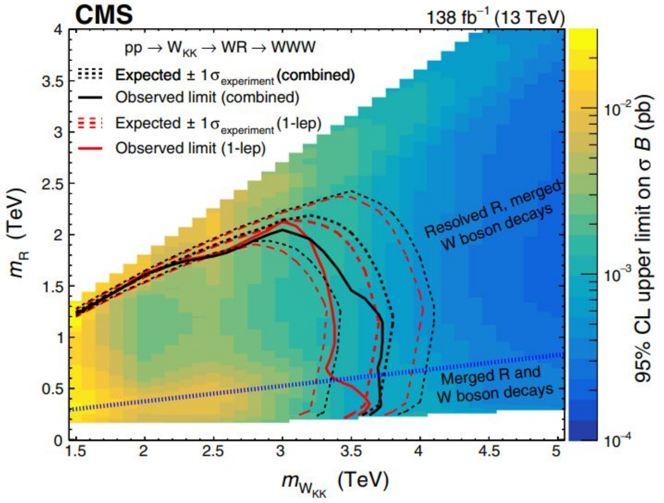


Experimental Motivation

- Non-standard final state
- Two resonances
- Limits: Diphoton searches
  - Photon shower shape selections reject partially merged photons
  - No “resonance” for resolved scenarios
- Limits: Three W Searches
  - Triphoton final state suppressed in models with electroweak gauge bosons in the extended bulk



[Phys. Rev. D 98, 092001 (2018) - Search for physics beyond the standard model in high-mass diphoton events from proton-proton collisions at $\sqrt{s}=13\text{ }\text{ }\mathrm{TeV}$ (aps.org)]

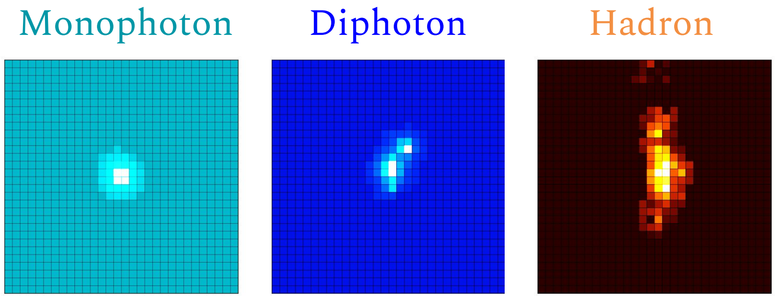


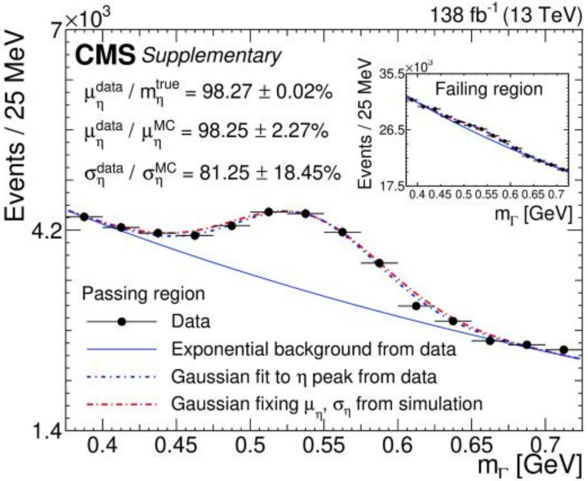
[Phys. Rev. Lett. 129, 021802 (2022) - Search for Resonances Decaying to Three W Bosons in Proton-Proton Collisions at $\sqrt{s}=13\text{ }\text{ }\mathrm{TeV}$ (aps.org)]

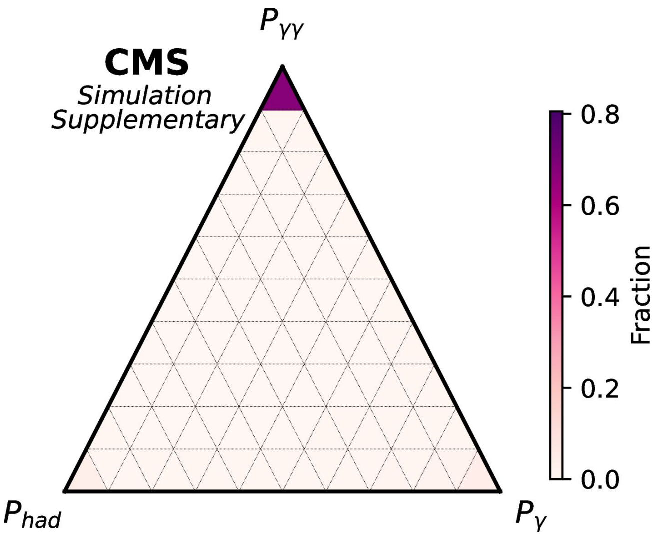
Tagger

[Steven's talk]

