Eighth Machine Learning in High Energy Physics Summer School 2023

12 April 2023, Erice, Italy

MLHEP 2023

Two machine learning applications for LHC searches

Yevgeny Kats

אוניברסיטת בן-גוריון בנגב

جامعة بن غوريون في النقب

Ben-Gurion University of the Negev

Plan

Two examples of possible ML uses in searches for physics beyond the Standard Model at the LHC:

1 **Searching for periodic signals in kinematic distributions using continuous wavelet transforms**

Eur. Phys. J. C 80 (2020) 3, 192 [arXiv:1907.03676]

in collaboration with Hugues Beauchesne

+ implementation in ATLAS

ATLAS-CONF-2023-010 (presented at Moriond 2023)

ML is used for image analysis: supervised / unsupervised.

2 **Searching for dark jets with displaced vertices using weakly supervised machine learning**

M.Sc. thesis by Noam Wunch (2022) [not yet published]

with help from Debjyoti Bardhan and Hugues Beauchesne

ML is used to achieve model-independence.

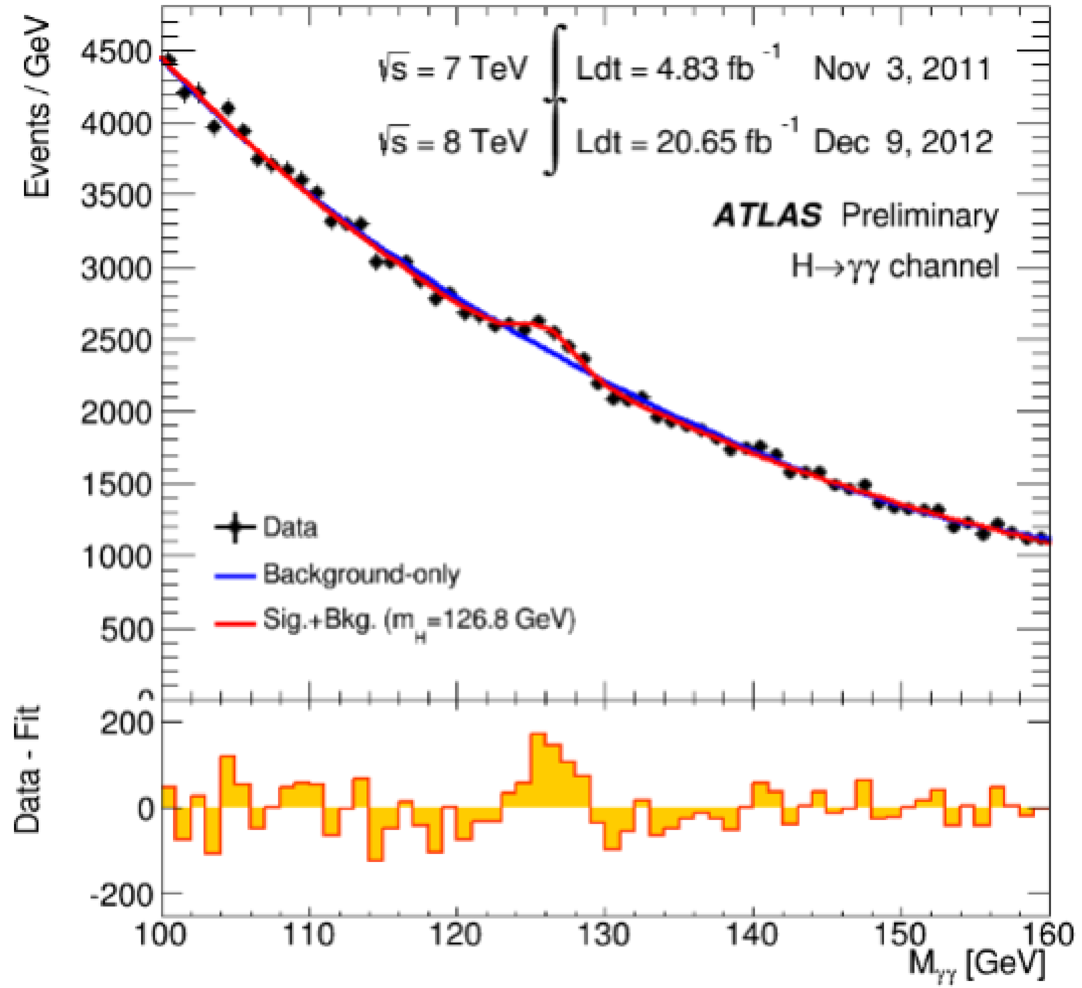

CENTRO ETTORE MAJORANA
ANTONINO

Particle discovery with invariant mass

Example

Higgs boson discovery

via $H \rightarrow \gamma\gamma$



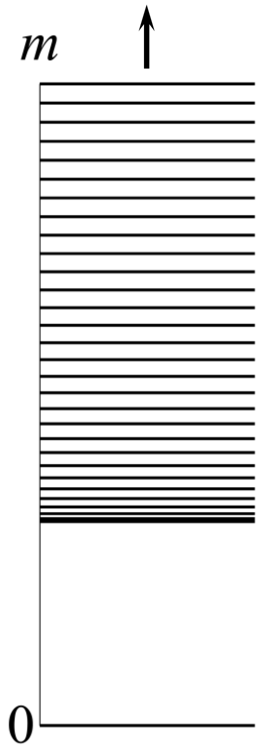
Diphoton invariant mass

$$M_{\gamma\gamma} = \sqrt{(E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2} = \sqrt{E_H^2 - \vec{p}_H^2} = M_H$$

Motivation for wavelet transform analysis

Towers of new particles with small (but resolvable) mass splittings e.g., Kaluza-Klein modes, clockwork gears.

- Clockwork / linear dilaton (CW/LD)
  Antoniadis, Arvanitaki, Dimopoulos, Giveon [1102.4043]; Baryakhtar [1202.6674]; Giudice, McCullough [1610.07962]; Giudice, $\mathcal{YK}$, McCullough, Torre, Urbano [1711.08437]
- RS models with low k / LED with small curvature
  Giudice, Plehn, Strumia [hep-ph/0408320]; Kisselev [0804.3941]; Franceschini, Giardino, Giudice, Lodone, Strumia [1101.4919]
- Discrete clockwork models
  Giudice, McCullough [1610.07962]
- Warped dark sector
  Brax, Fichet, Tanedo [1906.02199]



Signature: recurring peaks with comparable cross sections.

Motivates looking at the *Fourier transform*, or a *wavelet transform*.

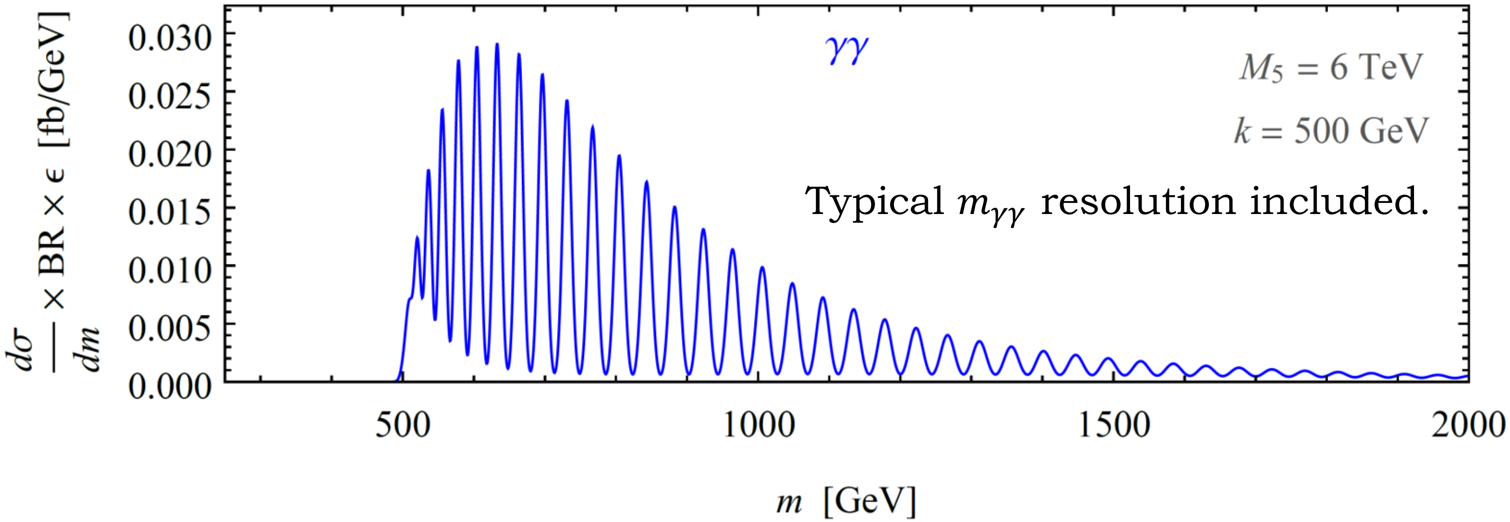
Motivation for wavelet transform analysis

Towers of new particles with small (but resolvable) mass splittings e.g., Kaluza-Klein modes, clockwork gears.

- Clockwork / linear dilaton (CW/LD)

  Antoniadis, Arvanitaki, Dimopoulos, Giveon [1102.4043]; Baryakhtar [1202.6674]; Giudice, McCullough [1610.07962]; Giudice, $\mathcal{YK}$, McCullough, Torre, Urbano [1711.08437]

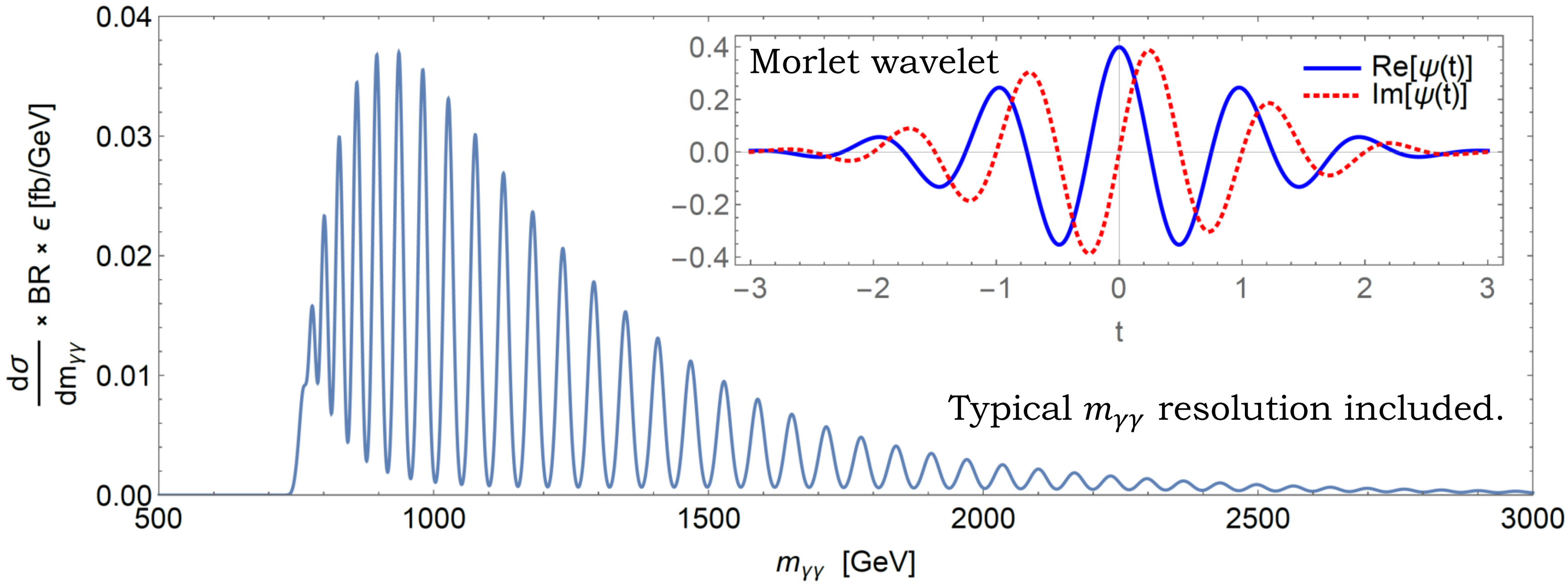
Can address the electroweak-Planck hierarchy problem.



Signature: recurring peaks with comparable cross sections.

Motivates looking at the *Fourier transform*, or a *wavelet transform.*

Wavelet transform

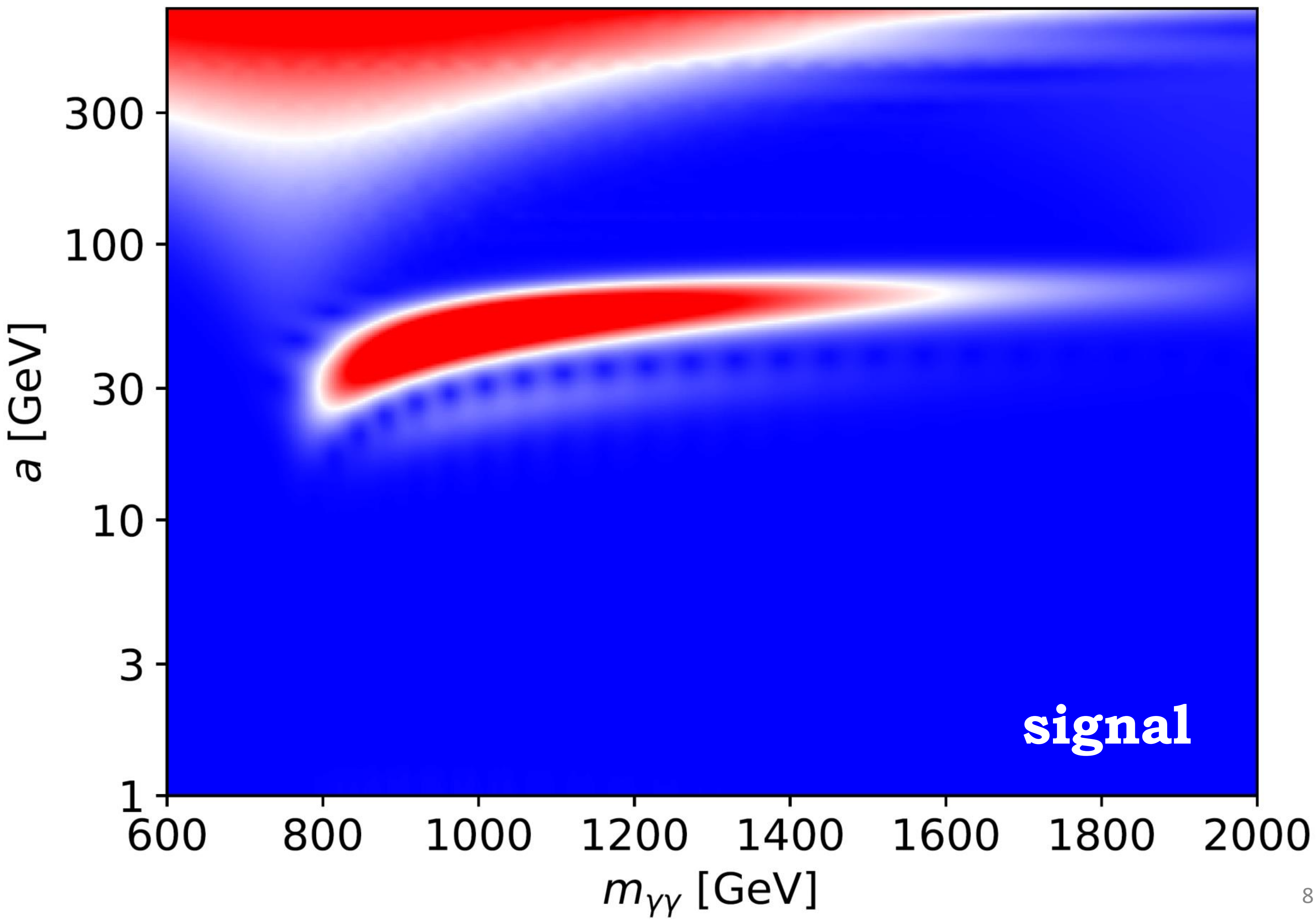


$$W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t)\, \psi^* \left(\frac{t-b}{a}\right) dt$$

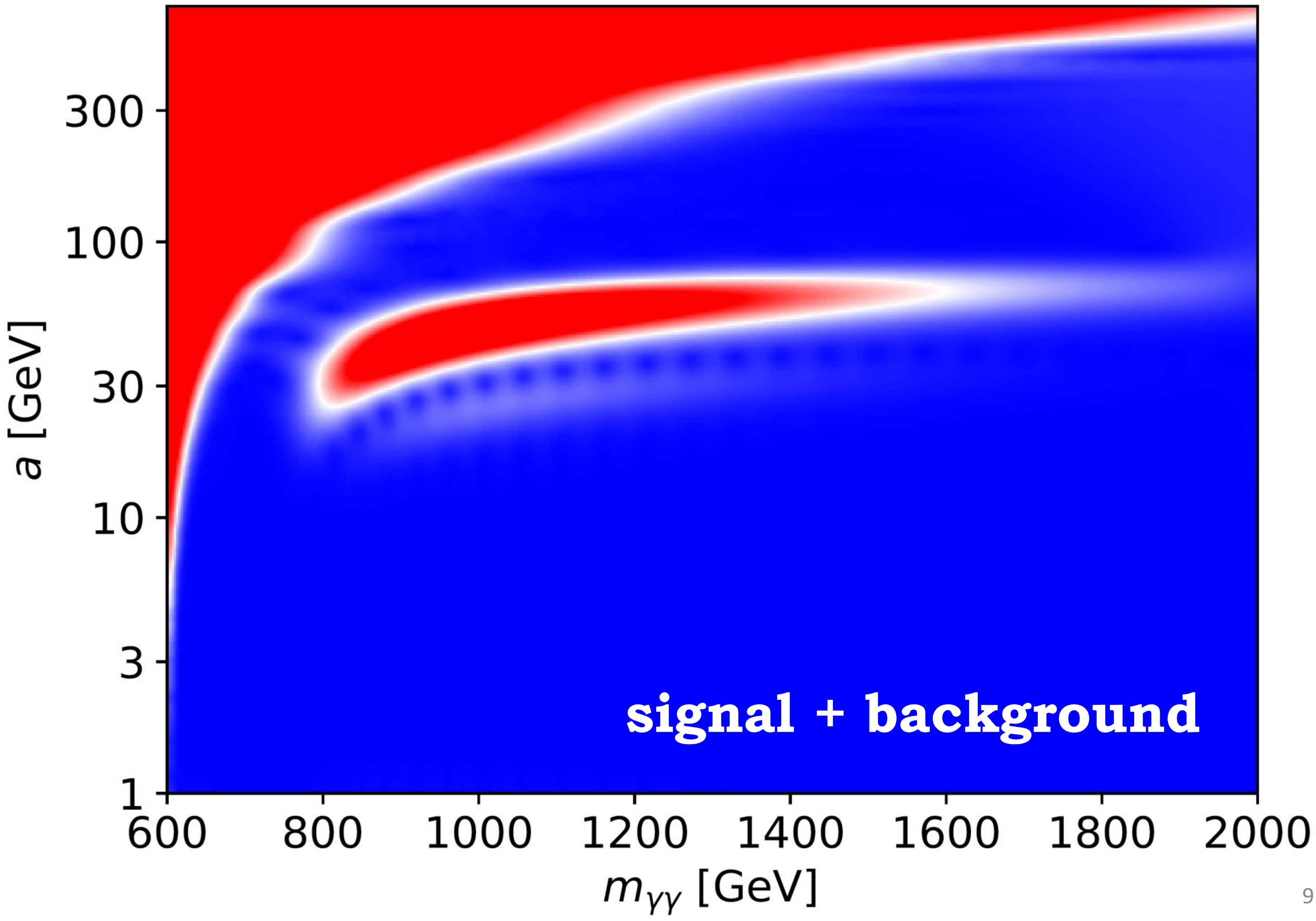
Morlet wavelet: $\psi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2} (e^{2\pi i t} - e^{-2\pi^2})$

