

My Thesis

by

Name

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Department)
in the University of Michigan
2020

Doctoral Committee:

Thomas Schwarz, Chair
Dante Amidei
Junjie Zhu
James Liu
Eric Bell

Name

name@umich.edu

ORCID ID: My Orcid ID

© Name 2020

“Questing Physicks isn’t like the Quiet and Queer branches. You can’t do it at home in a comfortable chair—you have to be out in the thick of the business, with your tools on your belt and your heart on your sleeve!” - Catherynne Valente, *The Girl Who Fell Beneath Fairyland and Led The Revels There ??.*

For

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

LIST OF TABLES

TABLE

LIST OF APPENDICES

APPENDIX

LIST OF ABBREVIATIONS

ABSTRACT

My Abstract.

CHAPTER 1

An Overview of the Standard Model

1.1 The Standard Model

The Standard Model of Particle Physics is arguably one of the crowning achievements of the last century of physics research. Though it is doubtless incomplete (it does not, for instance, explain the dark energy or dark matter observed in cosmological experiments, nor does it provide a satisfactory quantum-mechanical model of gravity), all predictions it has made have yet to be falsified ??, ??, ??.

The Standard Model is a quantum field theory, meaning that it describes the behavior of fields (physical quantities that are defined at all points in spacetime; common examples of fields include electric and magnetic fields) and their discrete, quantized excitations, referred to as particles (common examples of particles are the electron and the photon, which are excitations of the "electron field" and the electromagnetic field respectively). More about the mathematics of field dynamics will be discussed in section ??.

The Standard Model divides these fundamental particles (and their corresponding fields) into two major categories: the bosons and the fermions. Fermions carry half-integer spin, while bosons carry integer spin. In a general sense, the elementary fermions can be thought of as the particles that comprise matter, while the elementary bosons can be thought of as the particles that correspond to the behavior of the four fundamental forces (electromagnetism, the strong nuclear force, the weak nuclear force, and gravity).

The elementary fermions of the Standard Model all have spin 1/2. They are divided into two major subgroupings, quarks and leptons; each of these is further divided into three generations. Each generation of quarks contains one up-type quark (up, charm, or top) and one down-type quark (down, strange, and bottom), while each generation of leptons contains one charged lepton (electron, muon, or tau) and one neutrino (electron neutrino, muon neutrino, or tau neutrino). Fermions of successive generations behave similarly to each other, though those of each subsequent generation are more massive than the last.

The quarks are the only fermions that undergo both the strong and weak nuclear interactions, and are also the only particles in the Standard Model that have fractional electromagnetic charge (+2/3 for up-type quarks and -1/3 for down-type quarks). However, because quarks are never found in isolation and are always bound to other quarks in composite states called hadrons, all observable quark final states have integer charge. Conversely, the charged leptons all have an electromagnetic charge of -1, while neutrinos are chargeless. While the quarks and charged leptons have precisely-measured masses, the mass of the neutrinos is vanishingly small, and, to date, only the differences between the different neutrino masses have been conclusively measured.

In addition to these 12 fermion species, each fermion species has a corresponding antimatter "antifermion" species, which has the opposite charge and parity quantum numbers. This is discussed at length in section ??.

The bosons of the standard model are the "messengers" through which the fundamental forces operate (with the exception of the Higgs boson, which is detailed at length in section ??). The photon and the gluons are massless, while the two W-bosons, the Z-boson, and the Higgs boson are all massive. The photon is the mediator of the electromagnetic force and couples to charged particles; the gluon is the mediator of the strong nuclear force and couples to particles with "color charge", and the W and Z bosons are the mediators of the weak nuclear interaction, coupling to all fermions with left-handed parity and antifermions with right-handed parity. In addition, the three weak bosons (W^+ , W^- , and Z) can all couple to each other, as can the gluons.

The strong nuclear interaction binds quarks together, while the weak interaction governs the decay of one species of fermion into another. Weak interactions operate primarily on fermion doublets, coupling each up-type quark to its corresponding down-type quark and each charged lepton to its corresponding neutrino. However, some intergenerational coupling does occur, the rarity of which is governed by a matrix of coefficients called the Cabibbo-Kobayashi-Makasawa (or CKM) matrix.

The Higgs boson is the only scalar (spin-0) boson in the Standard Model. Though it does not mediate any force directly, its existence is a consequence of the unification of the electromagnetic and weak nuclear interactions into one "electroweak" interaction at high energy scales. Without the role of the Higgs boson in this process, the fermions, the W-bosons, and the Z-boson would all be massless; thus, the Higgs can be said to "give mass" to the particles of the Standard Model. It couples to all massive particles in the Standard Model, namely, the fermions and the W and Z bosons.