- Convolutional Neural Network
  - Classification (Photon, Diphoton, Hadron)
  - Regression (mass/energy)



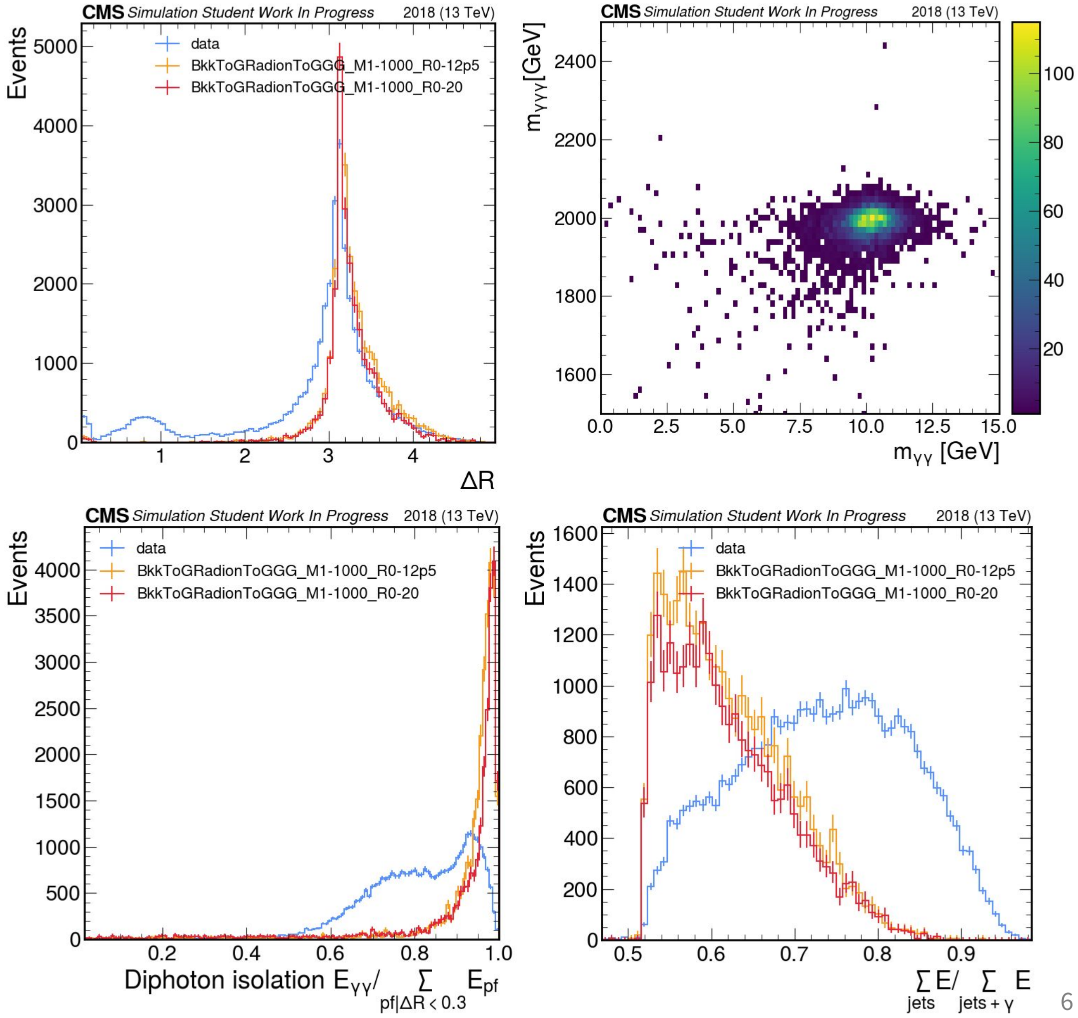




[[2405.00834] Search for new resonances decaying to pairs of merged diphotons in proton-proton collisions at $\sqrt{s}$ = 13 TeV (arxiv.org)]