Another work considered a discrete wavelet transform:
Lillard, Plehn, Romero, Tait [1906.10890]

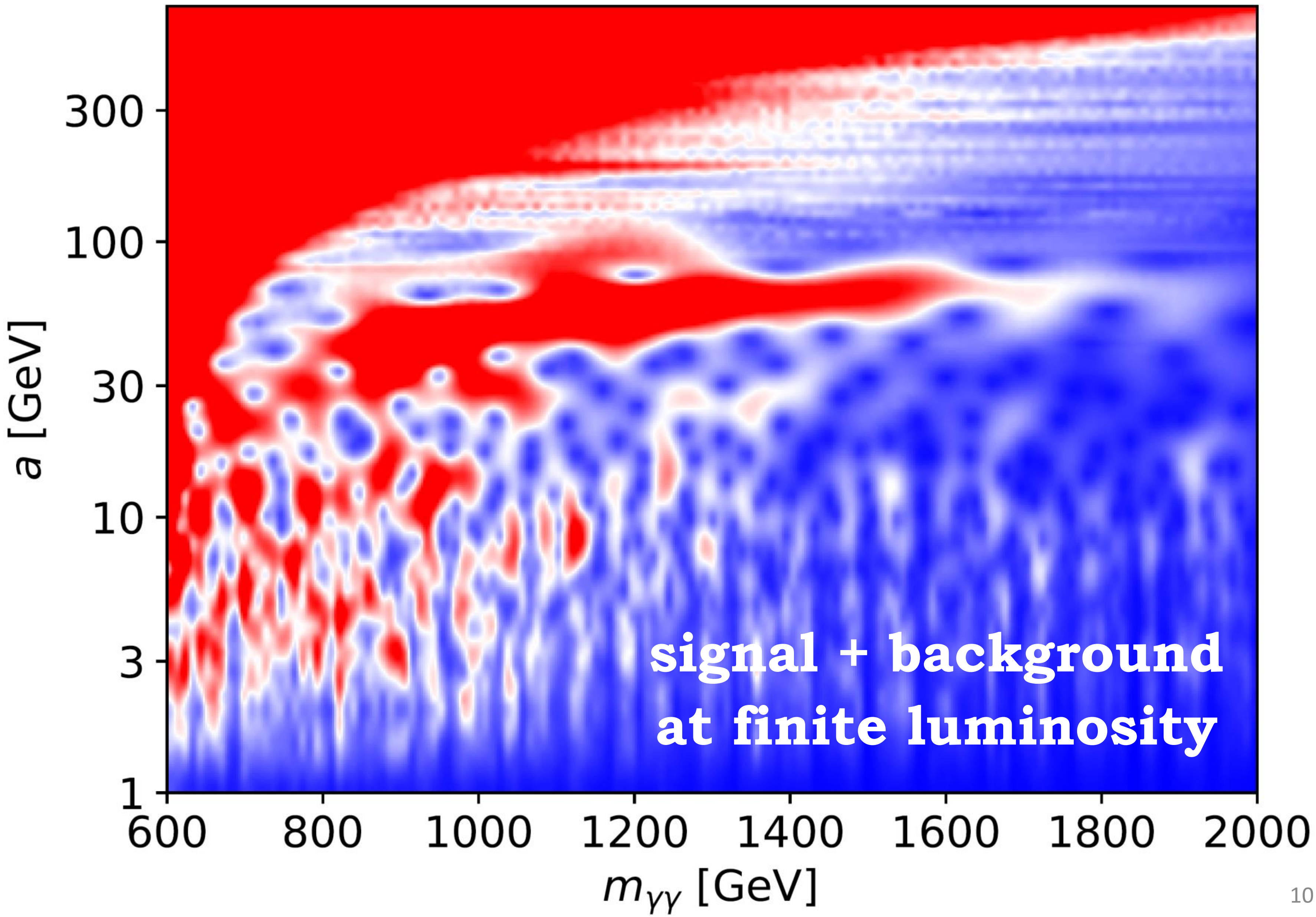
Wavelet transform



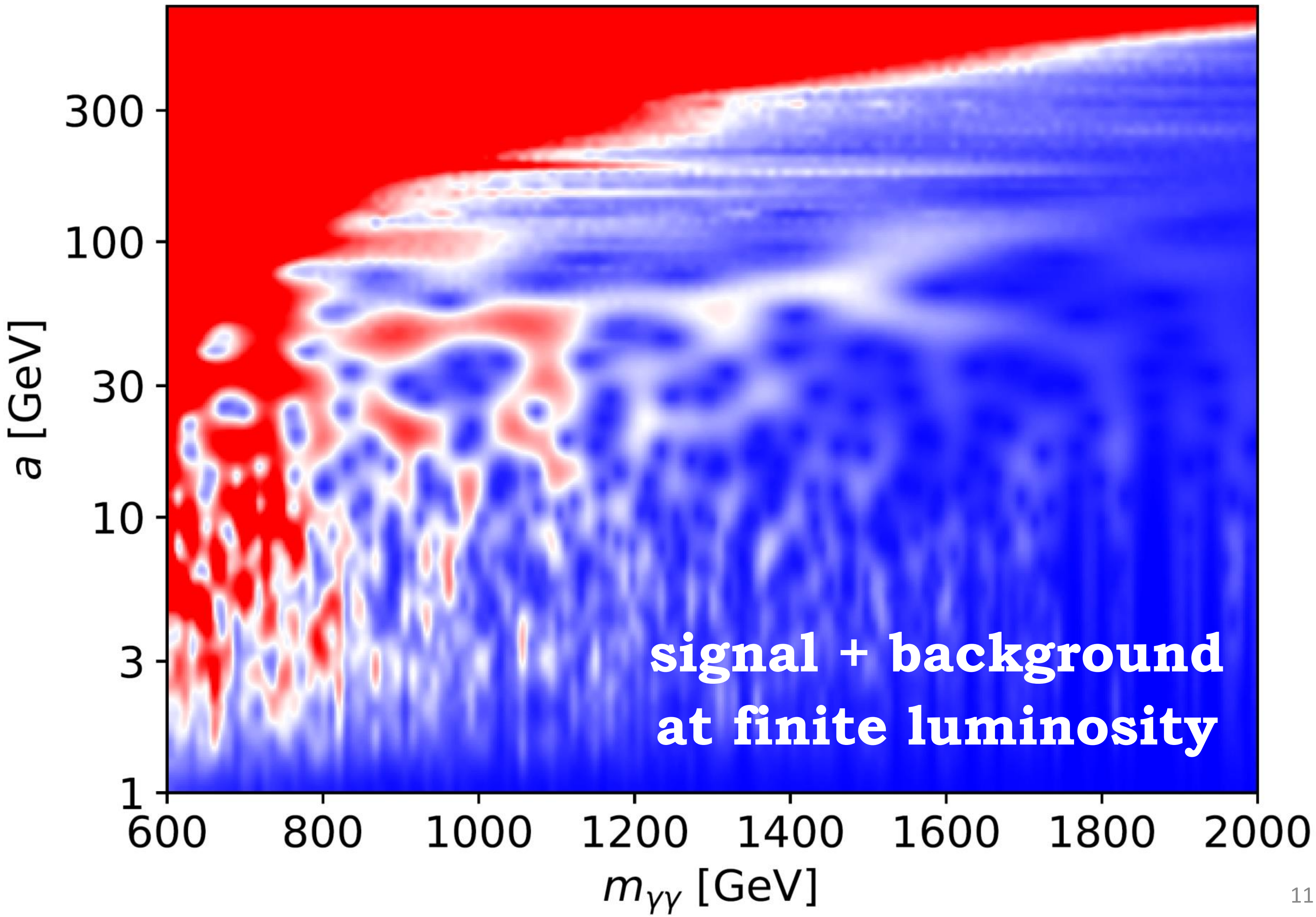
Wavelet transform



Wavelet transform



Wavelet transform



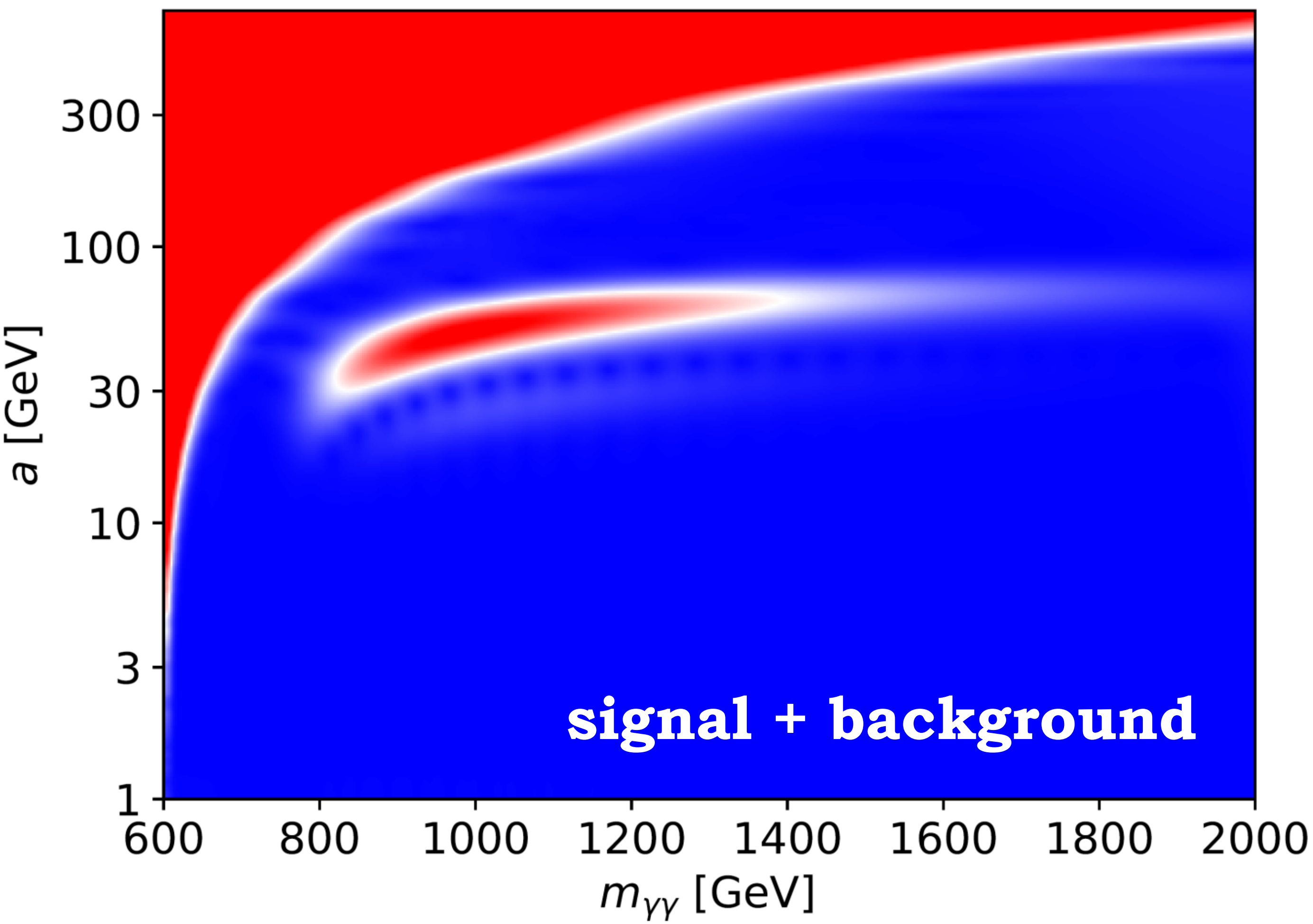
Five methods for analyzing the scalogram

1. **Model-specific search with a simple test statistic**
2. **Model-independent search with a simple test statistic**
3. **Neural network and a simple test statistic**
4. **Classifier neural network**
5. **Autoencoder neural network**

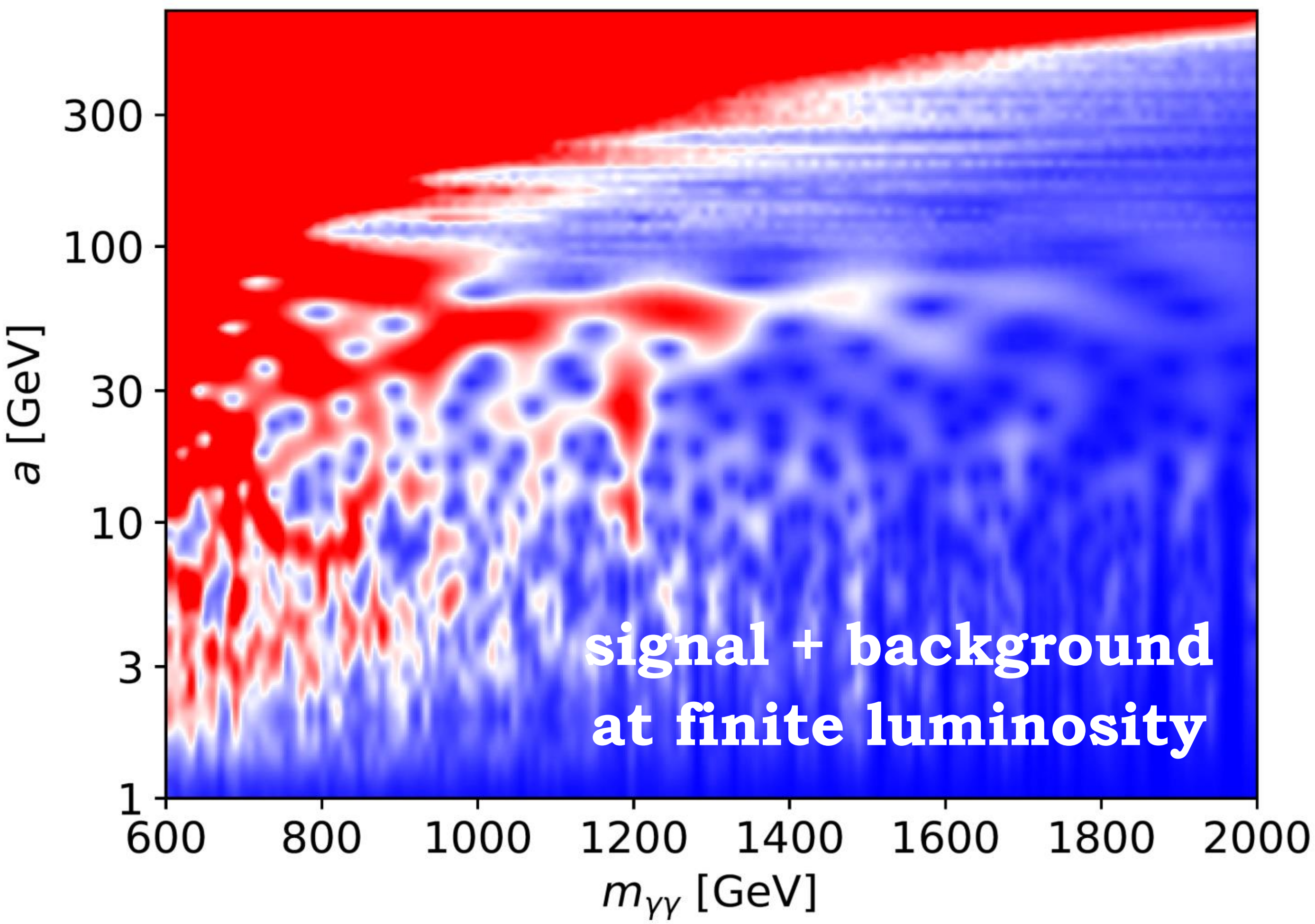
+ comparison with Fourier transform

Giudice, $\mathcal{YK}$, McCullough, Torre, Urbano [1711.08437]