Standard Model of Elementary Particles

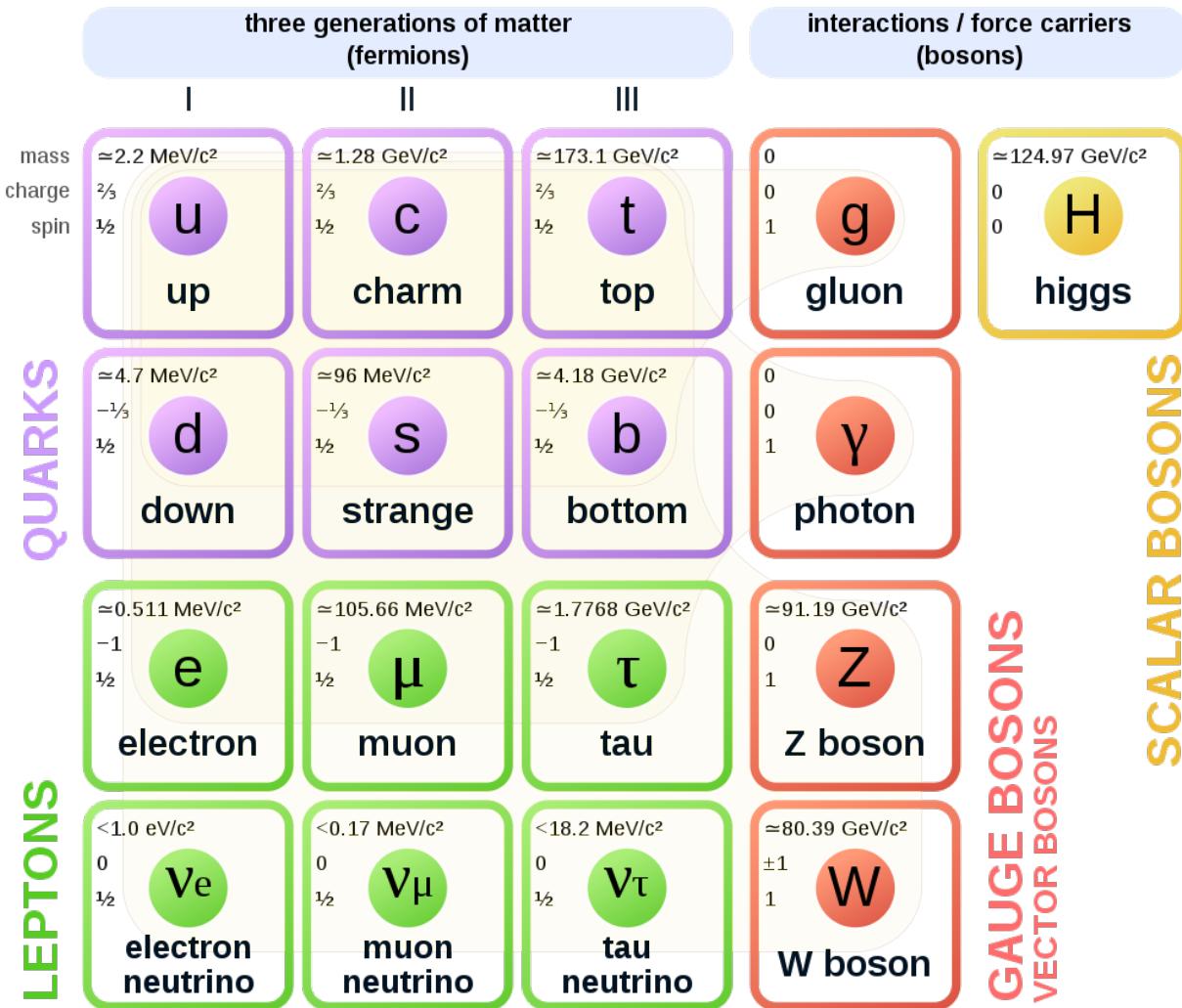


Figure 1.1: The "periodic table" of the Standard Model, depicting the three generations of fermions, the gauge bosons, and the Higgs. Taken from reference ??.

1.1.1 Lagrangians, Fields, and Gauge Transformations

In order to fully explain the Higgs mechanism, we must first discuss the mathematical language of quantum field theory. Both quantum and classical field theories are often discussed using Lagrangian dynamics, where the Lagrangian is defined as $\mathcal{L} = T - V$, the difference of the kinetic and the potential energy. Physical properties will always evolve in such a way that the integral of this property with respect to time, $S = \int \mathcal{L} dt$, called the action, is a constant. Lagrangians are also incredibly useful in that they give rise to conservation laws: Noether's Theorem states that operations performed on a system that do not change the Lagrangian are each associated to conserved quantities of a system (i.e., systems with translationally invariant Lagrangians must respect conservation of momentum, systems with temporally-invariant Lagrangians must respect conservation of energy, etc.).

A variety of types of fields exist, but we will discuss three at length: Klein-Gordon fields, which are scalar fields (single quantities defined everywhere), Dirac spinor fields (vector quantities defined everywhere), and gauge fields (additional vector fields that must be introduced in order to preserve certain physical symmetries). We begin with the Klein-Gordon field.

Klein-Gordon fields, one of the simplest examples of a field, are real scalar-valued quantities often denoted using the symbol ϕ . A non-interacting "free" Klein-Gordon field evolves according to the Lagrangian:

$$\begin{aligned} \mathcal{L} &= T - V \\ &= \frac{1}{2} \frac{\partial^2 \phi}{\partial t^2} - \frac{1}{2} (\nabla \phi)^2 - \frac{1}{2} m^2 \phi^2 \\ &= -\frac{1}{2} (\partial^\mu \phi) (\partial_\mu \phi) - \frac{1}{2} m^2 \phi^2 \end{aligned} \tag{1.1}$$

where we utilize Einstein sum notation in the last line to compress the four derivatives in the preceding expression into a single shorthand symbol. The Lagrangian of an interacting Klein-Gordon field would look similar, but would possess additional V potential terms depending on the nature of the interaction. As the only scalar in the Standard Model, the observable Higgs boson is the only elementary particle to follow an interacting Klein-Gordon field equation.

While the concept of a complex-valued scalar field does not directly correspond to any of the physical elementary particles that are present in the Standard Model, it plays an important role in the understanding of the Higgs mechanism. A complex-valued scalar field behaves similarly to a real-valued one, with the Lagrangian:

$$\mathcal{L} = (\partial^\mu \phi^*) (\partial_\mu \phi) - \frac{1}{2} m^2 \phi^* \phi \tag{1.2}$$

Where $*$ denotes the complex conjugate of the scalar field.

Finally, Dirac fields describe all Standard Model fermions (with the possible exception of neutrinos). They are vector fields as opposed to scalar fields, and behave according to the Lagrangian:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \quad (1.3)$$

where γ^μ denotes the set of Dirac gamma matrices, and $\bar{\psi} = \psi^\dagger \gamma^0$ denotes the transpose of the complex conjugate of the vector field multiplied by one of these matrices, defined as such in order to preserve invariance under relativistic Lorentz boosts.

A four-component Dirac vector field can be written in a variety of representations, but one of the most useful is that of a doublet of two two-component Weyl spinors, one left-handed and one right-handed, that is, $\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$. Each of these components transforms differently under Lorentz boosts.