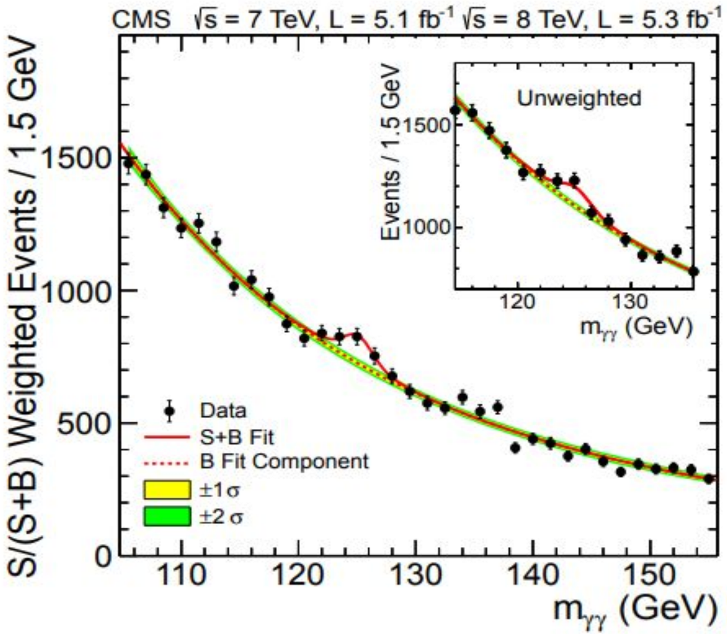
Event Selection

- Double photon trigger
  - $E_T > 60, 70\ \mathbf{GeV}$
- Discriminating variables
  - Diphoton classifier score
  - Diphoton isolation
  - Hadronic energy fraction
  - Delta R, Delta Eta
- Candidate selection
  - Combination with highest diphoton score



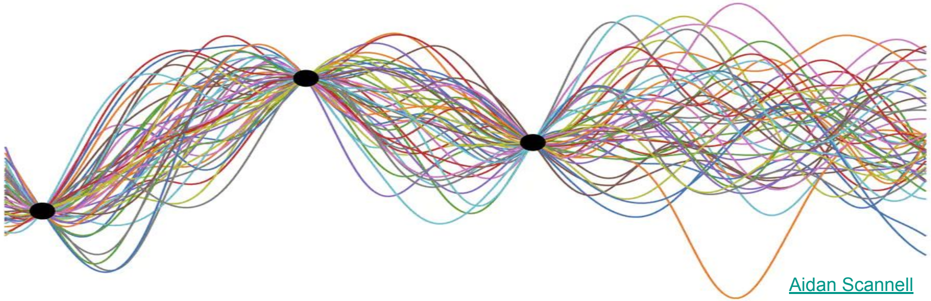
Parametric Models

- Characteristic quantities
  - Invariant Mass
- Poisson Process
  - Intensity function $\lambda(x) = f(x) + \mu s(x)$
- Model
  - Falling Background
  - Binned or unbinned
  - Parametric Functions
    - Dijet $f(x) = p_0 \frac{(1-x)^{p_1}}{x^{p_2}}$
    - Diphoton $f(x) = p_0 x^{p_1 + p_2 \log x}$
    - Power $f(x) = p_0 p_1^{p_2 x + p_3 / x}$
- Parametric form depends on background composition and data size



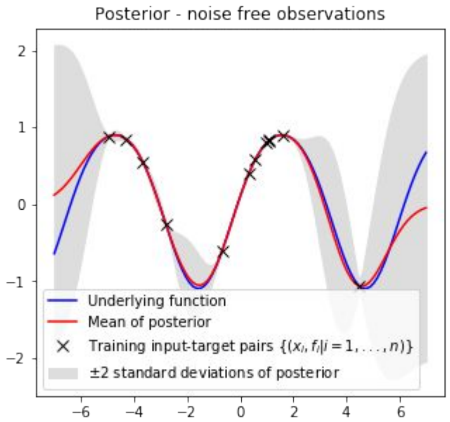
[1207.7235] Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC (arxiv.org)