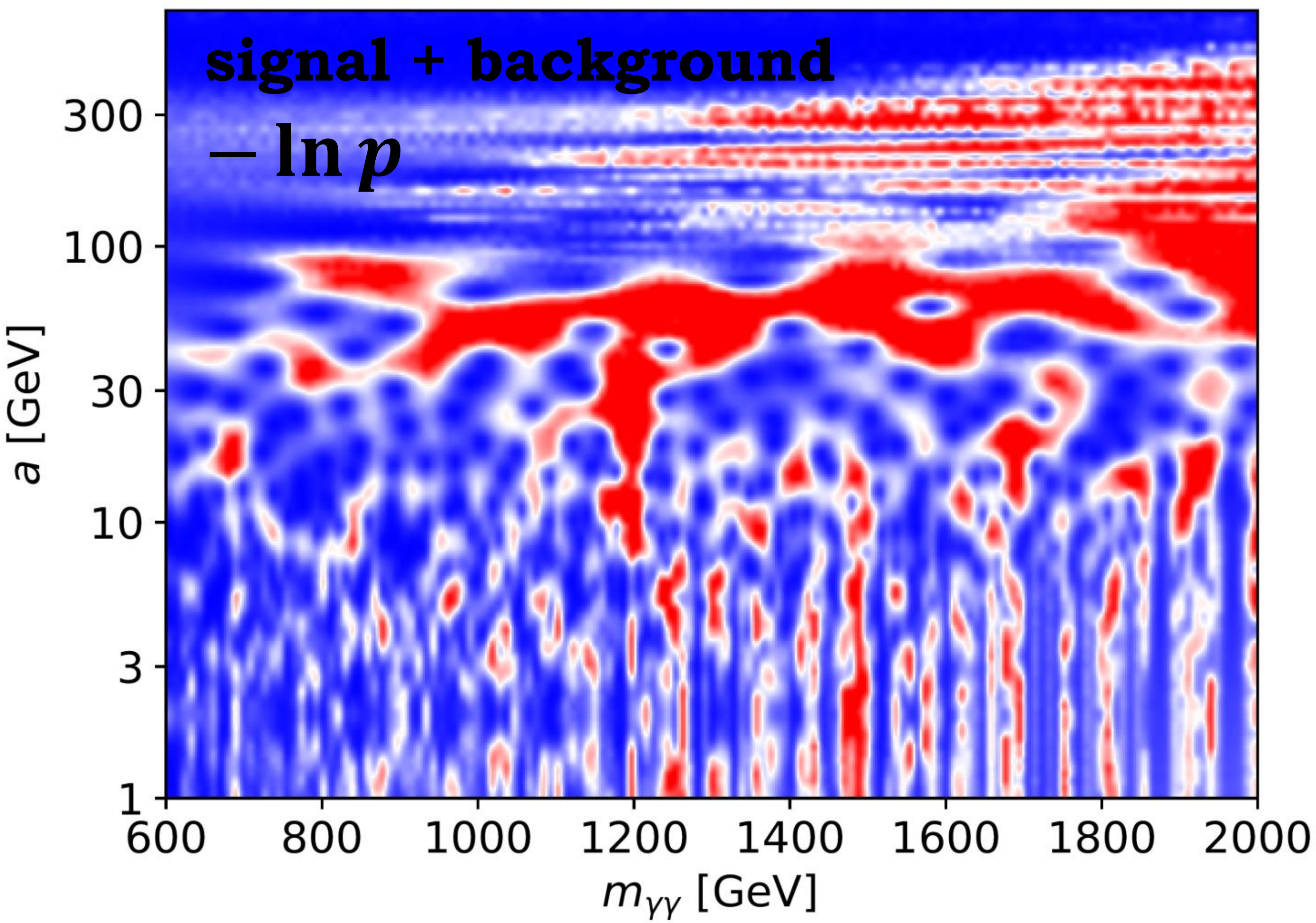
Method 1
Model-specific search with a simple test statistic



Method 1
Model-specific search with a simple test statistic

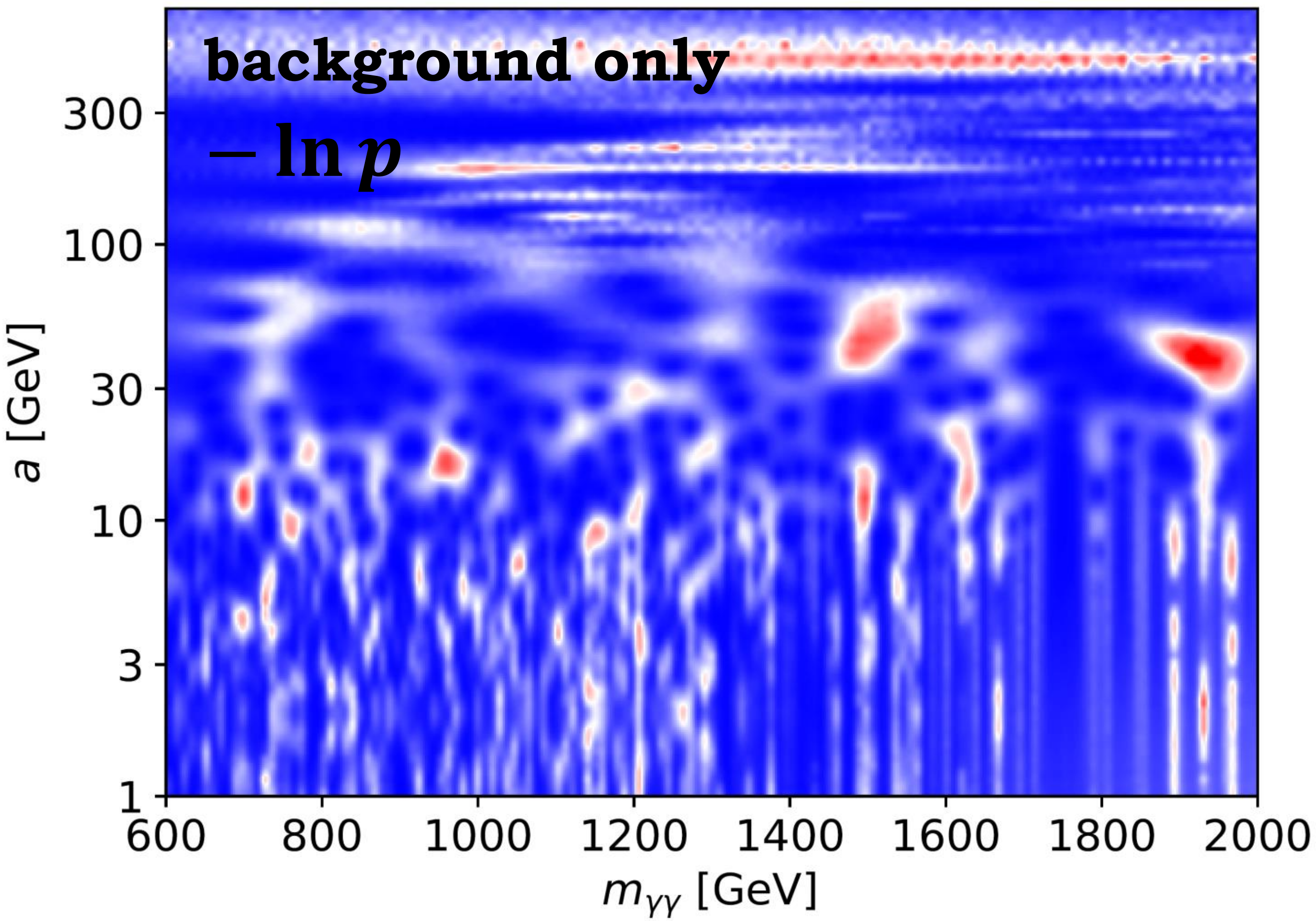


Method 1
Model-specific search with a simple test statistic



Method 1

Model-specific search with a simple test statistic



Method 1

Model-specific search with a simple test statistic

Test statistic – a function of the local p values:

$$t = -\sum_{i=i_{\min}}^{i_{\max}} \frac{1}{a_i} \ln p_i(a_i)$$

Range of lattice points to use is optimized for the particular signal.

Method 2

Model-independent search with a simple test statistic

Test statistic – a function of the local p values:

$$t = -\sum_{i=i_{\min}}^{i_{\max}} \frac{1}{a_i} \ln p_i(a_i)$$

Range of lattice points to use is determined as follows:

1. Select lattice points with p value below a threshold (we used 10%).
2. Identify continuous regions of such points.
3. Compute t for each region using the lowest p-value point for each mass.
4. Consider the region with the highest t.

Method 3
Neural network and a simple test statistic

Train a neural network (NN) to identify signal-like features in the $|W|/\langle|W_B|\rangle$ scalogram.

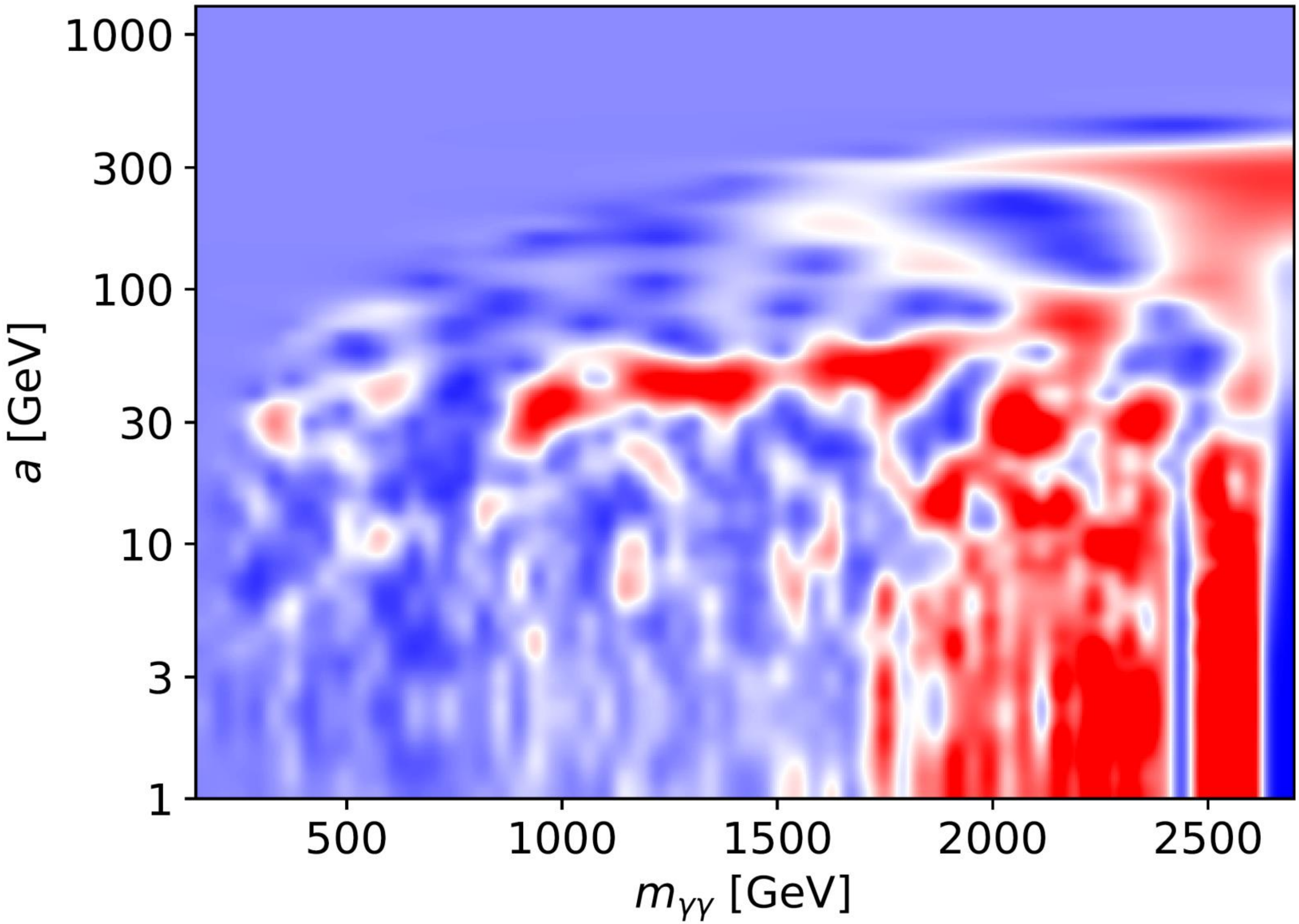
1. Train on samples with signals (from random points in the parameter space) to return $|W_S^P|/\langle|W_B|\rangle$.

   Train on background samples to return 0.

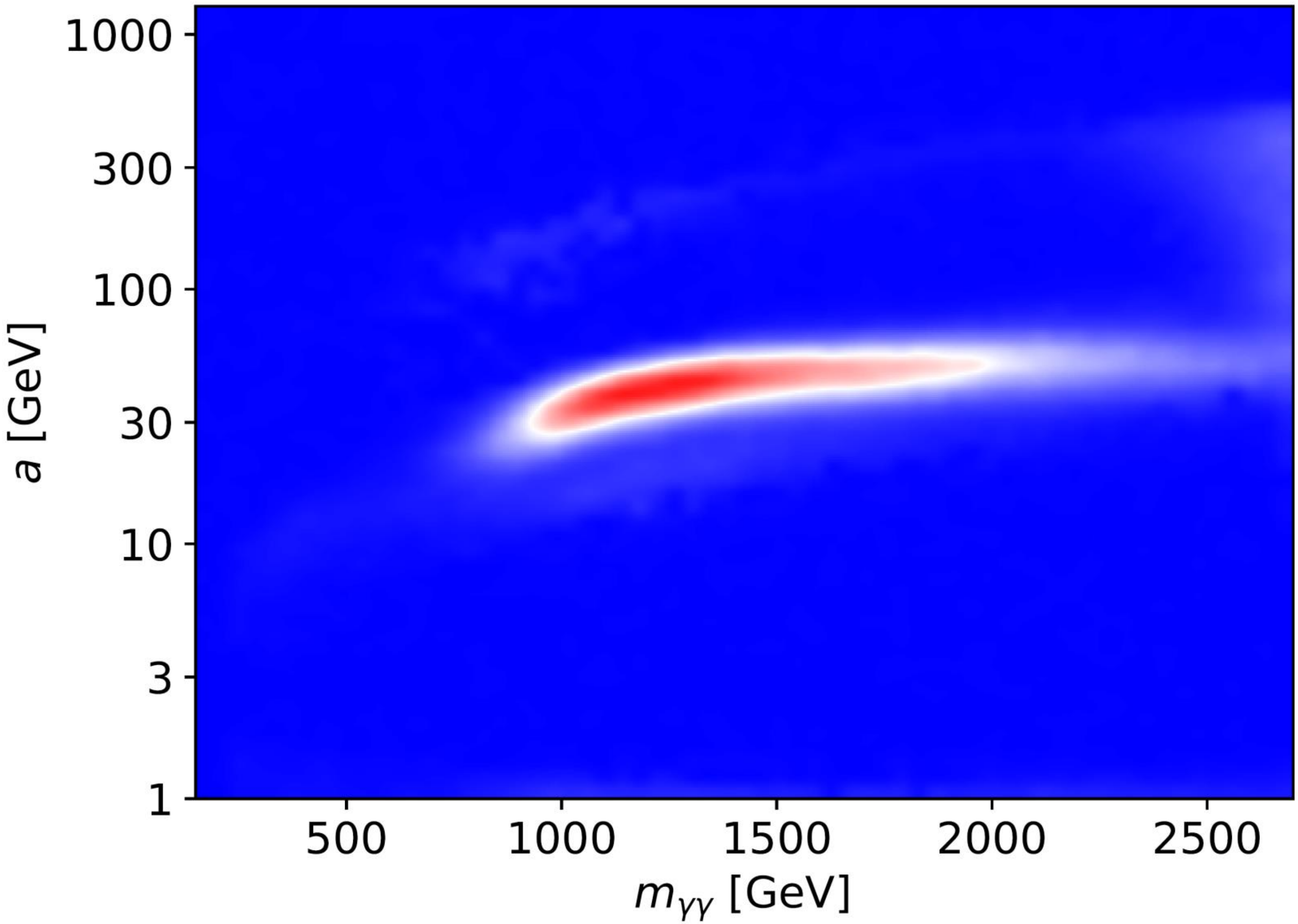
2. Select bins with NN output above a threshold and apply the test statistic

$$t = -\sum_{i=i_{\min}}^{i_{\max}} \frac{1}{a_i} \ln p_i(a_i)$$

Method 3
Neural network and a simple test statistic



Method 3
Neural network and a simple test statistic



Method 3
Neural network and a simple test statistic

Layer	Parameters
Input layer	63 mass bins $\times$ 56 scale bins
Convolutional layer 1	# filters = 32
	kernel size = (5, 5)
	Activation: Softplus
MaxPooling 1	Pooling size = (2, 2)
Convolutional layer 2	# filters = 64
	kernel size = (5, 5)
	Activation: Softplus
MaxPooling 2	Pooling size = (2, 2)
Convolutional layer 3	# filters = 128
	kernel size = (5, 5)
	Activation: Softplus
Dense 1	# of nodes = 5000
	Activation: Softplus
Output layer	# of nodes = 3528

Setting	Choice
Optimizer	Adam
Loss function	Mean squared error
# training experiments	5000
Validation split	0.2
Batch size	1000
# epochs	200
Callback	Smallest validation loss function

Method 4
Classifier neural network

1. For a particular signal, train a neural network to return:
   - **1** for scalograms containing the signal
   - **0** for scalograms with background only

2. Use the neural network output as a test statistic.

Method 4
Classifier neural network

Layer	Parameters
Input layer	63 mass bins $\times$ 56 scale bins
Convolutional layer 1	# filters = 4
	kernel size = (3, 3)
	Activation: Elu
MaxPooling 1	Pooling size = (2, 2)
Convolutional layer 2	# filters = 8
	kernel size = (3, 3)
	Activation: Sigmoid
MaxPooling 2	Pooling size = (2, 2)
Convolutional layer 3	# filters = 16
	kernel size = (3, 3)
	Activation: Sigmoid
Dense 1	# of nodes = 200
	Activation: Sigmoid
Dense 2	# of nodes = 100
	Activation: Sigmoid
Output layer	# of nodes = 1

Setting	Choice
Optimizer	Adam
Loss function	Binary cross entropy
# training experiments	4000
Validation split	0.2
Batch size	1000
# epochs	500
Callback	Smallest validation loss function

Method 5
Autoencoder neural network

Basic idea: train a NN to encode an input image in a small number of variables (a bottleneck) and then decode it.