In order to discuss the gauge fields corresponding to the Standard Model bosons, we must first discuss the concept of gauge symmetries. Consider a single Dirac vector field described by equation ???. We transform the field by rotating it by a constant phase α : $\psi \rightarrow e^{i\alpha}\psi$, $\bar{\psi} \rightarrow e^{-i\alpha}\bar{\psi}$.

$$\begin{aligned} \mathcal{L} &= -(e^{-i\alpha}\bar{\psi})(i\gamma^\mu \partial_\mu - m)(e^{i\alpha}\psi) \\ &= -(e^{-i\alpha}e^{i\alpha})\bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \\ &= -\bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \end{aligned} \quad (1.4)$$

We see that we recover the original Dirac Lagrangian. Thus, the lagrangian is invariant under a constant phase rotation. A transformation of this character is called a global gauge transformation. A one-dimensional rotation is a unitary transformation, so we call this a global $U(1)$ gauge symmetry.

We next consider the concept of a local gauge transformation, that is, one in which the phase α may vary with position. The field transforms, as before, like $\psi \rightarrow e^{i\alpha(x)}\psi$, $\bar{\psi} \rightarrow e^{-i\alpha(x)}\bar{\psi}$. However, we note that the dependence on position means that we can no longer factor out the exponentials, and we thus have

$$\begin{aligned} \mathcal{L} &= -(e^{-i\alpha(x)}\bar{\psi})(i\gamma^\mu \partial_\mu - m)(e^{i\alpha(x)}\psi) \\ &= -ie^{-i\alpha(x)}\bar{\psi}\gamma^\mu(e^{i\alpha(x)}\partial_\mu\psi + ie^{i\alpha(x)}\psi\partial_\mu\alpha - m\bar{\psi}\psi) \\ &= -\bar{\psi}(i\gamma^\mu \partial_\mu + \gamma^\mu \partial_\mu\alpha - m)\psi \end{aligned} \quad (1.5)$$

i.e., the Lagrangian is not invariant under this sort of transformation. However, local gauge invariance is an important physical symmetry, so in order to attempt to preserve it, we add an additional term to the original Lagrangian: an extra vector field A_μ that transforms like $A_\mu \rightarrow A_\mu - \frac{1}{q}\partial_\mu\alpha(x)$ for a constant q . We also define a new operator, called the covariant derivative: $D_\mu = \partial_\mu + iqA_\mu$. Given how A_μ transforms, we see that D_μ transforms like

$$\begin{aligned}
D_\mu &= \partial_\mu + iqA_\mu \\
\rightarrow \partial_\mu &+ iq(A_\mu - \frac{1}{q}\partial_\mu\alpha(x)) \\
&= \partial_\mu + iqA_\mu - i\partial_\mu\alpha(x) \\
&= D_\mu - i\partial_\mu\alpha(x)
\end{aligned} \tag{1.6}$$

Let us replace the $\partial_\mu\phi$ terms in the initial Lagrangian with $D_\mu\phi$ and transform.

$$\begin{aligned}
\mathcal{L} &= -(e^{-i\alpha(x)}\bar{\psi})(i\gamma^\mu(D_\mu - i\partial_\mu\alpha(x)) - m)(e^{i\alpha(x)}\psi) \\
&= -(e^{-i\alpha(x)}\bar{\psi})(i\gamma^\mu(\partial_\mu + iqA_\mu - i\partial_\mu\alpha(x)) - m)(e^{i\alpha(x)}\psi) \\
&\quad = -\bar{\psi}(i\gamma^\mu D_\mu + \gamma^\mu\partial_\mu\alpha - \gamma^\mu\partial_\mu\alpha - m)\psi \\
&\quad = -\bar{\psi}(i\gamma^\mu D_\mu - m)\psi
\end{aligned} \tag{1.7}$$

Thus, by introducing an additional vector field that corresponds to the local U(1) gauge symmetry, we have restored the invariance of our lagrangian. Physically, this field is analogous to the introduction of electromagnetism to our single Dirac fermion model, with the A_μ field playing the role of the photon: it is a vector quantity and so has spin-1, it must be massless (as adding an A_μ mass term to the Lagrangian would violate the symmetry again), and couples to the fermion fields according to their electromagnetic charge. We note that we can also still preserve invariance if we add an additional term $L_{kin} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$ to the Lagrangian, where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$: this corresponds to the energy inherent in electromagnetic fields themselves.

Each of the fundamental forces in the Standard Model can be understood in terms of these sorts of gauge symmetries. The photon is the particle excitation of the electromagnetic field, which corresponds to a one-dimensional rotation "U(1)" transformation. In order to be invariant under three dimensional transformations of the type "SU(3)" (the set of all volume-preserving, $\bar{\psi}\psi$ -preserving transformations in a 3D vector space), we must add eight new vector fields (these are the eight gluons), leading to the incorporation of the strong interaction into the Standard Model.

Similarly, in order to be invariant under two-dimensional transformations of the type "SU(2)" (the set of all volume-preserving, $\bar{\psi}\psi$ -preserving transformations in a 2D vector space), we must add three new vector fields. However, these cannot be the observed weak-interaction bosons, the W^+, W^- and Z : as mentioned before, the new fields must be massless, as adding a mass term for these bosons would violate the local gauge symmetry. How, then, can the masses of the weak bosons fit into the Standard Model? The answer lies in the introduction of yet another field, called the Higgs field, the particle excitation of which is the much-lauded Higgs boson.

1.2 CP-Symmetry

In addition to the $SU(2)_L \times U(1) \times SU(3)$ gauge symmetry, the Standard Model also respects a discrete symmetry called CPT. This is the combined product of three separate symmetries: charge conjugation (C), parity inversion (P), and time reversal (T). While some interactions may respect a limited combination of these symmetries (i.e., just C, or the product CP), all Standard Model processes must respect the product of the three: that is, for a given physics process, if we invert all the charges, flip all the parities, and run the interaction backward, the resulting new process must also be physically allowed.

The strong nuclear, electromagnetic, and gravitational interactions are symmetric under C, P, and T individually. The weak nuclear interaction, however, is invariant only under the combination of all three: it violates C, P, and, in some cases, CP. The CP violation that occurs in the weak interaction is through the CKM off-diagonal quark-mixing matrix mentioned briefly in section ??.