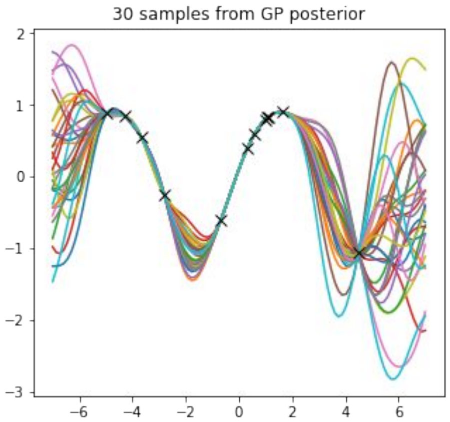
Semi-parametric Models



Aidan Scannell

- Semi-parametric alternative: Gaussian process (GP)
  - Specified by a mean and covariance function
  - A GP is said to specify a distribution over functions
  - Parameters can be interpretable (e.g. smoothness)
  - Natural uncertainty estimation
  - **Non-trivial choice of mean and covariance functions**





$$\mathrm{m} = 0$$

$$K_{i,j} = \sigma \exp \frac{|x_i - x_j|^2}{2l^2}$$

Building on previous work

Model selection and signal extraction using Gaussian Process regression

Abhijith Gandrakota,[a] Amit Lath,[b] Alexandre V. Morozov[b] and Sindhu Murthy[c]

[a] *Fermi National Accelerator Laboratory, Batavia, IL, U.S.A.*

[b] *Department of Physics & Astronomy, Rutgers, The State University of New Jersey, 136 Frelinghuysen Rd., Piscataway, NJ, U.S.A.*

[c] *Department of Physics, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA, U.S.A.*

E-mail: abhijith@fnal.gov, lath@physics.rutgers.edu, morozov@physics.rutgers.edu, sindhum@andrew.cmu.edu

ABSTRACT: We present a novel computational approach for extracting localized signals from smooth background distributions. We focus on datasets that can be naturally presented as binned integer counts, demonstrating our procedure on the CERN open dataset with the Higgs boson signature, from the ATLAS collaboration at the Large Hadron Collider. Our approach is based on Gaussian Process (GP) regression — a powerful and flexible machine learning technique which has allowed us to model the background without specifying its functional form explicitly and separately measure the background and signal contributions in a robust and reproducible manner. Unlike functional fits, our GP-regression-based approach does not need to be constantly updated as more data becomes available. We discuss how to select the GP kernel type, considering trade-offs between kernel complexity and its ability to capture the features of the background distribution. We show that our GP framework can be used to detect the Higgs boson resonance in the data with more statistical significance than a polynomial fit specifically tailored to the dataset. Finally, we use Markov Chain Monte Carlo (MCMC) sampling to confirm the statistical significance of the extracted Higgs signature.

[Model selection and signal extraction using Gaussian Process regression | Journal of High Energy Physics (springer.com)]

Modeling Smooth Backgrounds & Generic Localized Signals with Gaussian Processes

Meghan Frate,[1] Kyle Cranmer,[2] Saarik Kalia,[3] Alexander Vandenberg-Rodes,[4] and Daniel Whiteson[1]

[1] *Department of Physics and Astronomy, University of California, Irvine, CA 92697*
[2] *Department of Physics and Astronomy, New York University, New York, NY 10003*
[3] *Department of Physics, MIT, Boston, MA*
[4] *Obsidian Security Inc., Newport Beach, CA 92660*

We describe a procedure for constructing a model of a smooth data spectrum using Gaussian processes rather than the historical parametric description. This approach considers a fuller space of possible functions, is robust at increasing luminosity, and allows us to incorporate our understanding of the underlying physics. We demonstrate the application of this approach to modeling the background to searches for dijet resonances at the Large Hadron Collider and describe how the approach can be used in the search for generic localized signals.

[1709.05681] Modeling Smooth Backgrounds and Generic Localized Signals with Gaussian Processes (arxiv.org)

A Bayesian Search for the Higgs Particle

Shirin Golchi*[1] and Richard Lockhart[2]

[1] University of British Columbia - Okanagan, Statistics
[2] Simon Fraser university, Department of Statistics and Actuarial Science

Abstract

The statistical procedure used in the search for the Higgs boson is investigated in this paper. A Bayesian hierarchical model is proposed that uses the information provided by the theory in the analysis of the data generated by the particle detectors. In addition, we develop a Bayesian decision making procedure that combines the two steps of the current method (discovery and exclusion) into one and can be calibrated to satisfy frequency theory error rate requirements.