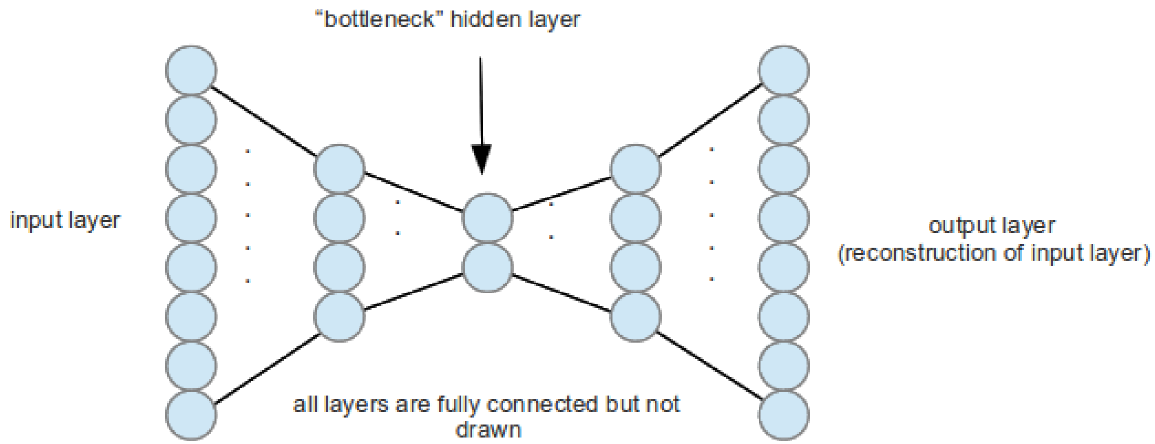


Image by Nghia Ho.

If the bottleneck is sufficiently large to capture typical backgrounds but too small to work for an arbitrary image, it may often fail to decode samples with signals.

Method 5
Autoencoder neural network

Our implementation:

1. Train an autoencoder NN on scalograms of backgrounds, with $-\ln p$ as inputs, asking it to minimize the difference between the output and input.
2. Use the NN loss function (the difference between the output and input) as a test statistic.

Method 5
Autoencoder NN

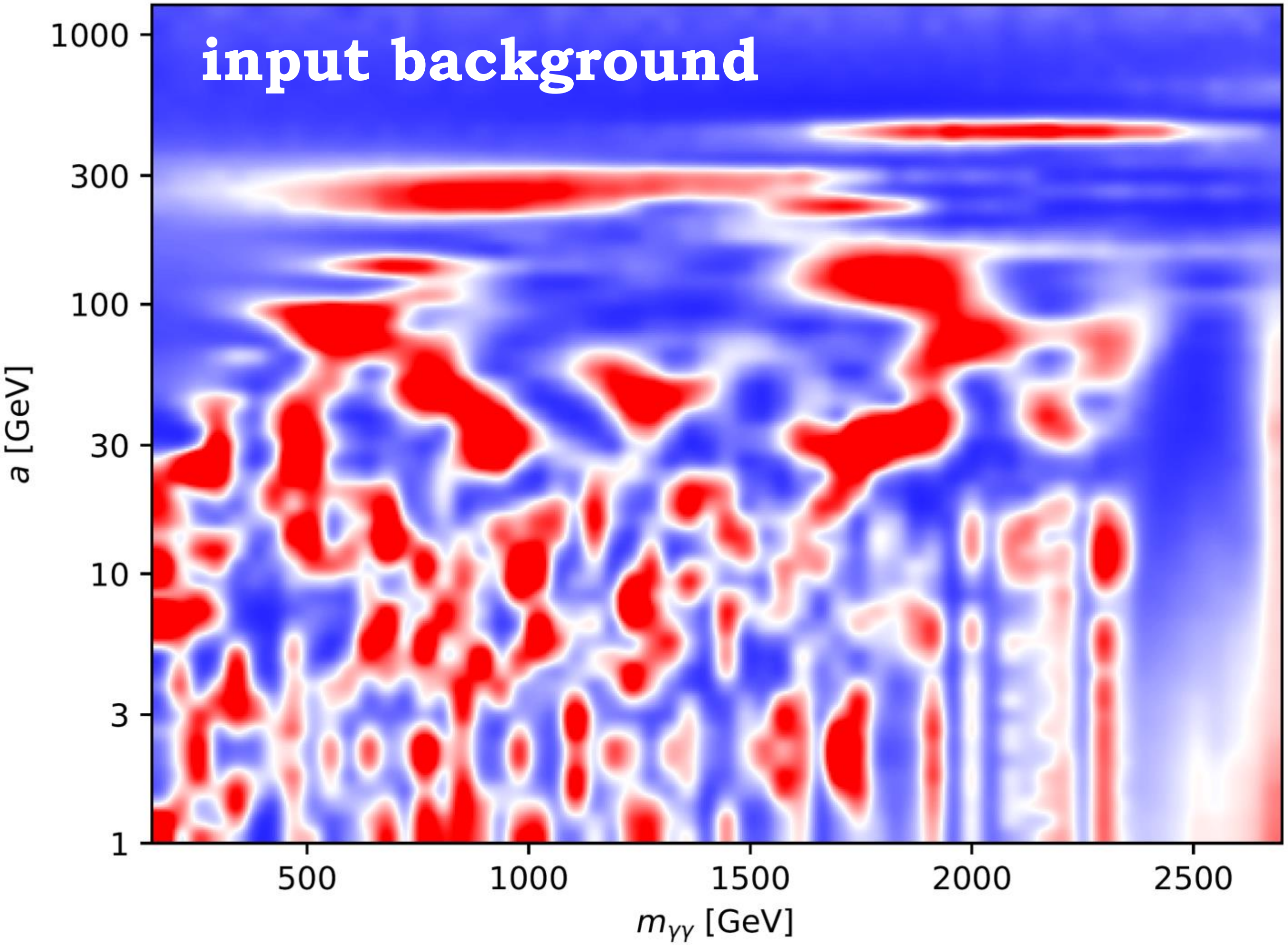
Setting	Choice
Optimizer	ADAM
Loss function	Mean squared error
# training experiments	5000
Validation split	0.2
Batch size	1000
Padding	Same
# epochs	100
Callback	Smallest validation loss function

Layer	Parameters
Input layer	$\boxed{\text{60 mass bins} \times \text{56 scale bins}}$
Convolutional layer 1	# filters = 128 kernel size = (3, 3) Activation: Elu
MaxPooling 1	Pooling size = (2, 2)
Convolutional layer 2	# filters = 128 kernel size = (3, 3) Activation: Elu
MaxPooling 2	Pooling size = (2, 2)
Convolutional layer 3	# filters = 128 kernel size = (3, 3) Activation: Elu
Dense 1	# of nodes = 40 Activation: Elu
Dense 2 (*Encoded*)	$\boxed{\text{\# of nodes} = 20}$ Activation: Elu
Dense 3	# of nodes = 40 Activation: Elu
Convolutional layer 4	# filters = 128 kernel size = (3, 3) Activation: Elu
UpSampling 1	Upsampling factors = (2, 2)
Convolutional layer 5	# filters = 128 kernel size = (3, 3) Activation: Elu
UpSampling 2	Upsampling factors = (2, 2)
Convolutional layer 6	# filters = 1 kernel size = (3, 3) Activation: Elu
Output layer	$\boxed{\text{60 mass bins} \times \text{56 scale bins}}$