A parity inversion is equivalent to reversing a particle's momentum without reversing its spin. For a fermion field $\psi(t, x)$, a parity transform P takes $\psi(t, x)$ to $\psi(t, -x)$, and looks like $P\psi(t, x)P = e^{i\theta}\gamma^0\psi(t, -x)$ for some constant matrix θ .

Given this, we find that the expression $P\bar{\psi}\psi P = \bar{\psi}\psi$, so $\bar{\psi}\psi$ is a *scalar* under parity, while the expression $iP\bar{\psi}\gamma^5\psi P = -i\bar{\psi}\gamma^5\psi$ acquires a minus sign, so $i\bar{\psi}\gamma^5\psi$ is a *pseudoscalar* under parity.

A time-reversal is equivalent to "running a process backward". For example, a swinging pendulum is a process that is approximately T symmetric, as it looks similar "played" in forward or reverse, while a plate shattering on the ground is not a T symmetric process, as reversing it does not make physical sense (broken plates do not spontaneously reform). For a fermion field $\psi(t, x)$, a time-reversal transform T takes $\psi(t, x)$ to $\psi(-t, x)$, and looks like $T\psi(t, x)T = -\gamma^1\gamma^3\psi(-t, x)$. Given this, we find that the expression $T\bar{\psi}\psi T = \bar{\psi}\psi$, so $\bar{\psi}\psi$ is a *scalar* under parity, while the expression $iT\bar{\psi}\gamma^5\psi T = -i\bar{\psi}\gamma^5\psi$ acquires a minus sign, so $i\bar{\psi}\gamma^5\psi$ is a pseudoscalar under time-reversal.

A charge conjugation is a reversal of all charges (electric charge, weak hypercharge, and strong-nuclear color charge). For a fermion field $\psi(t, x)$, a charge conjugation transform C takes ψ to $\bar{p}\bar{\psi}$, and looks like $C\psi C = -i(\bar{\psi}\gamma^0\gamma^2)^T$ (where T here denotes the matrix transpose). Given this, we find that the expression $C\bar{\psi}\psi C = \bar{\psi}\psi$, so $\bar{\psi}\psi$ is a *scalar* under charge conjugation, as is the expression $iC\bar{\psi}\gamma^5\psi C = i\bar{\psi}\gamma^5\psi$ as well. Thus, Lagrangians containing terms of the form $i\bar{\psi}\gamma^5\psi$ are invariant under C and CPT, but not P, T or CP.

Applying C and P together is equivalent to switching particles and antiparticles. However, we also know that our universe is composed almost entirely of matter, and not equal parts matter and antimatter (living in a universe in which matter and antimatter are equally abundant would be very unpleasant, as the two would be constantly annihilating each other, bathing us in a near-constant

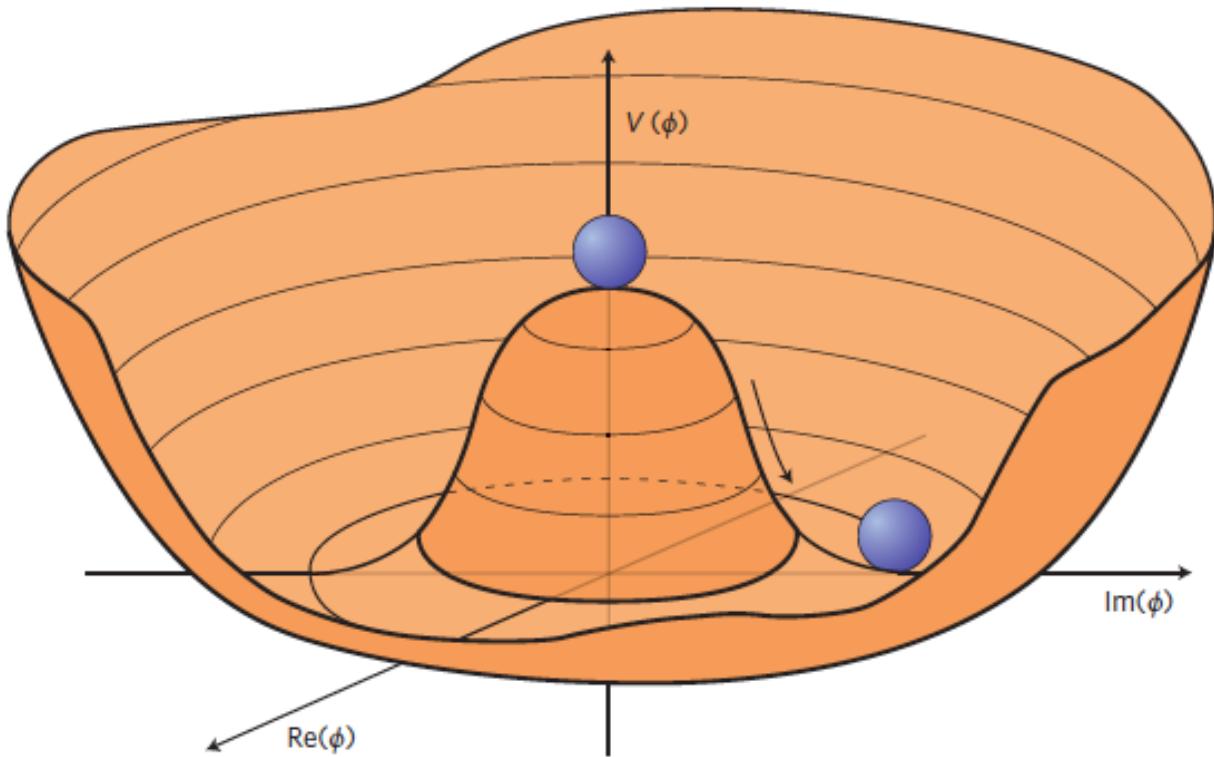


Figure 1.2: The "wine bottle" Higgs potential hill, from reference ??

shower of gamma rays). Thus, some physics process that occurs in the high-energy regime near the Big Bang must violate CP in a substantial way. The CP-invariance that occurs through the CKM mixing is not enough to account for the matter-antimatter asymmetry, so searching for this source of CP violation is of pressing interest to experimental physicists.