Keywords: Bayes rule, decision set, Higgs boson, linear loss function

[1501.02226] A Bayesian Search for the Higgs Particle (arxiv.org)

Other works using GPs

PHYSICAL REVIEW D **96**, 032002 (2017)

Measurement of the top quark mass in the dileptonic $t\bar{t}$ decay channel using the mass observables $M_{b\ell}$, M_{T2}, and $M_{b\ell\nu}$ in pp collisions at $\sqrt{s} = 8$ TeV

A. M. Sirunyan *et al.**
(CMS Collaboration)
(Received 20 April 2017; published 22 August 2017)

A measurement of the top quark mass (M_t) in the dileptonic $t\bar{t}$ decay channel is performed using data from proton-proton collisions at a center-of-mass energy of 8 TeV. The data was recorded by the CMS experiment at the LHC and corresponds to an integrated luminosity of $19.7 \pm 0.5\ \mathrm{fb}^{-1}$. Events are selected with two oppositely charged leptons ($\ell = e, \mu$) and two jets identified as originating from b quarks. The analysis is based on three kinematic observables whose distributions are sensitive to the value of M_t. An invariant mass observable, $M_{b\ell}$, and a "stransverse mass" observable, M_{T2}, are employed in a simultaneous fit to determine the value of M_t and an overall jet energy scale factor (JSF). A complementary approach is used to construct an invariant mass observable, $M_{b\ell\nu}$, that is combined with M_{T2} to measure M_t. The shapes of the observables, along with their evolutions in M_t and JSF, are modeled by a nonparametric Gaussian process regression technique. The sensitivity of the observables to the value of M_t is investigated using a Fisher information density method. The top quark mass is measured to be $172.22 \pm 0.18(\mathrm{stat})\,^{+0.89}_{-0.93}(\mathrm{syst})$ GeV.

DOI: 10.1103/PhysRevD.96.032002

[Phys. Rev. D 96, 032002 (2017) - Measurement of the top quark mass in the dileptonic $t\overline{t}$ decay channel using the mass observables ${M}_{b\ensuremath{\ell}}$, ${M}_{\mathrm{T}2}$, and ${M}_{b\ensuremath{\ell}\ensuremath{\nu}}$ in pp collisions at $\sqrt{s}=8\text{ }\text{ }\mathrm{TeV}$ (aps.org)]

Searches for pair-produced multijet resonances using data scouting in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration*

Abstract

Searches for pair-produced multijet signatures using data corresponding to an integrated luminosity of $128\,\mathrm{fb}^{-1}$ of proton-proton collisions at $\sqrt{s} = 13$ TeV are presented. A data scouting technique is employed to record events with low jet scalar transverse momentum sum values. The electroweak production of particles predicted in R-parity violating supersymmetric models is probed for the first time with fully hadronic final states. This is the first search for prompt hadronically decaying mass-degenerate higgsinos, and extends current exclusions on R-parity violating top squarks and gluinos.