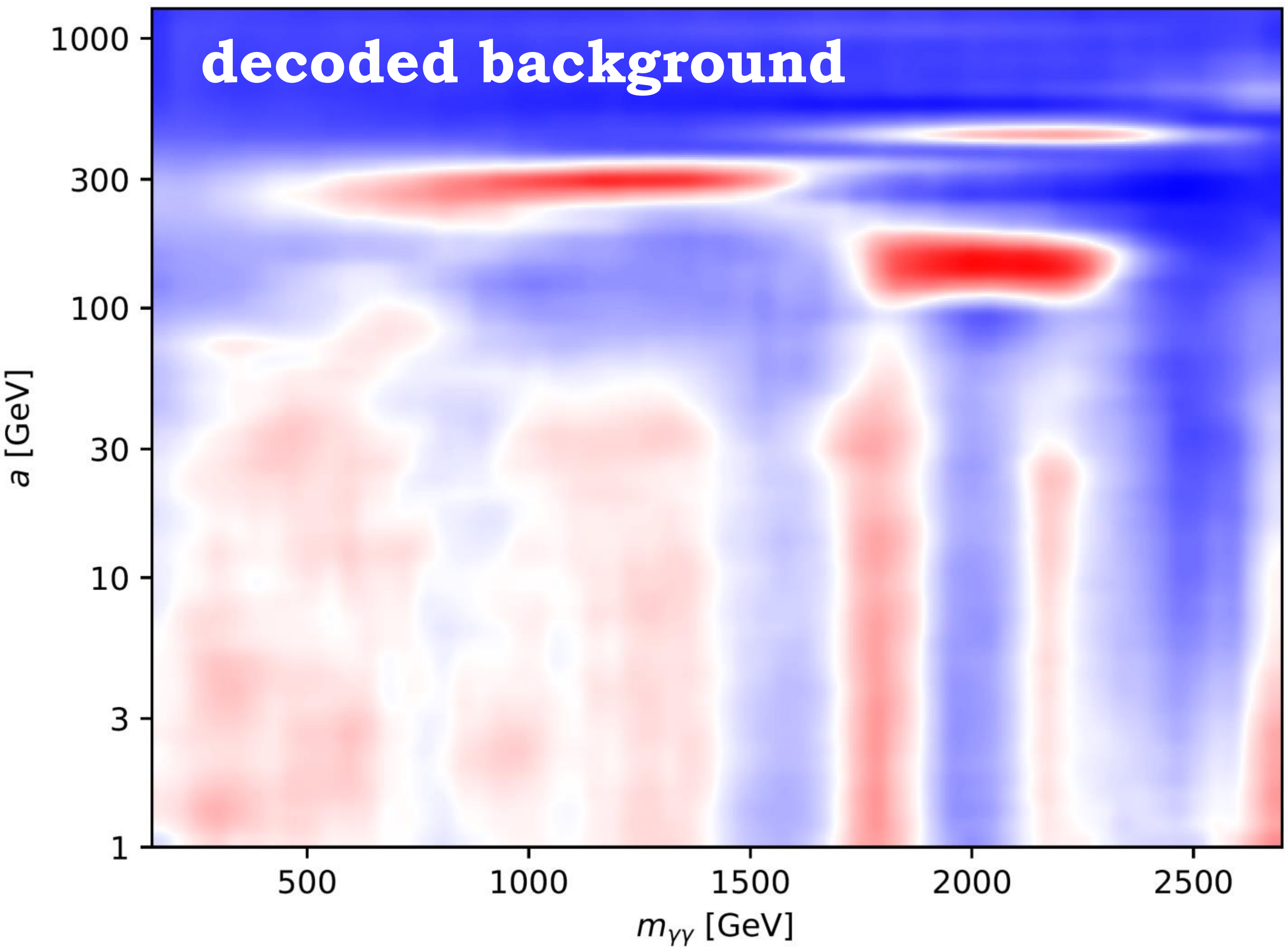
encoding

decoding

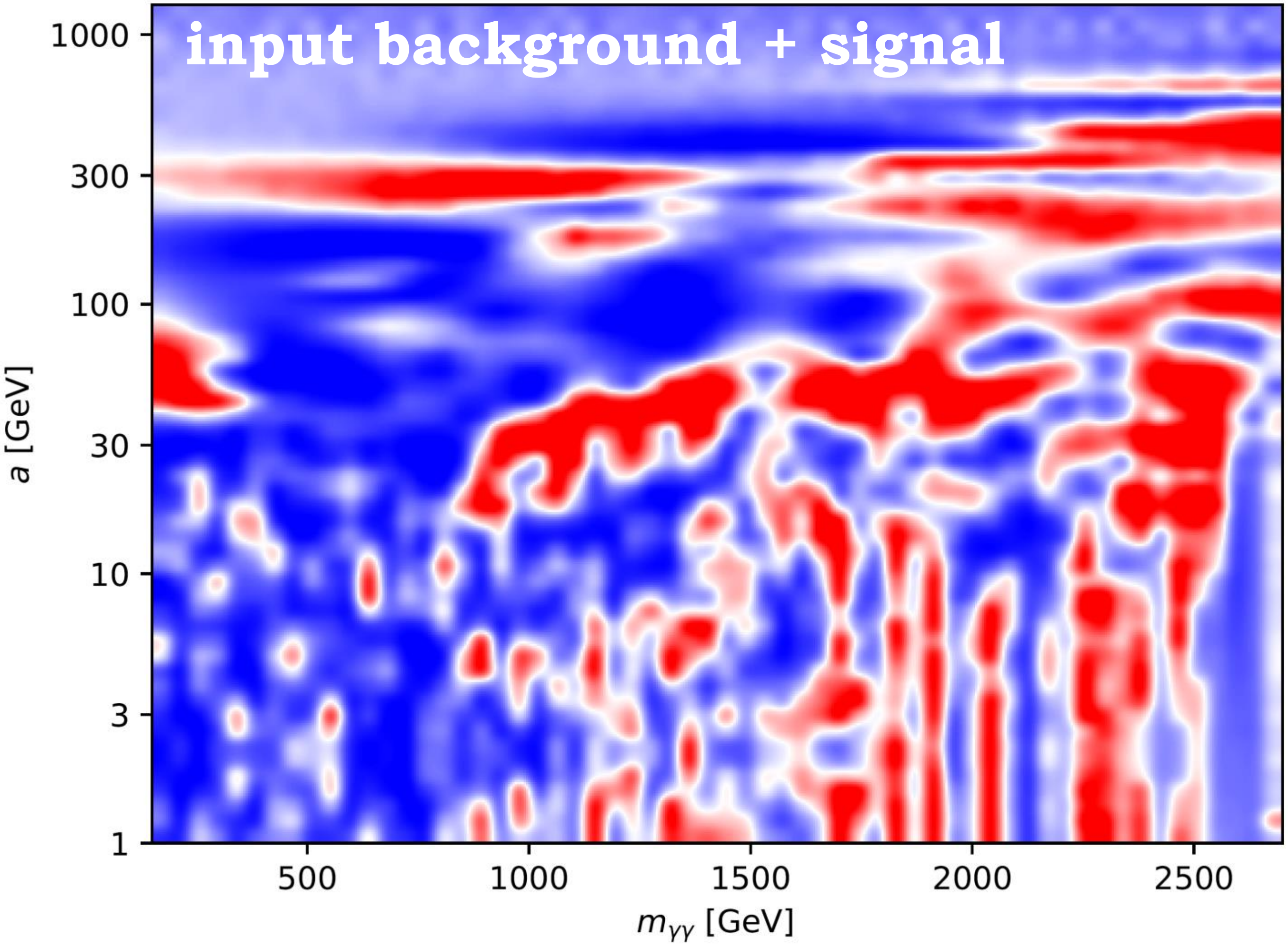
Method 5
Autoencoder neural network



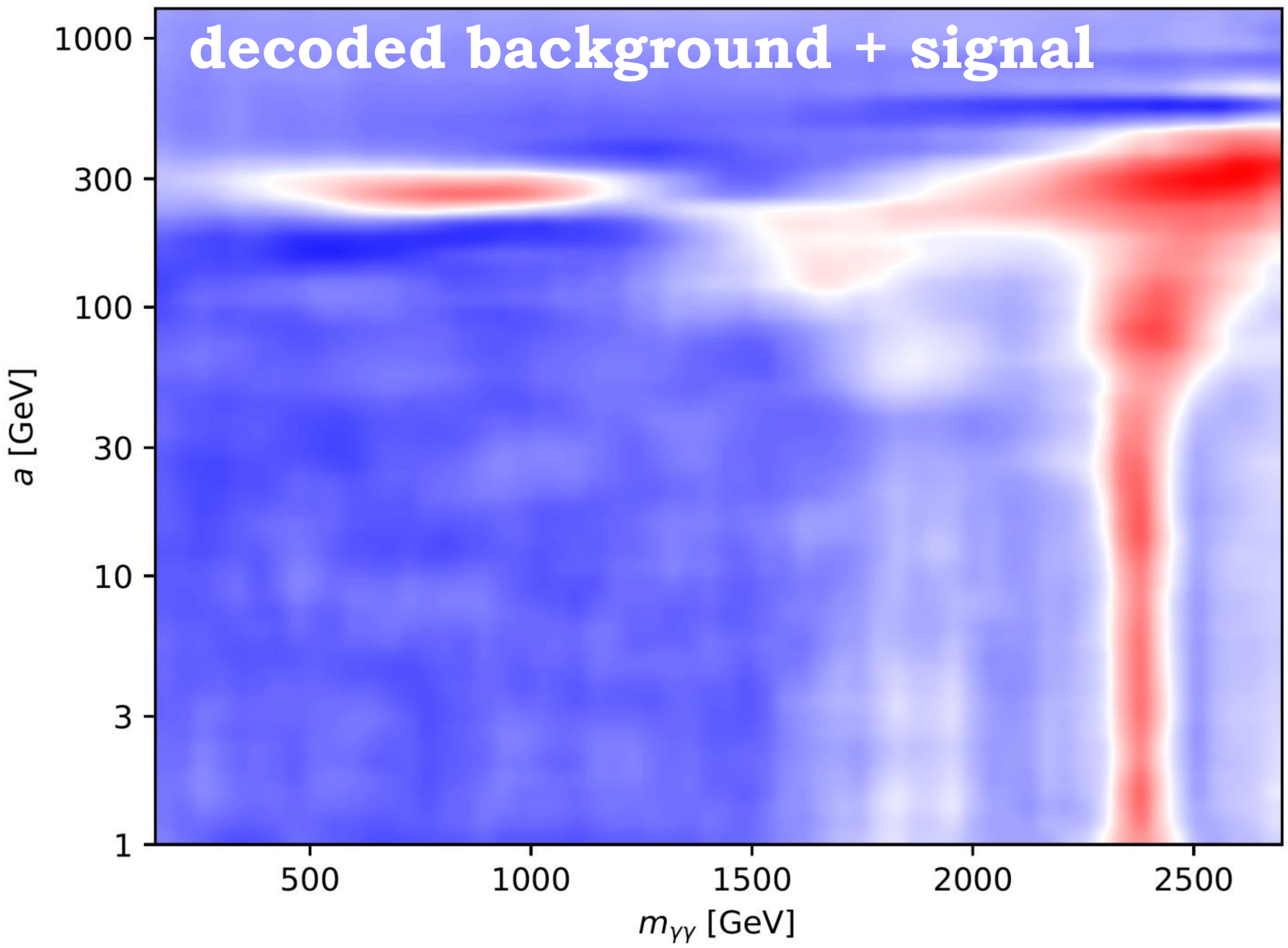
Method 5
Autoencoder neural network



Method 5
Autoencoder neural network



Method 5
Autoencoder neural network



Fourier transform for a comparison

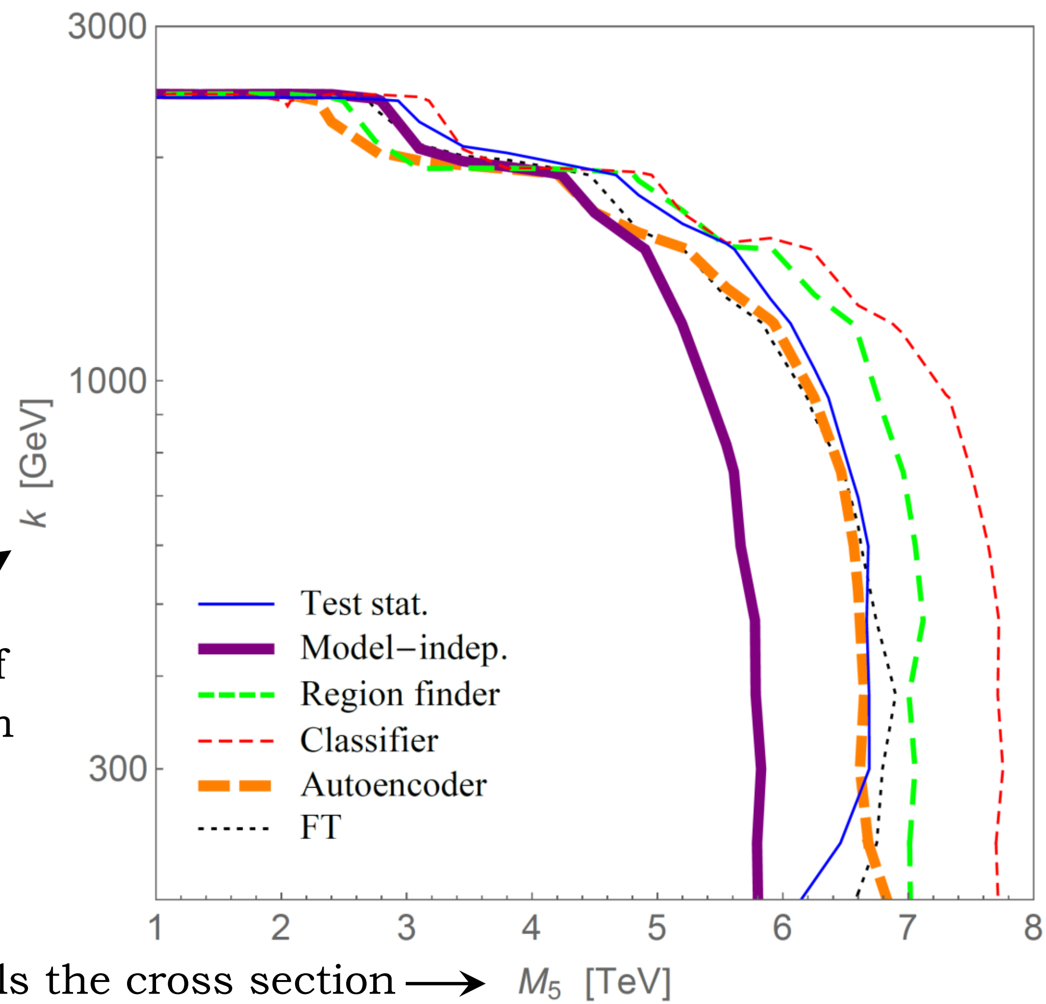
Define the windowed power spectrum:

$$P(T) = \left| \frac{1}{\sqrt{2\pi}} \int_{k}^{m_{\max}} dm \frac{d\sigma}{dm} \frac{1}{L} \exp\left(i \frac{2\pi g(m)}{T} \right) \right|^2$$

1. In the particular model, the spectrum is periodic in $g(m) = R\sqrt{m^2 - k^2}$, not simply m.
2. Use $P(1)$ as the test statistic.
3. Upper limit of integration is optimized for each point in the parameter space.

Sensitivity to the CW/LD scenario

Example corresponding to
37 fb^{-1} at 13 TeV
in the $\gamma\gamma$ channel
based on the ATLAS dataset
Phys. Lett. B 775 (2017) 105

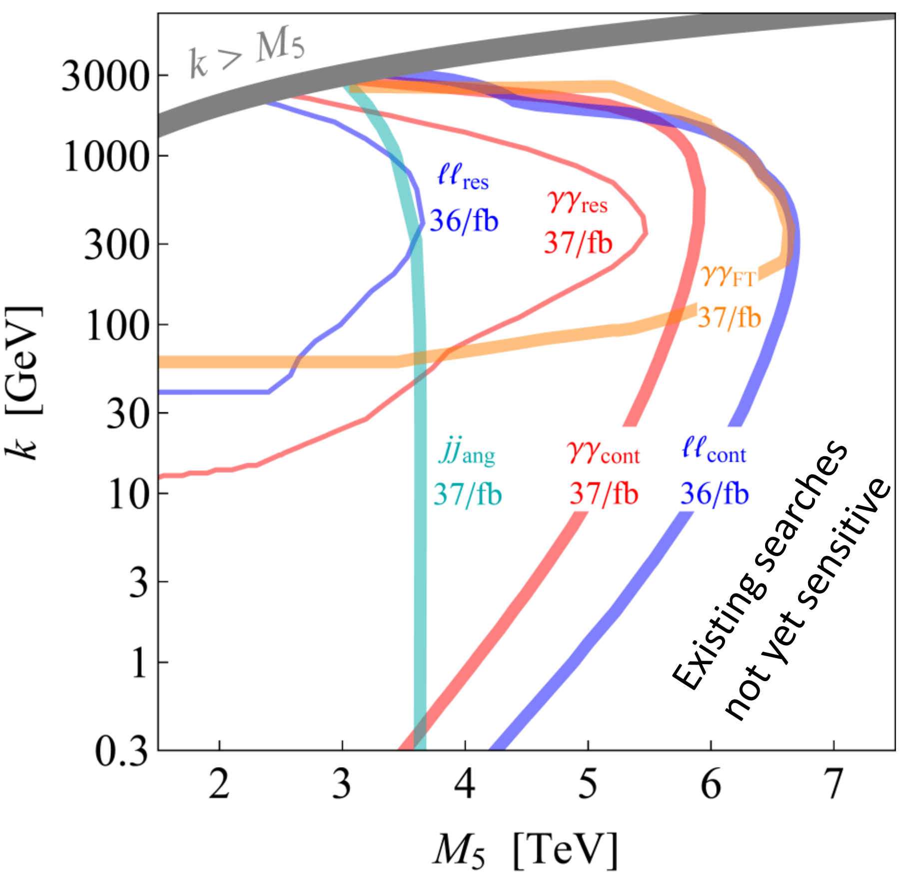


beginning of
the spectrum

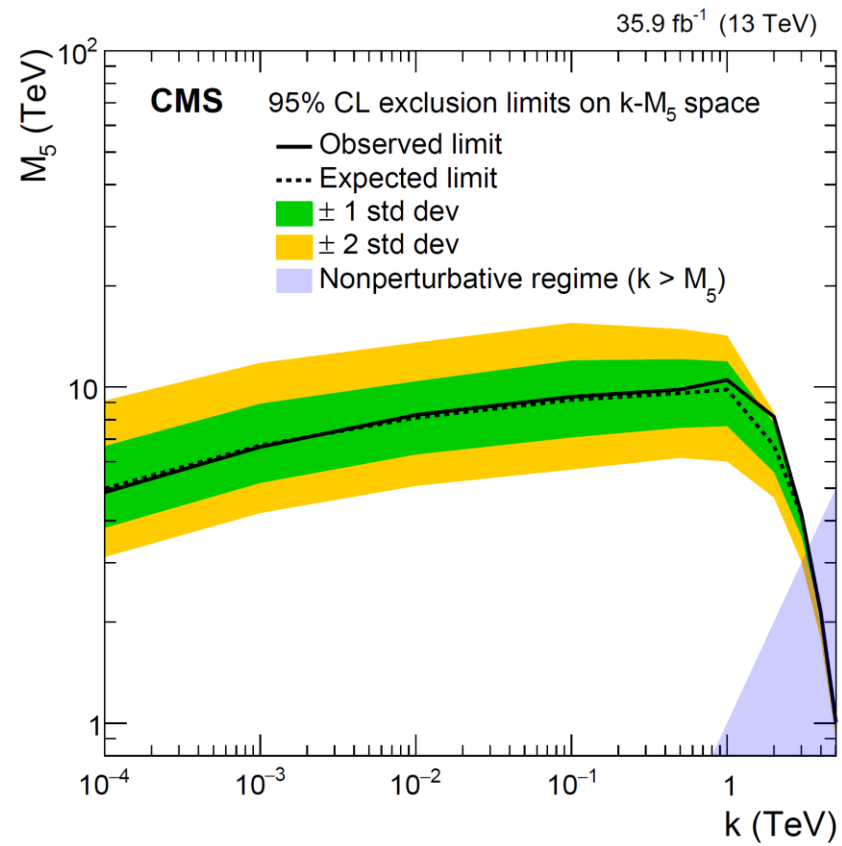
controls the cross section →

When comparing the methods, keep in mind that some (thick lines) are more model-independent than others (thin lines).

Sensitivity of other searches to the CW/LD scenario



Giudice, $\mathcal{YK}$, McCullough, Torre, Urbano
JHEP 06 (2018) 009 [arXiv:1711.08437]



CMS Collaboration
Phys. Rev. D 98 (2018) 092001
[arXiv:1809.00327]

Application to ATLAS data

ATLAS CONF Note

ATLAS-CONF-2023-010

23rd March 2023

Search for periodic signals in the dielectron and diphoton invariant mass spectra using 139 fb^{-1} of pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector

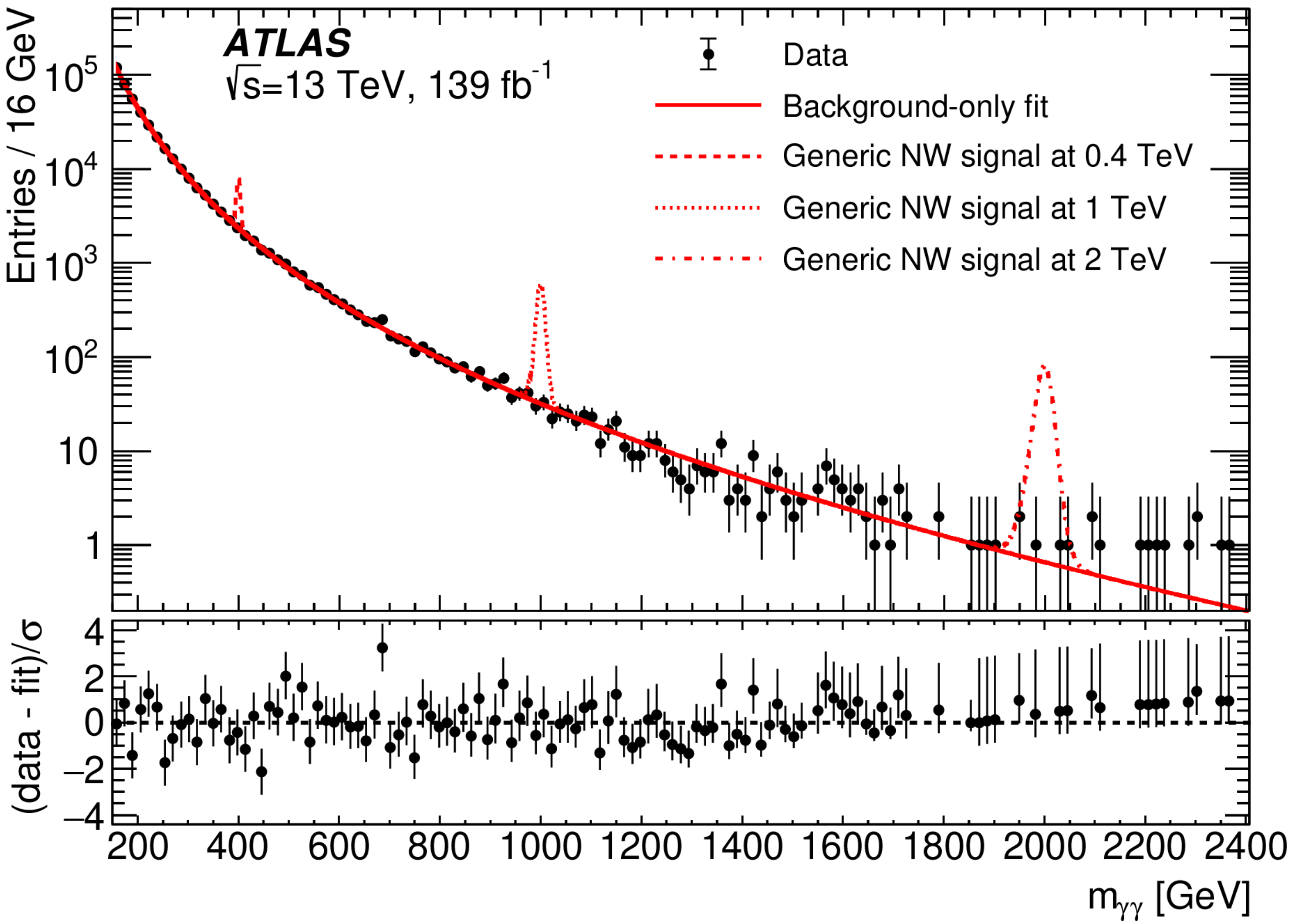
The ATLAS Collaboration

A search for physics beyond the Standard Model inducing periodic signals in the dielectron and diphoton invariant mass spectra is presented using 139 fb^{-1} of $\sqrt{s}$ = 13 TeV pp collision data collected by the ATLAS experiment at the LHC. Novel search techniques based on continuous wavelet transforms are used to infer the frequency of periodic signals from the invariant mass spectra and neural network classifiers are used to enhance the sensitivity to periodic resonances. In the absence of a signal, exclusion limits are placed at the 95% confidence level in the two-dimensional parameter space of the clockwork gravity model. Model-independent searches for deviations from the background-only hypothesis are also performed.