1.3 The Higgs Mechanism and Electroweak Symmetry Breaking

Given the results of previous sections, it should be clear that the Higgs field is important not simply because it "gives particles mass" (an often-made claim which is, in a sense, true), but because it is a vital missing piece of the Standard Model that is necessary to reconcile the elegant mathematical language of the fundamental interactions with the particles we observe in real-world experiments.

To understand the Higgs mechanism, we must first devote a brief detour to the concept of spontaneous symmetry breaking. This occurs when an unstable, continuous symmetric state spontaneously changes into a more stable, asymmetric one. Consider, for example, the "wine-bottle" potential shape depicted in figure ??.

When the ball is balanced at the top of the potential "hill", the configuration is spatially symmetric: that is, no direction is privileged over any other. However, when this delicate balance is even slightly disturbed, the ball will roll down the hill in some direction, resulting in a final state that is *not* symmetrical. This is the phenomenon known as "spontaneous symmetry breaking".

Using the language of quantum field theory, we can easily add spontaneously-breakable symmetric potential terms to a Lagrangian. If we do so, the phenomenon of symmetry breaking allows us to rewrite these terms as a combination of massless fields (one for each "choice" the symmetry breaking must make) and massive fields (corresponding to the leftover degrees of freedom). If we rewrite the "wine-bottle" potential in this way, for example, the one new massless field corresponds to the ball's angular position along the circle at the base, while the massive field corresponds to the ball's radial position "up" or "down" the hill. The massless particles that arise from these massless fields are known as Goldstone bosons.

We are now ready to discuss the Higgs mechanism. We can combine the electromagnetic and weak nuclear forces as different manifestations of the same force, called the electroweak force, that transforms like $U(1) \times SU(2)$. We write the terms of the Standard Model electroweak Lagrangian, noting that, since the weak interaction is observed to be chiral, it couples differently to left- and right-handed fermions.

In this case, the generator of the $U(1)$ symmetry ($\alpha(x)$) is not the electric charge Q as in our previous example, but is instead the "Weak Hypercharge" $Y = 2(Q - I_3)$, where I_3 is the third component of the "Weak Isospin" I . For right-handed particles, $I_3 = 0$, for left-handed up/charm/top quarks and neutrinos, $I_3 = +\frac{1}{2}$, and for left-handed down/strange/bottom quarks and charged leptons, $I_3 = +\frac{1}{2}$. We will call our $U(1)$ boson B_μ rather than A_μ , to distinguish it from the photon. For a fermion ψ (u/c/s/e/ μ/τ) that couples weakly to another fermion ψ' , we can write the left-and-right-handed components of f as

$$L = \begin{pmatrix} \psi_L \\ \psi'_L \end{pmatrix}, R = \begin{pmatrix} \psi_R \end{pmatrix}.$$

Similarly, the generators of the $SU(2)$ symmetry are the three matrices τ_i , one corresponding to each weak boson W_i^μ . The self-interaction term for the weak interaction is $L_{kin} = -\frac{1}{4}G^{\mu\nu}G_{\mu\nu}$, where $G_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu + \frac{g}{2}W_\mu \times W_\nu$.

We can define a new covariant derivative capturing the weak interaction terms as well: $D^\mu = \partial^\mu - \frac{1}{2}iYg_1B^\mu - \frac{1}{2}iIg_2\Sigma\tau_iW_i^\mu$. Thus, the full electroweak $SU(2)_L \times U(1)$ Lagrangian is:

$$\mathcal{L} = i\bar{R}\gamma^\mu D_\mu R + i\bar{L}\gamma^\mu D_\mu L - \frac{1}{4}G^{\mu\nu}G_{\mu\nu} - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \quad (1.8)$$

However, we note that this Lagrangian does not allow for gauge boson mass or fermion mass, as a fermion mass term $m\bar{\psi}\psi = m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R)$ is not an $SU(2)$ singlet and would thus not be invariant under the $SU(2)$ symmetry.

To rectify this, we can introduce a new complex SU(2) doublet scalar field $\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}$ called the *Higgs field*. It follows the Lagrangian

$$\mathcal{L} = T - V = (D^\mu \phi)^\dagger (D_\mu \phi) - \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2 \quad (1.9)$$

which, if $mu^2 < 0$ has a minimum at

$$\phi^\dagger \phi = \frac{-\mu^2}{2\lambda} = \frac{v^2}{2} \quad (1.10)$$

We note that the potential is symmetric, so we can pick a direction to "break the symmetry" in. Let us define the vacuum as $\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}$. Then, for any arbitrary configuration of the field, we can expand the potential about the minimum as $\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ (v + h(x)) \end{pmatrix}$.

We can expand the kinetic term in the Lagrangian as

$$(D^\mu \phi)^\dagger (D_\mu \phi) = \frac{1}{2} (\partial_\mu h)(\partial^\mu h) + \phi^\dagger \left(\frac{1}{2} i Y g_1 B_\mu - \frac{1}{2} i I g_2 \sum \tau_i W_{i\mu} \right)^\dagger \left(\frac{1}{2} i Y g_1 B^\mu - \frac{1}{2} i I g_2 \sum \tau_i W_i^m u \right) \phi \quad (1.11)$$

Which, using the definitions of the three tau matrices, and putting $Y = I = 1$, becomes

$$\begin{aligned} & \frac{1}{2} (\partial_\mu h)(\partial^\mu h) + \frac{1}{8} \left| \begin{bmatrix} g_1 B_\mu + g_2 W_\mu^3 & g_2 (W_\mu^1 - i W_\mu^2) \\ g_2 (W_\mu^1 + i W_\mu^2) & g_1 B_\mu + g_2 W_\mu^3 \end{bmatrix} \begin{bmatrix} 0 \\ v \end{bmatrix} \right|^2 \\ &= \frac{1}{2} (\partial_\mu h)(\partial^\mu h) + \frac{1}{8} v^2 g_2^2 ((W_\mu^1)^2 + (W_\mu^2)^2) + \frac{1}{8} v^2 (g_1 B_\mu - g_2 W_\mu^3)^2 \end{aligned} \quad (1.12)$$