Submitted to Physical Review Letters

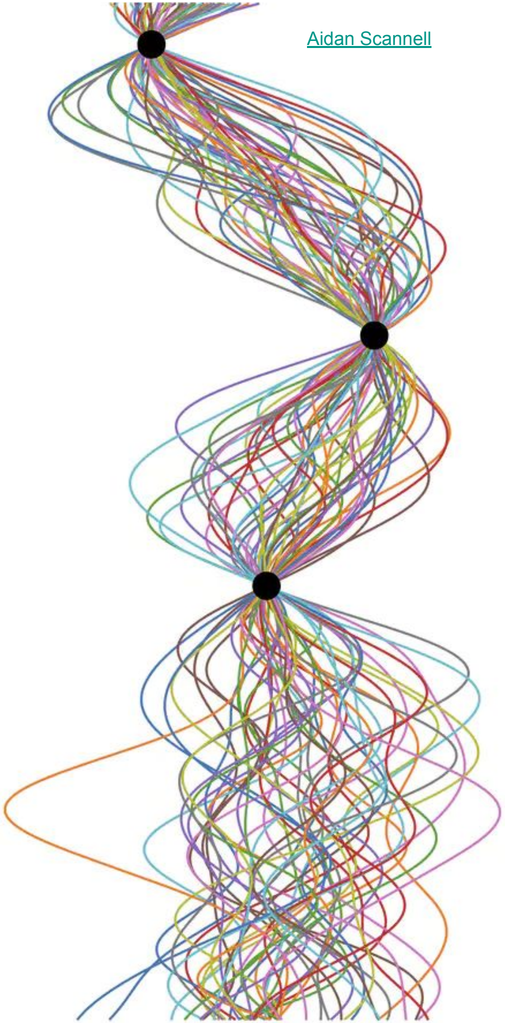
[2404.02992] Searches for pair-produced multijet resonances using data scouting in proton-proton collisions at $\sqrt{s}$ = 13 TeV (arxiv.org)

Aidan Scannell

Semi-parametric Models

Remark: Count data is not Gaussian

- Gaussian likelihood with GP prior -> Gaussian predictive distribution
- Poisson likelihood with GP prior -> **Intractable integral**
- Estimate integral with Markov Chain Monte Carlo (MCMC)
- GP Approaches with count data
    - Unbinned: Doubly stochastic Poisson process (Cox process)
    - Binned: Latent variable GP with Poisson likelihood (this work)



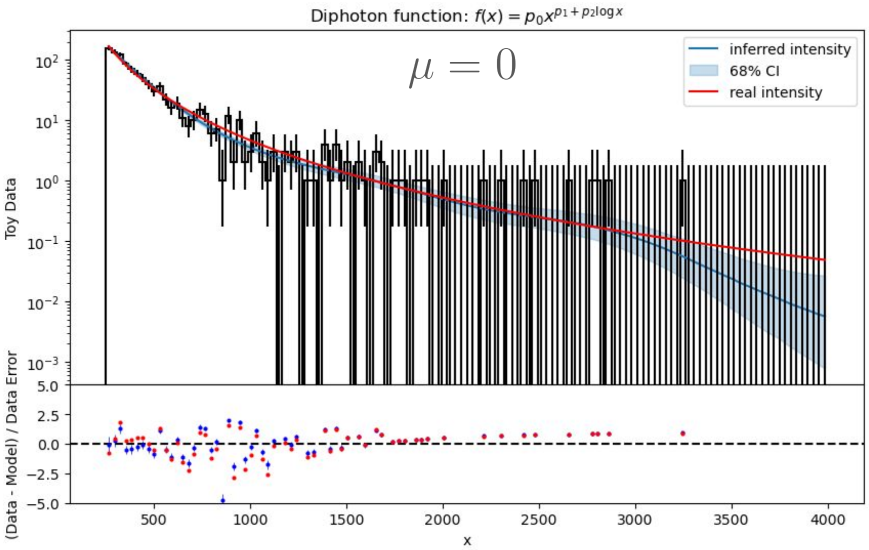
Semi-parametric Models

Inspired by: [1501.02226] A Bayesian Search for the Higgs Particle (arxiv.org)

- Latent Gaussian process with Poisson likelihood

$$y_i \sim Poisson(\lambda_i)$$
$$log(\lambda_i) = m(x_i) + \epsilon_i + \mu s(x_i)$$
$$\epsilon \sim GP(\mathbf{0}, K)$$

- Log transform to enforce positivity
- Mean function for stationarity



$$m(x) = \beta_0 + \beta_1 x$$

$$K_{i,j} = \sigma \exp \frac{|x_i - x_j|^2}{2l^2}$$

GP: Open Questions

- Bias/Coverage
- Range of fit
- Mean vs. Covariance functions
  - How to model trends and smoothness
  - Priors are important
- Negative binomial likelihood?
- Monotonic GP’s?
  - SDE’s
  - Normalizing flow
  - Conditional gradient

Next Steps

- Define realistic/non-informative priors
- Extend to 2D
  - Signal kernel model
  - Extend mean/covariance functions (ARD)
- Define statistical procedure(s) for:
  - Excluding null (no bump) hypothesis
  - Exclusion limits
- Model/measure systematic effects