Application to ATLAS diphoton data

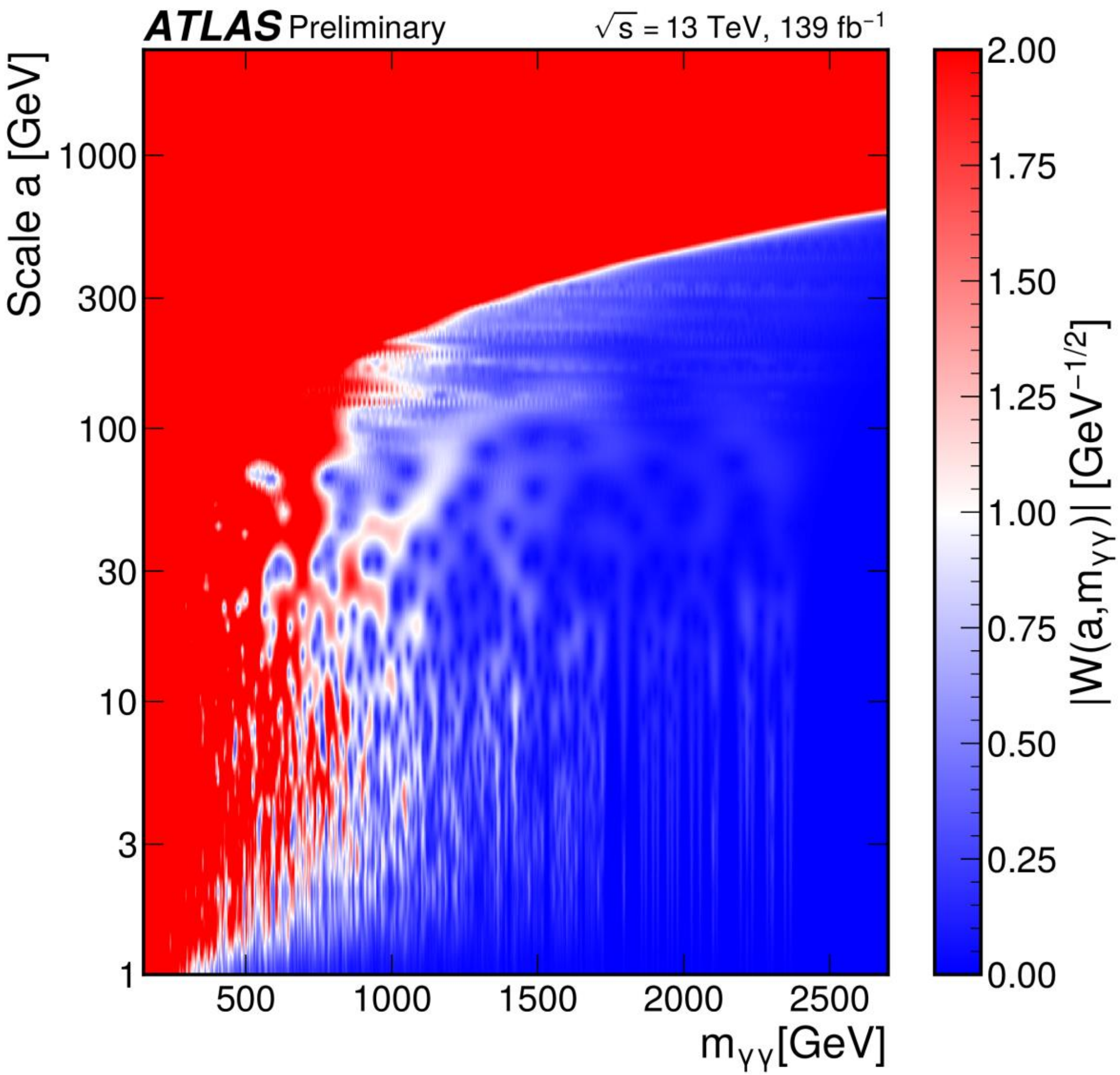
Original invariant mass spectrum

Phys. Lett. B 822 (2021) 136651



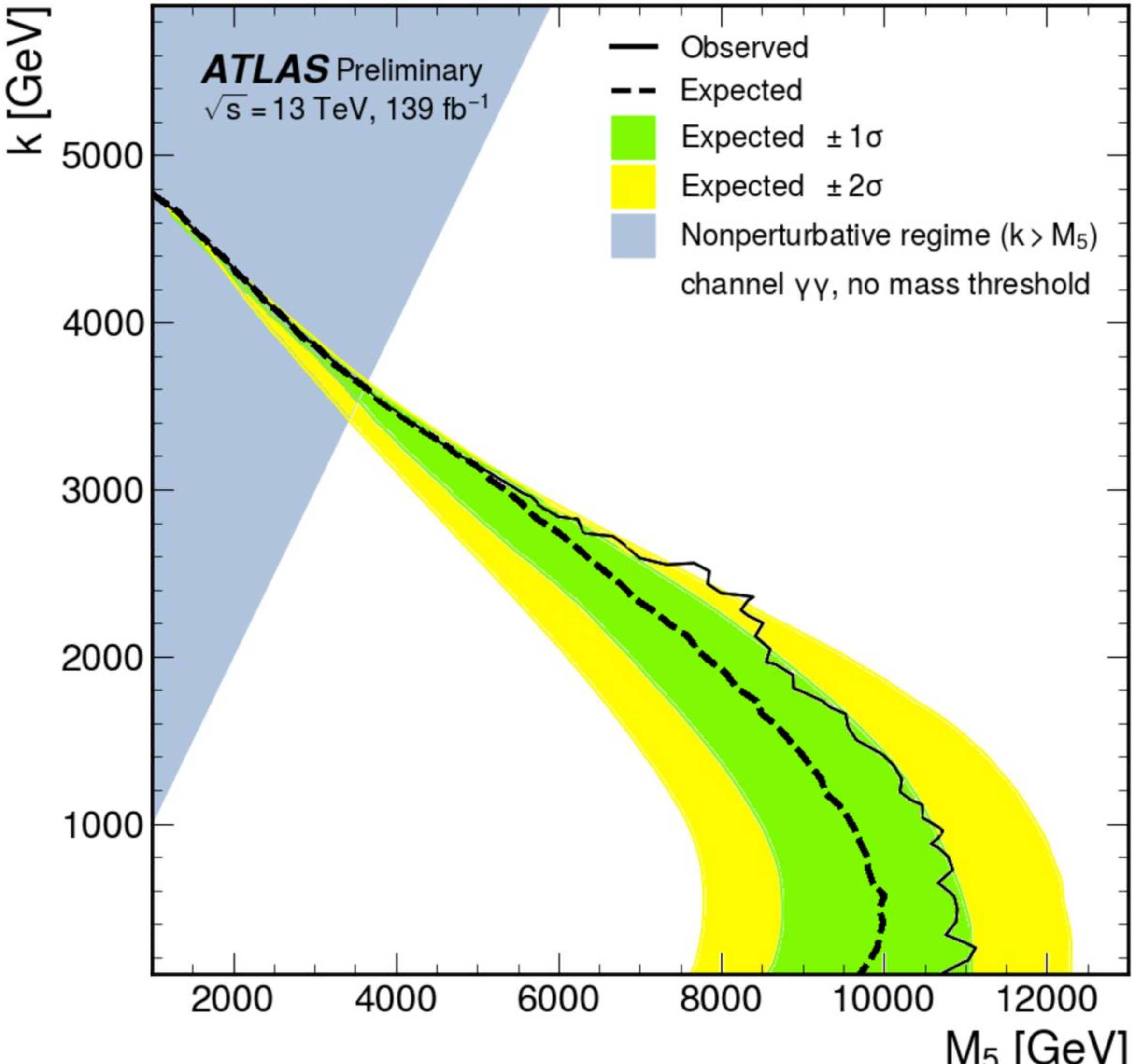
Application to ATLAS diphoton data

Wavelet transform



Application to ATLAS diphoton data

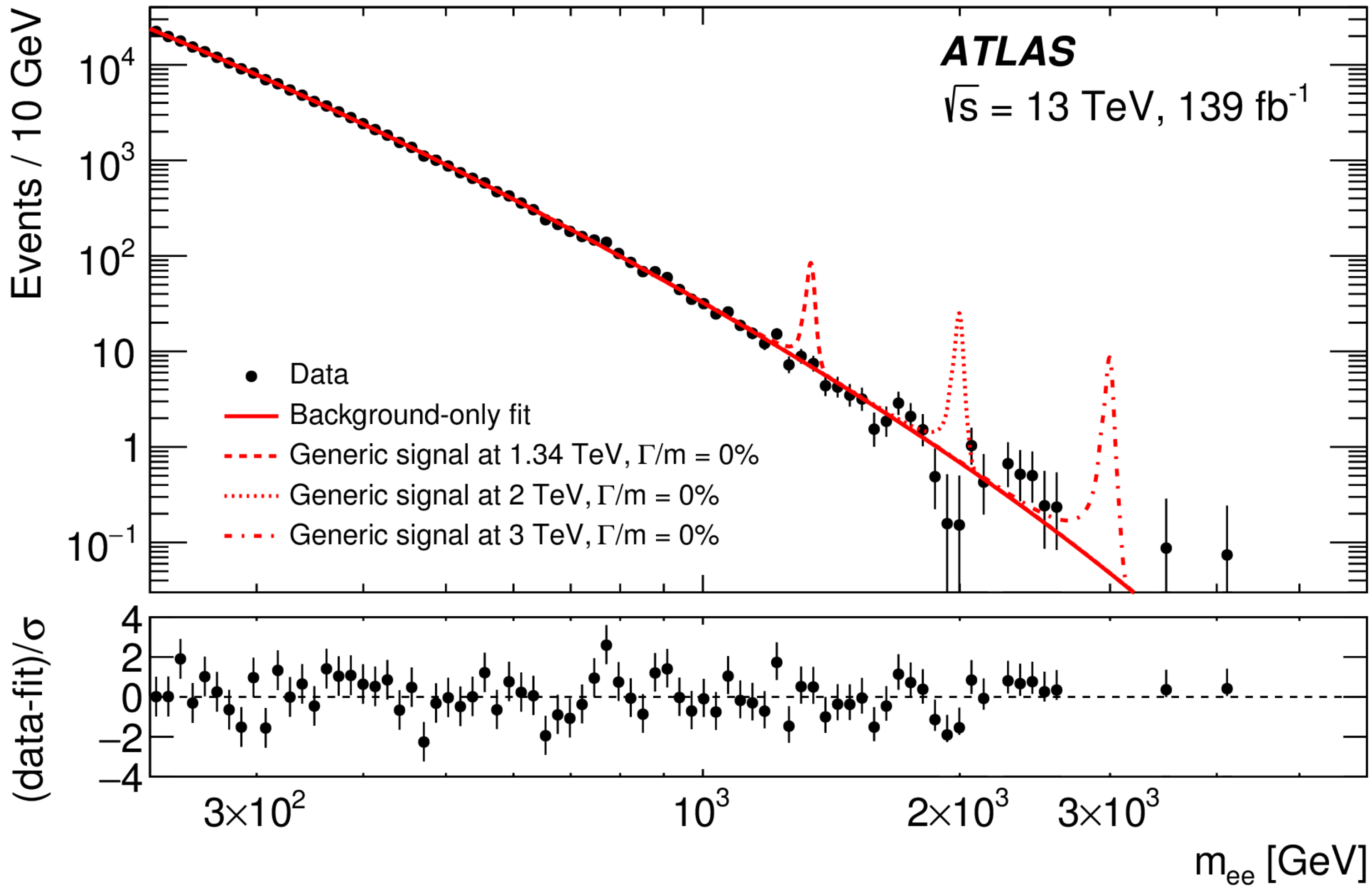
Classifier exclusion limit



Application to ATLAS dielectron data

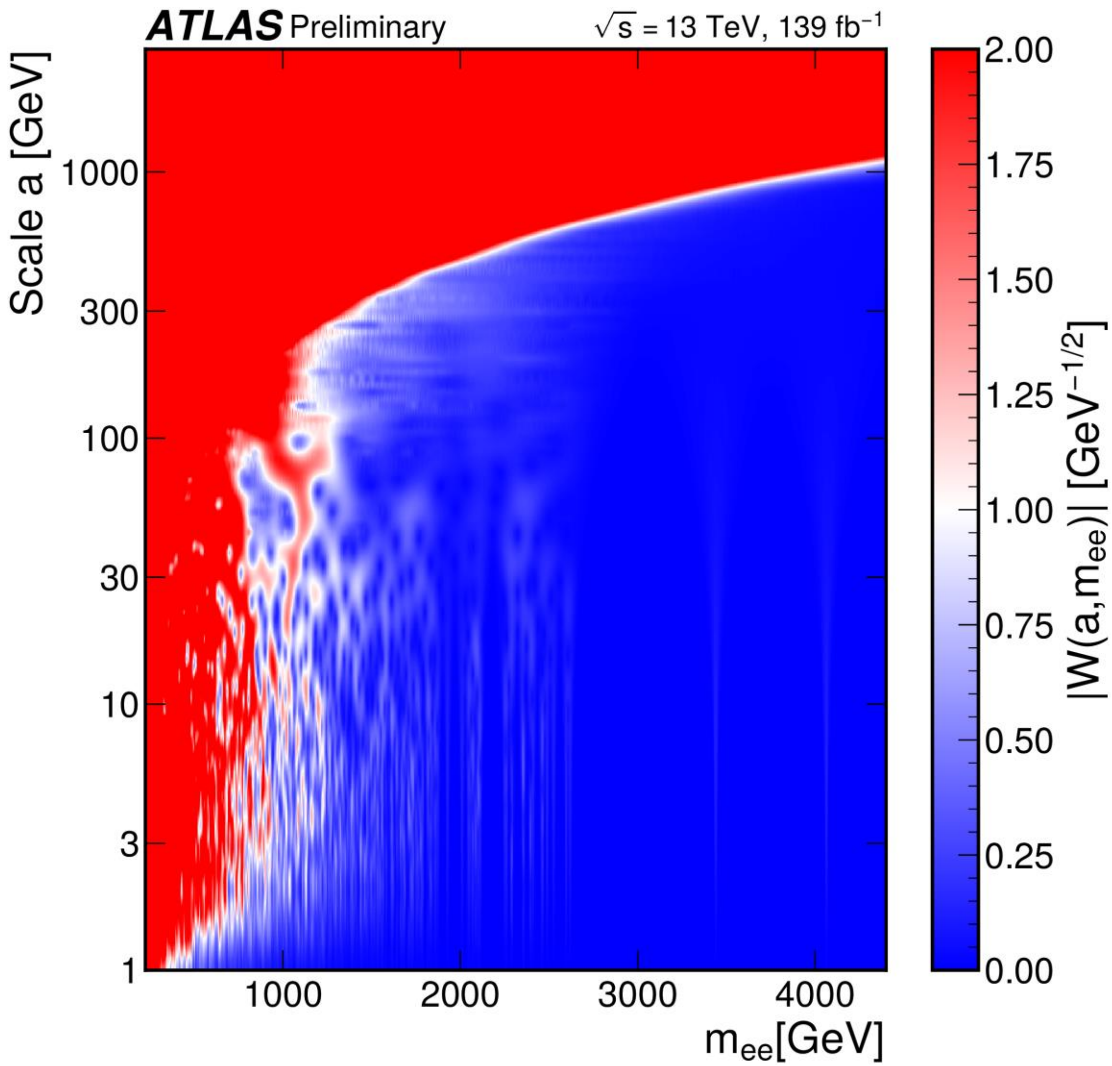
Original invariant mass spectrum

Phys. Lett. B 796 (2019) 68-87



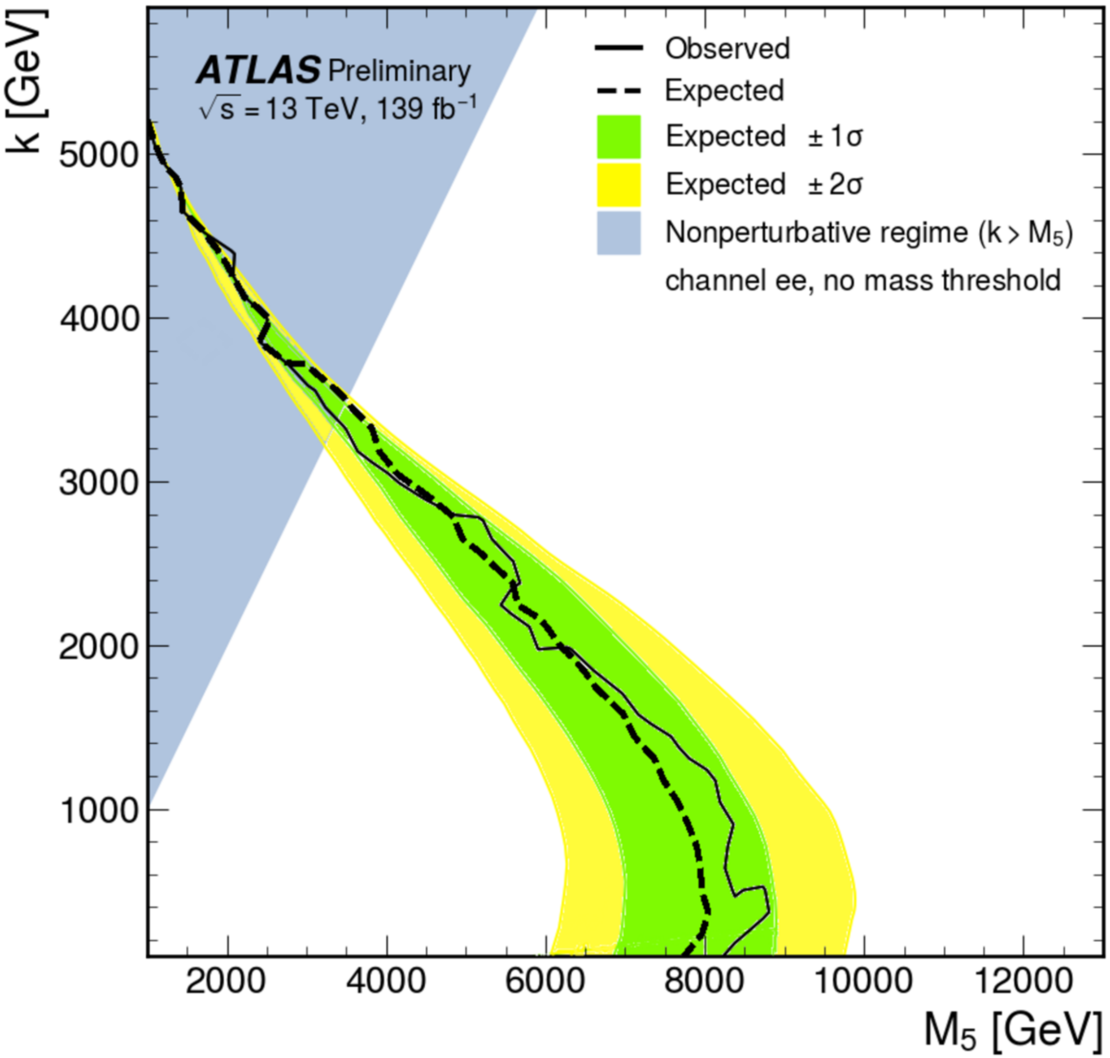
Application to ATLAS dielectron data

Wavelet transform



Application to ATLAS dielectron data

Classifier exclusion limit



Application to ATLAS data

Autoencoder results for both channels

Threshold	Dielectron Significance	Diphoton Significance
$R < 10\%$	0.4	-1.8
$R < 50\%$	1.5	-0.2
No threshold	0.7	-0.7
Scale threshold	1.5	-0.6

No significant anomaly detected.

2 Motivation for “dark jets” searches

- The Standard Model leaves many questions unanswered: electroweak-Planck hierarchy, dark matter, baryon asymmetry, neutrino masses, flavor structures etc.
    - Multiple additional particles likely exist.
        - Possible another (“dark”) QCD-like sector with its own gluons and quarks.
            - Couplings to the Standard Model could allow dark quark production, followed by showering and hadronization in the dark sector (“dark jets”), and observation via decays back to Standard Model particles as anomalous jets.

Strassler and Zurek, Phys. Lett. B 651 (2007) 374–379 and many others.

Motivation for “dark jets” searches

- Properties of dark jets are model-dependent:
    - Dark gauge group
    - Dark QCD scale
    - Number of dark quark flavors and their masses
    - Type of couplings to the Standard Model
    - Strength of couplings to the Standard Model

- Simulation of dark jets is challenging:
    - Only certain simple scenarios are implemented in Pythia etc.
    - QCD is strongly coupled at low energies. Even the simulation of our own QCD jets has large uncertainties, even after tuning to data (impossible for dark QCD).

- Therefore, model-independent and simulation-independent searches for anomalous jets are desirable.

Classes of scenarios

Scenarios may include:

- Different production mechanisms.
- Missing energy (due to long-lived dark hadrons) aligned with jets.
- Very / slightly displaced vertices (due to dark hadron decays).
- Decays to light/heavy quarks / leptons / photons.
- In each case – different backgrounds, different discriminators.
- Several proposals for searches, as well as actual searches, exist.

Schwaller, Stolarski, Weiler, JHEP 05 (2015) 059 *Emerging Jets*

Cohen, Lisanti, Lou, Phys. Rev. Lett. 115 (2015) 171804 *Semivisible Jets*

Cohen, Lisanti, Lou, Mishra-Sharma, JHEP 11 (2017) 196

Park, Zhang, Phys. Rev. D 100, (2019) 115009

CMS Collaboration, JHEP 02 (2019) 179 *first search for Emerging Jets*

Cohen, Doss, Freytsis, JHEP 09 (2020) 118

Bernreuther, Finke, Kahlhoefer, Kramer, Muck, SciPost Phys. 10 (2021) 046

Kar and Sinha, SciPost Phys. 10 (2021) 084

Canelli, de Cosa, Pottier, Niedziela, Pedro, Pierini, JHEP 02 (2022) 074

CMS Collaboration, JHEP 06 (2022) 156 *first search for Semivisible Jets*

and more…

Our class of scenarios

- The class of scenarios we address:
  - Resonant pair production of dark jets.
  - No missing energy in the jets.
  - Fully-hadronic decays to light quarks.
  - Decays are slightly displaced, with background due to b/c quarks.