If we again define some new field combinations here: $W_\mu^+ = \frac{1}{\sqrt{2}} (W_\mu^1 - i W_\mu^2)$, $W_\mu^- = \frac{1}{\sqrt{2}} (W_\mu^1 + i W_\mu^2)$, $Z_\mu = \frac{g_1 B_\mu + g_2 W_\mu}{\sqrt{g_1^2 + g_2^2}}$ and $A_\mu = \frac{g_1 B_\mu - g_2 W_\mu}{\sqrt{g_1^2 + g_2^2}}$, then we get:

$$\frac{1}{2} (\partial_\mu h)(\partial^\mu h) + \left(\frac{1}{2} v g_2 \right)^2 W_\mu^+ W_\mu^- + \left(\frac{1}{2} v (\sqrt{g_1^2 + g_2^2}) \right)^2 Z_\mu \quad (1.13)$$

And we see that Z field and the two W fields have acquired mass terms, while the A field has remained massless. This corresponds to what we observe in the real world: the W bosons and the Z boson are massive, while the photon remains massless!

For fermions, the coupling to the Higgs field looks like $g_f (\bar{f}_L \phi f_R + \phi^\dagger \bar{f}_R f_L)$ (for some constant g_f).

We note that, if we expand $\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ (v + h(x)) \end{pmatrix}$ as before, we find that the coupling term now looks like $\frac{g_f v}{\sqrt{2}} (\bar{f}_L f_R + \bar{f}_R f_L) + \frac{g_f}{\sqrt{2}} (\bar{f}_L f_R + \bar{f}_R f_L)$. Thus, the presence of the Higgs field allows us to have fermion mass terms in our Lagrangians, where the masses are $m_f = \frac{g_f v}{\sqrt{2}}$. However, we note that, at the current time, there is no way to derive the g_f coupling theoretically: thus, while the Higgs mechanism now allows us to have massive fermions in our theory, it does not actually determine what the masses of these fermions are.

Finally, we note that if we expand the potential term in the complex Klein-Gordon Lagrangian for the Higgs field about the minimum, that is,

$$\begin{aligned} V &= \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2 = \frac{1}{2} \mu^2 (v + h)^2 + \frac{1}{4} \lambda (v + h)^4 \\ &\quad = -\frac{1}{2} \lambda v^2 (v + h)^2 + \frac{1}{4} \lambda (v + h)^4 \\ &= -\frac{1}{2} \lambda (v^4 + 2v^3 h + v^2 h^2) + \frac{1}{4} \lambda (v^4 + 4vh^3 + 6v^2 h^2 + 4v^3 h + h^4) \end{aligned} \tag{1.14}$$

we see a mass term of the form $(\frac{M_h^2}{2})h^2 = \frac{6}{4} \lambda v^2 h^2 - \frac{1}{2} \lambda v^2 h^2 = \lambda v^2 h^2$; That is, $M_h = v\sqrt{2\lambda}$.

In sum: without a Higgs field, in order for the Standard Model electroweak Lagrangian to preserve $SU(2)_L \times U(1)$ symmetry, we are forbidden from including mass terms for the fermions or the gauge bosons. However, by adding this additional field to the model and breaking its symmetry, we find that we can include fermion mass terms that do not violate the $SU(2)_L \times U(1)$ symmetry, and that the electroweak gauge bosons now acquire mass terms by default. Furthermore, this new field also has a mass. This is a remarkable result: by simply adding one extra field to the model, we have solved several different problems in one fell swoop!

1.4 The Higgs Boson and Its Couplings

As we have seen, the Standard Model Higgs is massive, and couples directly to all other massive particles (with the possible exception of neutrinos; this is an active area of research). However, the existence of the Higgs is clearly not all there is to the story: the Standard Model is noticeably incomplete (What are dark matter and dark energy? What determines the exact masses of each of the fermions? Why have we never observed right-handed neutrinos? Can we unify other forces like we have unified electromagnetism and the weak force?). In the majority of proposed Beyond-the-Standard Model extensions, other particles that would interact with the Standard Model Higgs are introduced, including additional Higgs fields, as well as new massive fermions and bosons.

With the discovery of the Higgs, it can be argued that we have now entered an era of precision-measurement in collider physics. By closely studying interactions that are very sensitive to the

injection of new physics, we can potentially detect deviations from the Standard Model: in many cases, we may see hints of the existence of new particles through changes to quantities of interest well before we are actually able to produce these particles in colliders directly. Due to the Higgs coupling to all massive particles, Higgs interactions are especially sensitive to new physics, and thus can serve as a useful bellwether for physics beyond the standard model.

The Large Hadron Collider is a proton-proton collider; as such, physics processes that occur in it are almost always initiated by quarks or gluons. The predominant production modes are, in order of frequency, gluon-gluon fusion (ggF), vector boson fusion (VBF), vector-boson associated (VH) and top-associated (tH). The top-associated Higgs production mode is of particular interest to many experimentalists due to its direct dependence on the Higgs coupling to a fermion—by closely measuring the frequency and properties of the tH mode, we can better understand the nature of this variety of couplings. Rarer production modes, such as single-top associated Higgs production (tH), can shed still more light on the nature of the top-quark/Higgs coupling (called the top quark Yukawa coupling), but are much more difficult to observe. The tH process occurs in two final-state modes, tWH and tHj ; however, there is some nuance in their definitions which we will come to shortly.