- While the proposed search is model-independent, we test it on a set of toy benchmarks.

Production: $pp \to Z' \to q'\bar{q}'$

Decays:
$$\begin{aligned} \rho' &\to \pi'\pi' \\ \pi' &\to d\bar{d} \end{aligned}$$

gauge group	SU(3)
$\Lambda_{\mathrm{QCD}'}$	5 / 10 GeV
$n_{q'}$	1
$m_{q'}$	5 / 10 GeV
$m_{\pi'}$	5 / 10 GeV
$m_{\rho'}$	10.5 / 21 GeV
$c\tau_{\pi'}$	0.1 / 0.2 / 0.3 mm
r_{inv}	0

Weakly-supervised machine learning

Classification Without Labels (CWoLa)

Metodiev, Nachman, Thaler, JHEP 10 (2017) 174

Komiske, Metodiev, Nachman, Schwartz, PRD 98 (2018) 011502

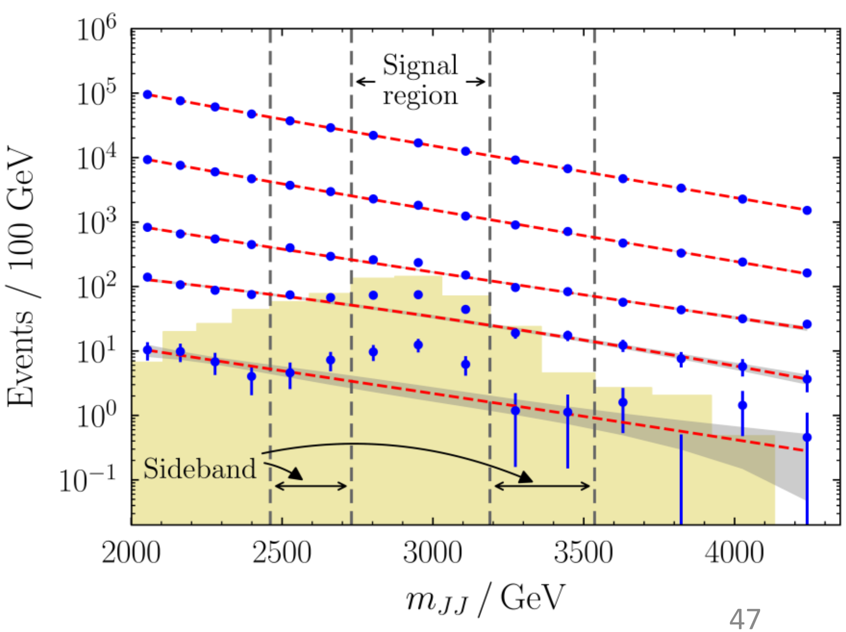
Samples with signal-rich and background-rich "weak labels" are defined, based on some assumed property of the signal. The classifier will figure out on its own which additional features can distinguish between the samples.

Example: Extended Bumphunt

The signal is assumed to be resonant in invariant mass.

Collins, Howe, Nachman, PRD 99 (2019) 014038

ATLAS collaboration, PRL 125 (2020) 131801

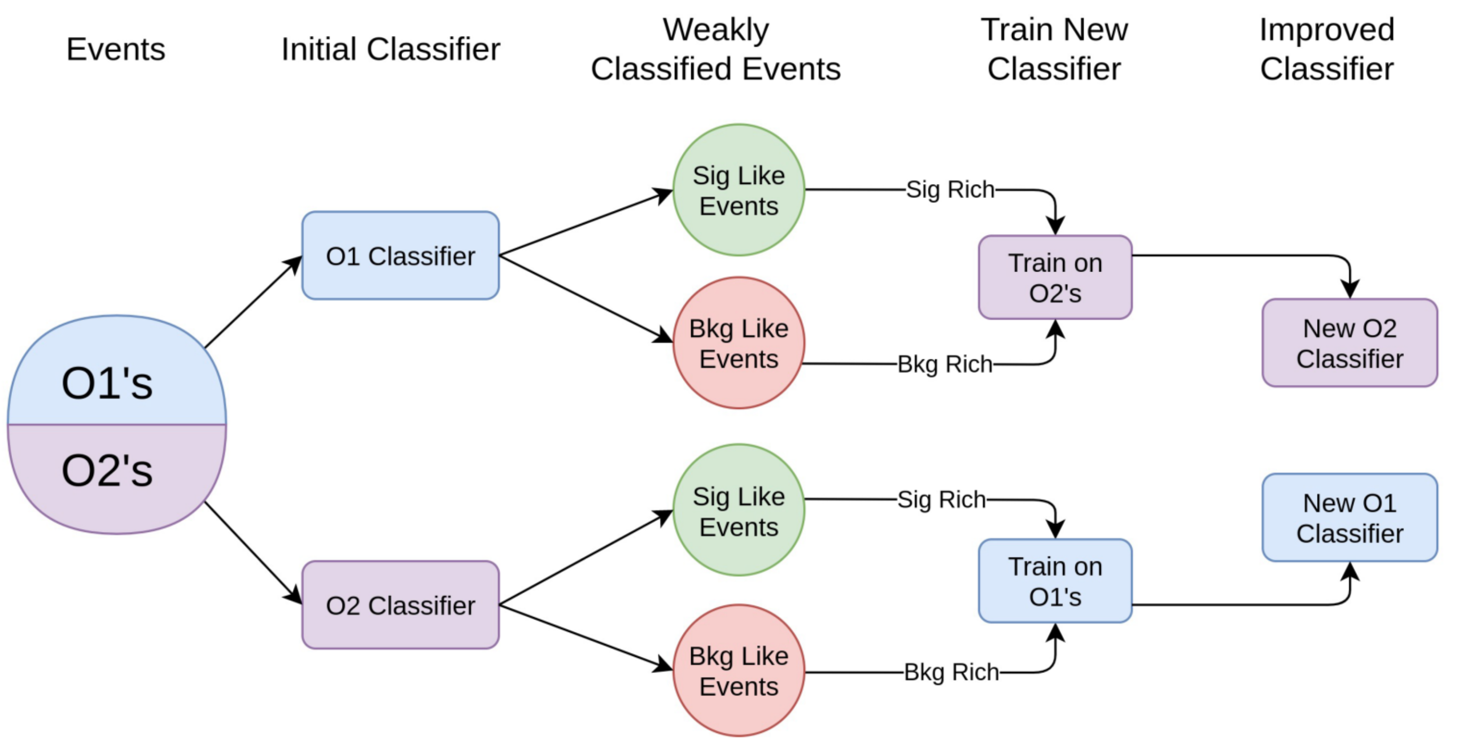


Weakly-supervised machine learning

Tag N' Train

Amram and Suarez, JHEP 01 (2021) 153

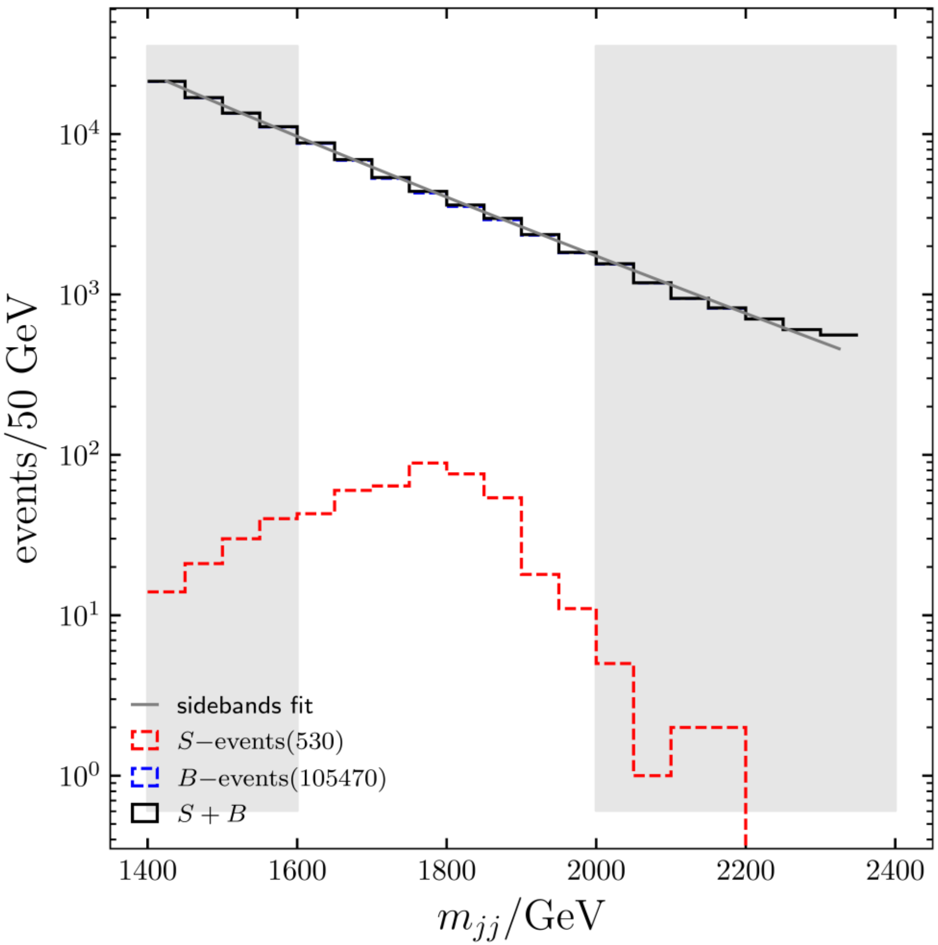
In cases when the events contain either a pair of signal objects or a pair of background objects, assign a weak label based on one of the objects and train on the features of the other.



Our work

General strategy

1. For an assumed mass for the particle producing the dark jets, select search region + sidebands in dijet invariant mass (m_{jj}).



Our work

General strategy

1. For an assumed mass for the particle producing the dark jets, select search region + sidebands in dijet invariant mass (m_{jj}).
2. To train and infer on the same dataset, split the data into k-folds to infer on each k-fold with a NN trained on the rest of the events.

Pizza in Erice

Our work

General strategy

1. For an assumed mass for the particle producing the dark jets, select search region + sidebands in dijet invariant mass (m_{jj}).
2. To train and infer on the same dataset, split the data into k-folds to infer on each k-fold with a NN trained on the rest of the events.
3. Consider the two leading jets in the event: j_1 and j_2.
   Use cuts on the multiplicity n_{obj} (number of objects) of j_2 to form weak labels for j_1:

   S-rich if $n_{obj} > n_{obj}^{S}$ and m_{jj} in the search region

   B-rich if $n_{obj} < n_{obj}^{B}$

Our work

General strategy

1. For an assumed mass for the particle producing the dark jets, select search region + sidebands in dijet invariant mass (m_{jj}).
2. To train and infer on the same dataset, split the data into k-folds to infer on each k-fold with a NN trained on the rest of the events.
3. Consider the two leading jets in the event: j_1 and j_2.
   Use cuts on the multiplicity n_{obj} (number of objects) of j_2 to form weak labels for j_1:
   S-rich if $n_{\mathrm{obj}} > n_{\mathrm{obj}}^{\mathrm{S}}$ and m_{jj} in the search region
   B-rich if $n_{\mathrm{obj}} < n_{\mathrm{obj}}^{\mathrm{B}}$
4. Use these labels to train a NN classifier on features of j_1.
   The specific features we used will be described in the following.

Our work

General strategy

1. For an assumed mass for the particle producing the dark jets, select search region + sidebands in dijet invariant mass (m_{jj}).
2. To train and infer on the same dataset, split the data into k-folds to infer on each k-fold with a NN trained on the rest of the events.
3. Consider the two leading jets in the event: j_1 and j_2.
   Use cuts on the multiplicity n_{obj} (number of objects) of j_2 to form weak labels for j_1:
   S-rich if $n_{\mathrm{obj}} > n_{\mathrm{obj}}^{\mathrm{S}}$ and m_{jj} in the search region
   B-rich if $n_{\mathrm{obj}} < n_{\mathrm{obj}}^{\mathrm{B}}$
4. Use these labels to train a NN classifier on features of j_1.
5. Do the same with the j_1 and j_2 roles interchanged.
6. Apply the NNs to data that wasn't used for training.
7. Cut on the product of the two NN outputs, interpolate the sidebands to the search region and check for a bump.

Simulation tools

- MadGraph model with Z' couplings to SM quarks and dark quarks. Cohen, Lisanti, Lou, Mishra-Sharma, JHEP 11 (2017) 196
- Pythia 8 for QCD and dark sector showering and hadronization.
- Detector simulation with Delphes, with added smearing of track parameters.
- Vertex reconstruction with the AVR algorithm implemented in RAVE.

Event selection

Event selection

Standard ATLAS/CMS dijet event selection

+ displaced vertices requirement.

p_T^{jet}	> 150 GeV
$\lvert\eta\rvert^{\text{jet}}$	< 2
m_{jj}	> 1133 GeV
$\lvert y^*\rvert$	< 0.8
$\Delta\phi(jj)$	> 1
$\sum_{\text{disp.vert.}} p_T^{\text{vertex}}/p_T^{\text{jet}}$	> 0.2

ATLAS Collaboration
JHEP 03 (2020) 145

displaced vertices requirement

Dominant backgrounds

$b\bar{b}$ (~50%)

jj w/gluon splitting (~37%)

Features

- Displaced vertex mass

$$m^2_{\text{vertex}} = \left(\sum_{\text{tracks}} \sqrt{\mathbf{p}^2_{\text{track}} + m^2_{\pi^\pm}} \right)^2 - \left(\sum_{\text{tracks}} \mathbf{p}_{\text{track}} \right)^2$$

- Displaced object lifetime estimate: vertex transverse displacement D_0 divided by $\gamma\beta_T = \frac{p_T^{\text{vertex}}}{m_{\text{vertex}}}$

- Momentum fraction in the vertex tracks: $p_T^{\text{vertex}}/p_T^{\text{jet}}$

- Vertex track count.

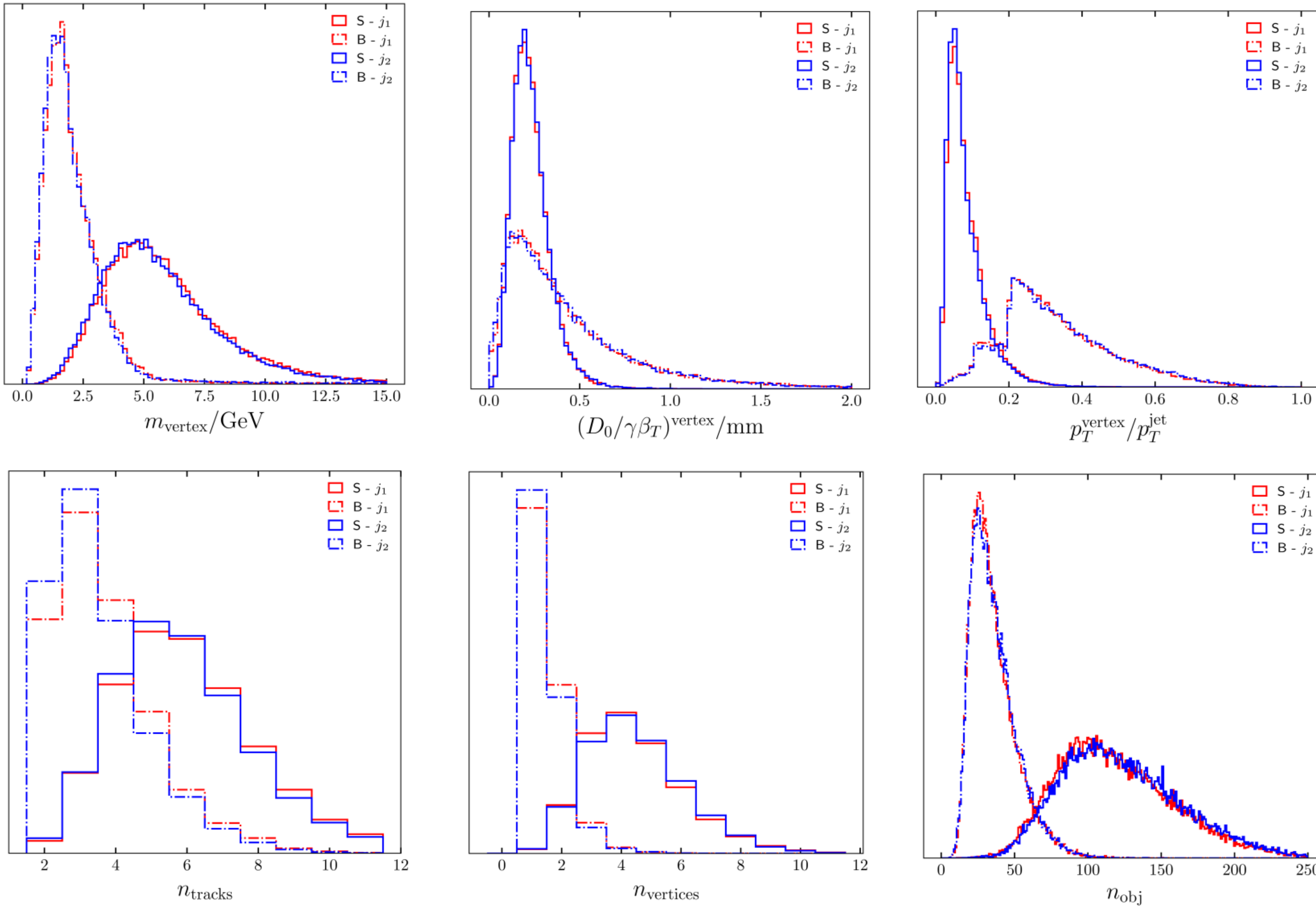
If multiple vertices are present, median values are taken for the above features.

- Number of displaced vertices in the jet.

- Number of objects in the jet (n_{obj}).