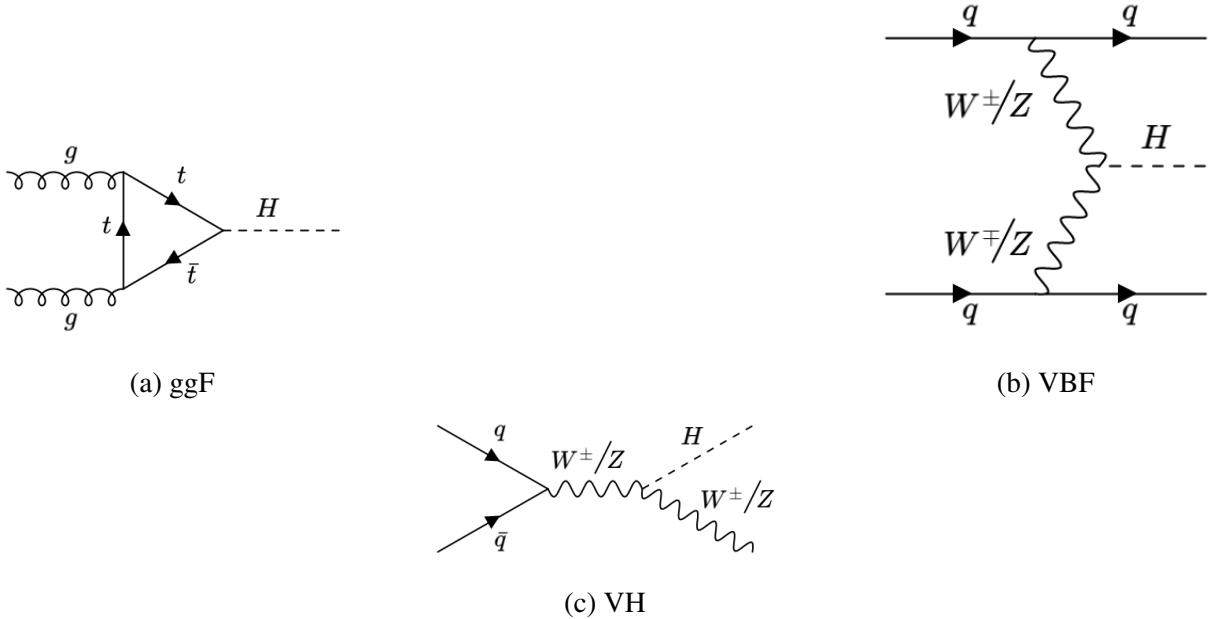
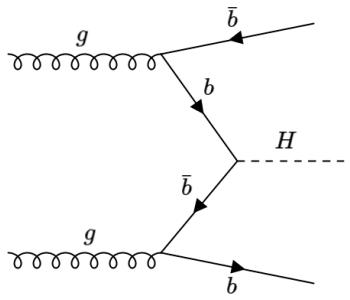
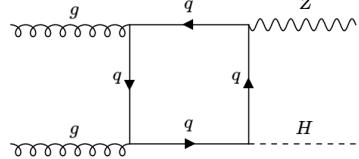


Figure 1.3: Feynman diagrams depicting the three leading Higgs production modes.

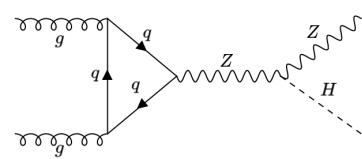
Similarly, the dominant decay mode of the Higgs is to bottom quarks ($H \rightarrow bb$), followed by to W bosons ($H \rightarrow WW$), gluons ($H \rightarrow WW$), tau leptons ($H \rightarrow \tau\tau$), charm quarks ($H \rightarrow bb$), Z bosons ($H \rightarrow ZZ$), photons ($H \rightarrow \gamma\gamma$), and a Z boson and a photon ($H \rightarrow Z\gamma$). Because the Higgs does not couple to massless particles directly, the decays to gg and $\gamma\gamma$ are mediated by loop



(a) bbH



(b) $gg \rightarrow ZH$



(c) Additional $gg \rightarrow ZH$

Figure 1.4: Feynman diagrams depicting relevant less-common Higgs production modes.

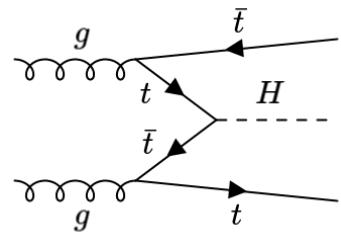
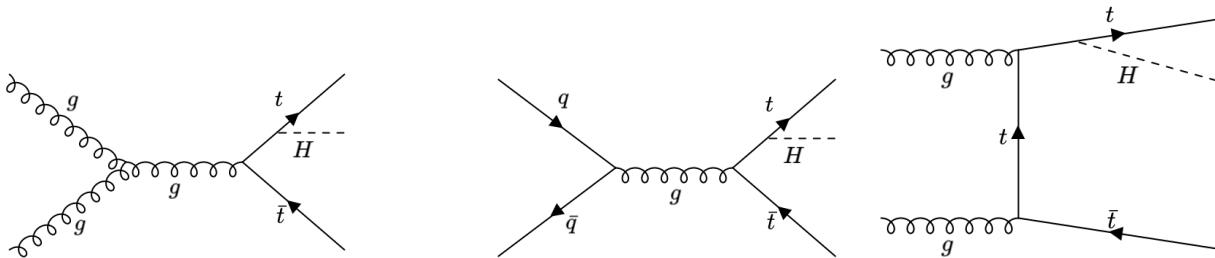
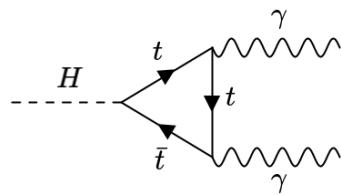
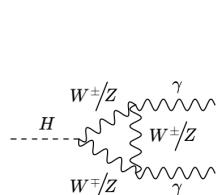


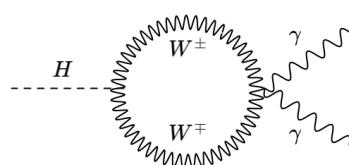
Figure 1.5: Feynman diagrams depicting ttH production modes.



(a) Top-mediated



(b) $W \rightarrow \gamma$



(c) $W \rightarrow \gamma\gamma$

Figure 1.6: Feynman diagrams depicting the leading-order processes contributing to the Higgs diphoton decay.

diagrams, most often involving top quarks.