Example distributions

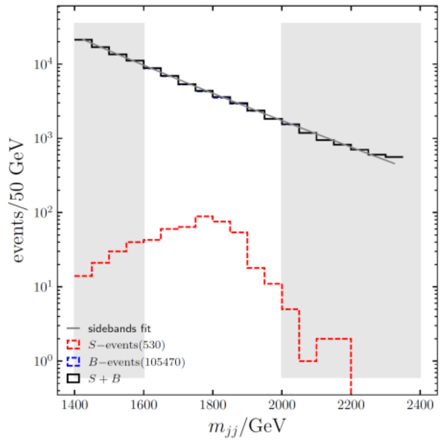
$$(m_{\pi'}, c\tau_{\pi'}) = (10\ \text{GeV},\ 0.2\ \text{mm})$$



NN architecture

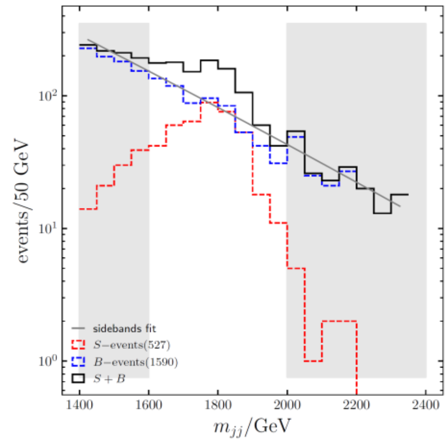
```
_________________________________________________________________
Layer (type)                 Output Shape              # Params
=================================================================
Layer-1 (Dense)              (None, 32)                224
_________________________________________________________________
activation (LeakyReLU)       (None, 32)                0
_________________________________________________________________
dropout (Dropout)            (None, 32)                0
Layer-2 (Dense)              (None, 16)                528
_________________________________________________________________
activation (ELU)             (None, 16)                0
_________________________________________________________________
dropout (Dropout)            (None, 16)                0
Layer-3 (Dense)              (None, 16)                272
_________________________________________________________________
activation (ELU)             (None, 16)                0
_________________________________________________________________
dropout (Dropout)            (None, 16)                0
Layer-4 (Dense)              (None, 4)                 68
_________________________________________________________________
activation (ELU)             (None, 4)                 0
Output (Dense)               (None, 1)                 5
=================================================================
Total params: 1,097
_________________________________________________________________
```

Bump hunt



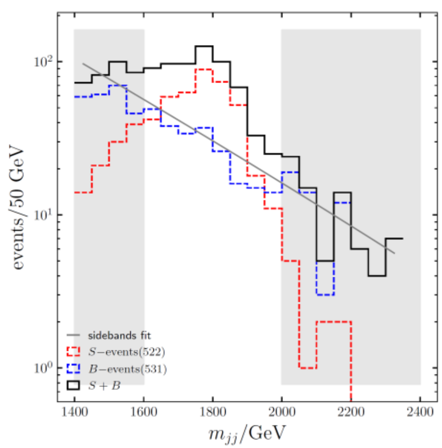
	$n_S^{\text{sig.reg.}}$	$n_B^{\text{sig.reg.}}$	significance
true	415	35851	2.2σ
sideband fit	-300	36566	-0.72σ

(a) all events



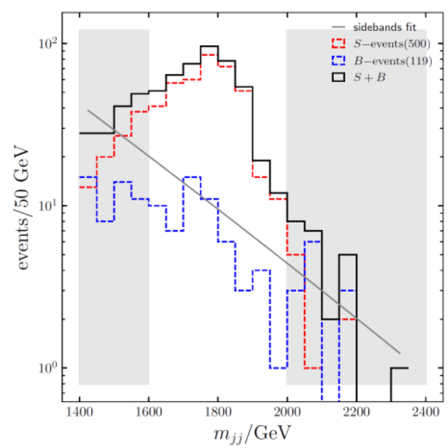
	$n_S^{\text{sig.reg.}}$	$n_B^{\text{sig.reg.}}$	significance
true	413	648	16σ
sideband fit	365	696	5.9σ

(b) $\epsilon_D = 2\%$



	$n_S^{\text{sig.reg.}}$	$n_B^{\text{sig.reg.}}$	significance
true	408	229	27σ
sideband fit	378	259	9.4σ

(c) $\epsilon_D = 1\%$



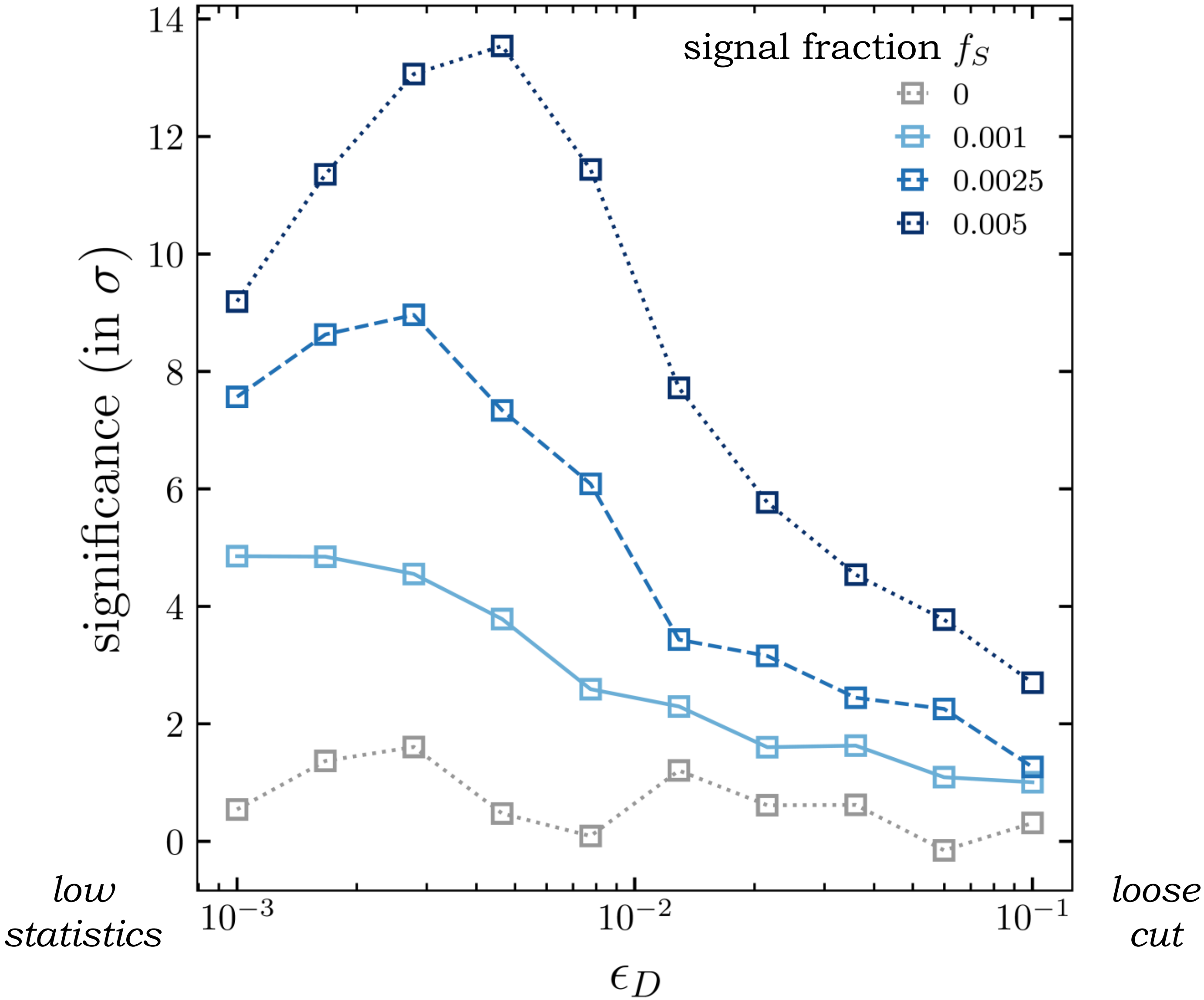
	$n_S^{\text{sig.reg.}}$	$n_B^{\text{sig.reg.}}$	significance
true	392	57	52σ
sideband fit	366	83	14σ

(d) $\epsilon_D = 0.6\%$

Example corresponds to:

- Z' mass: 2 TeV
- Dark hadron parameters:
  $$(m_{\pi'}, c\tau_{\pi'}) = (10\ \text{GeV},\ 0.2\ \text{mm})$$
- signal fraction (including sidebands): 0.5%
- Luminosity: 800 fb^{-1}

Bump significance




Thank You!

Supplementary Slides

KK gravitons in CW/LD scenario

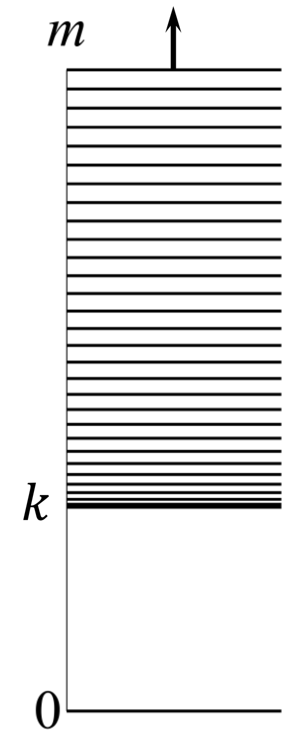
Masses

$$m_0^2 = 0 \qquad m_n^2 = k^2 + \frac{n^2}{R^2} \qquad n = 1, 2, 3, \ldots$$

Couplings to Standard Model

$$\mathcal{L} \supset -\frac{1}{\Lambda_n} h_{\mu\nu}^{(n)} T^{\mu\nu}$$

$$\Lambda_0^2 = M_P^2 \qquad \Lambda_n^2 = M_5^3 \pi R \left(1 + \left(\frac{kR}{n}\right)^2\right)$$

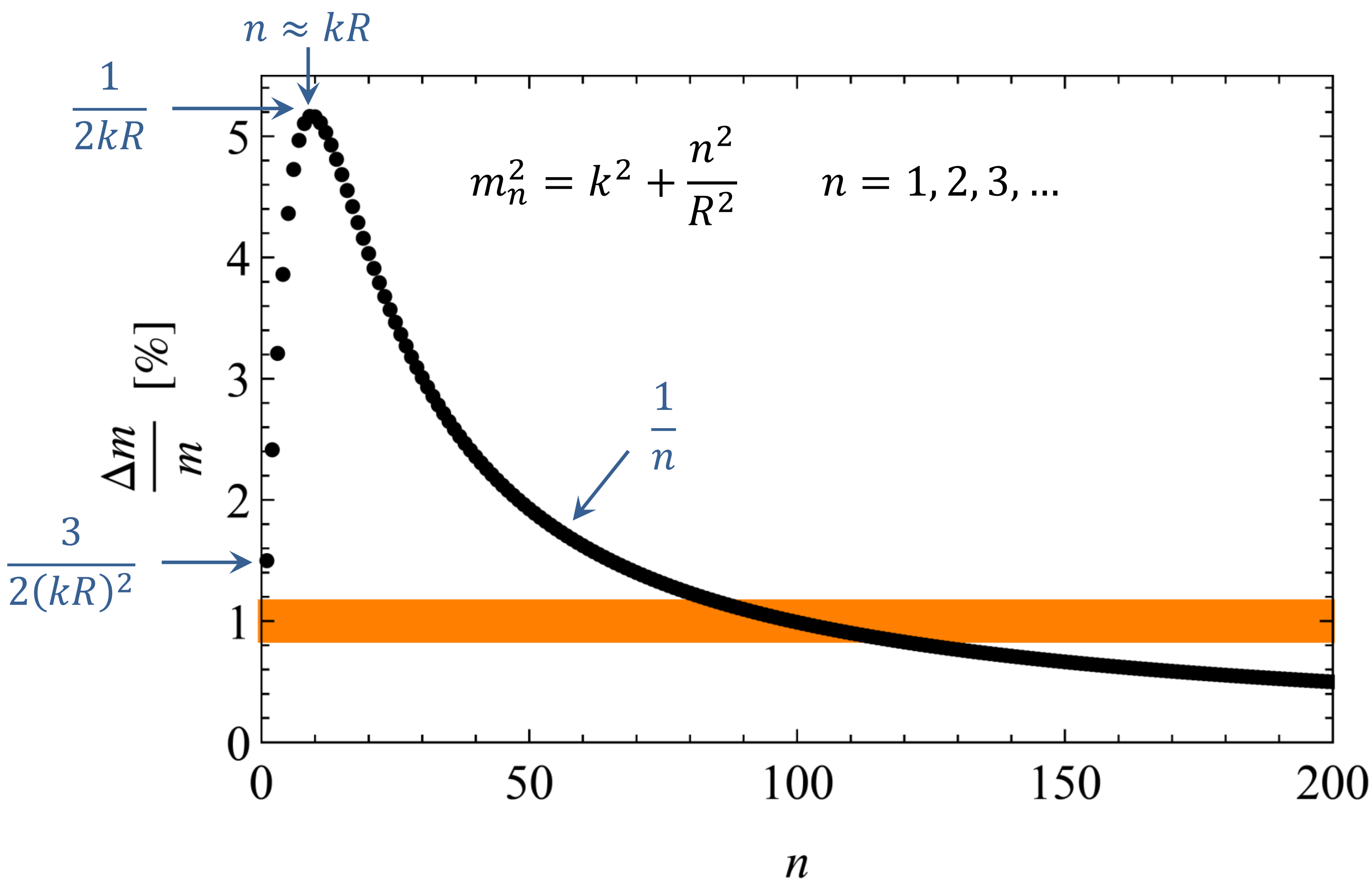


Production via $T_{\mu\nu}$ from gg and $q\bar{q}$.

Decays (1) To SM particle pairs via $T_{\mu\nu}$

(2) To pairs of lighter KK modes via 5D gravity self-interactions

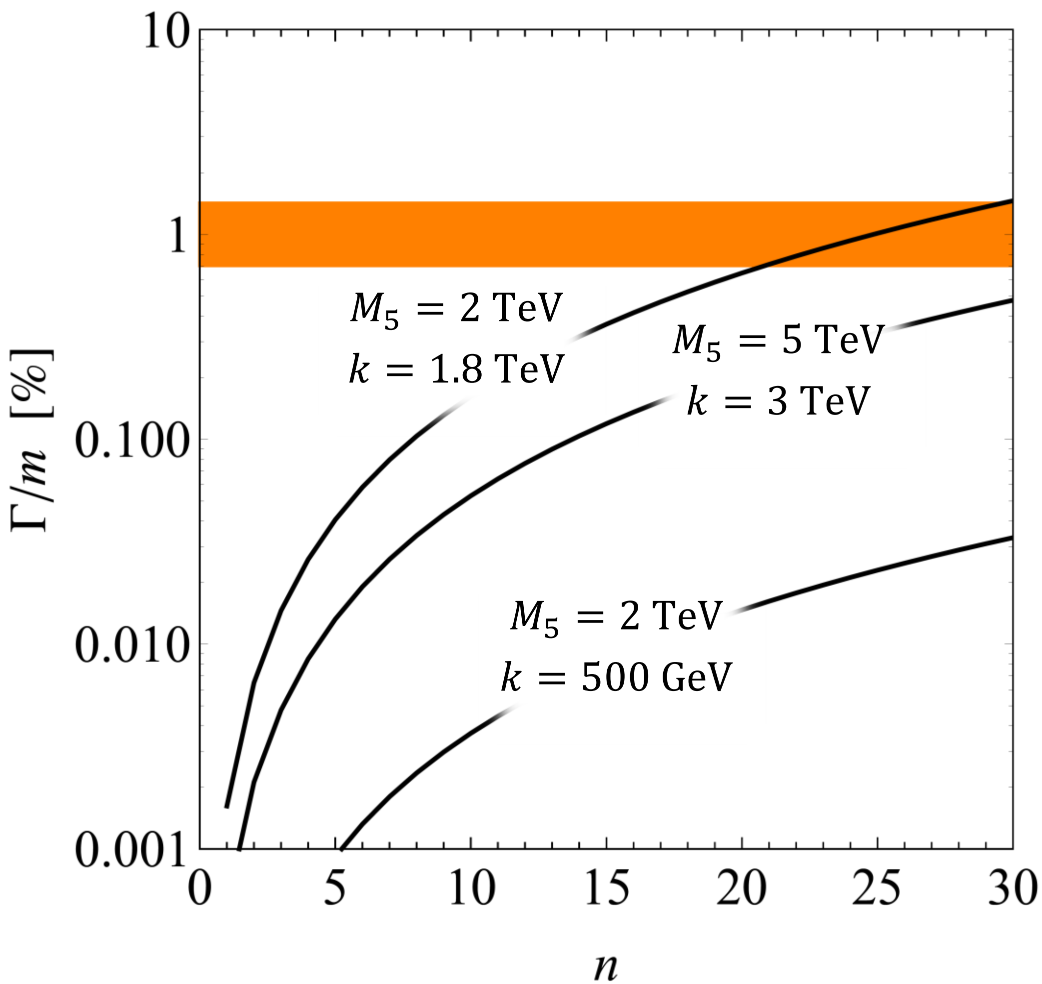
KK graviton mass splittings



For $n \lesssim 100$, i.e. $k \lesssim m_n \lesssim 10k$, the individual modes can be resolved in the $\gamma\gamma$ and e^+e^- channels in ATLAS and CMS!

KK graviton widths

The intrinsic widths are small. For at least the first ~30 modes, the width is dominated by the experimental resolution.



KK graviton decays

Decays to SM particles

gg	$\sum_i q_i\bar{q}_i$	W^+W^-	ZZ	hh	$\gamma\gamma$	$\sum_i \ell_i^+\ell_i^-$	$\sum_i \nu_i\bar{\nu}_i$
34%	38%	9.2%	4.6%	0.35%	4.2%	6.4%	3.2%

*when phase space suppressions are negligible

Easiest decays to see: $\gamma\gamma$, e^+e^-, $\mu^+\mu^-$

Total rate to SM particles (for $n \gg kR,\ m_n \gg m_t$):

$$\Gamma_{n\to\mathrm{SM}} \simeq \frac{283}{960\pi^2}\frac{m_n^3}{RM_5^3}$$

Decays to pairs of lighter KK gravitons

For $n \gg kR \gg 1$:

$$\Gamma_{n\to\mathrm{KK}} \simeq \frac{5\cdot 7\cdot 17}{3\cdot 2^{14}\pi^2}\frac{\sqrt{km_n}\,m_n^3}{kRM_5^3} \quad \Rightarrow \quad \frac{\Gamma_{n\to\mathrm{KK}}}{\Gamma_{n\to\mathrm{SM}}} \approx 0.04\sqrt{\frac{m_n}{k}}$$

A small effect on the BR, except for very low k.

Autoencoder loss local *p*-values