Investigating each of these decay modes has different benefits and detriments. While the Higgs decay to bottom quarks is the most common, correctly identifying and accurately reconstructing quarks and gluons using the jets that they produce in detectors is often very difficult. Decays to W bosons, Z bosons, and τ leptons provide "cleaner" channels, but because the W, Z and τ decay dominantly to hadrons, similar reconstruction issues occur unless we restrict ourselves to the rarer leptonic subchannels of these decays. The Higgs to diphoton channel occurs very rarely, but offers a much more unambiguous signal, as the odds of misidentifying high-energy gamma rays are fairly low.

Additionally, we must consider combinatorics: in the ttH production mode, for instance, the final state contains two bottom quarks from the decay of the associated top quarks. If we choose to examine the subchannel in which the Higgs decays to bottom quarks as well, we find that we have at least four final-state bottom quark jets in the event, with 12 different unique assignments of bottom quark jets to parent particles. Correctly matching up which jet originated from which parent particle thus further complicates the reconstruction problem.

The rate of a particular physics production mode or scattering process is parameterized using a quantity called the cross-section σ , which is measured in units of "barns". Similarly, the rate of a particular decay process is parameterized using a quantity called the decay width Γ , which measures the probability of a particular decay occurring per unit time. We can also define the "Branching Fraction" of a particular process: if Γ_H denotes the sum of all Higgs decay widths, for a given process $H \rightarrow XX$ we define

$$Br(H \rightarrow XX) = \frac{\Gamma(H \rightarrow XX)}{\Gamma_H} \quad (1.15)$$

Decay Mode	Branching fraction [%]
$H \rightarrow bb$	57.5 ± 1.9
$H \rightarrow WW$	21.6 ± 0.9
$H \rightarrow gg$	8.56 ± 0.86
$H \rightarrow \tau\tau$	6.30 ± 0.36
$H \rightarrow cc$	2.90 ± 0.35
$H \rightarrow ZZ$	2.67 ± 0.11
$H \rightarrow \gamma\gamma$	0.228 ± 0.011
$H \rightarrow Z\gamma$	0.155 ± 0.014
$H \rightarrow \mu\mu$	0.022 ± 0.001

Table 1.1: Higgs decay modes and branching fractions, for a Standard Model Higgs with mass of 125.09 GeV ??

Cross-sections are calculated at a given perturbation-theoretic "order", indicating how many

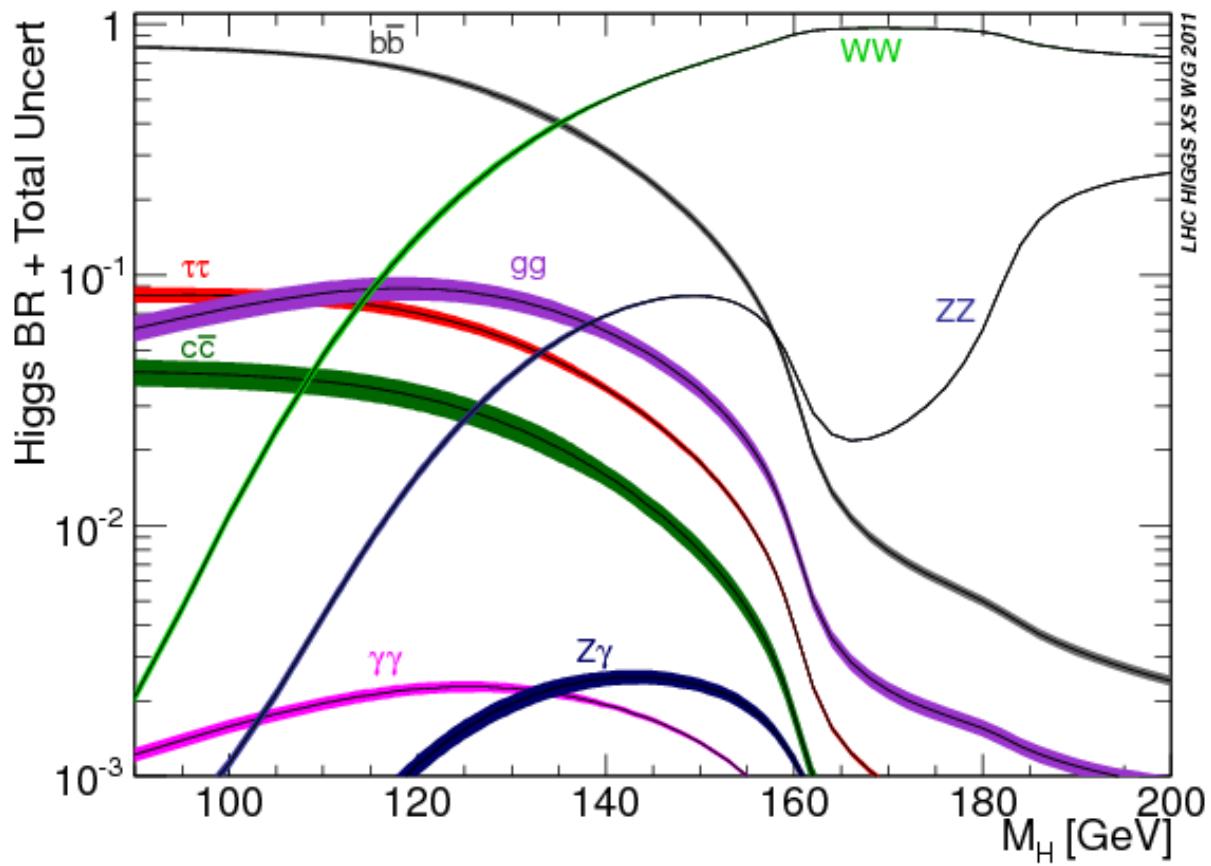


Figure 1.7: The branching ratio of the Higgs to various final state particles as a function of its mass (now known to be 125 GeV), from reference ??.

additional correction terms (often indicated by Feynman diagrams with internal loops and vertices) are accounted for in the calculation- for a given process, leading-order (or "LO"), next-to-leading order ("NLO"), and next-to-next-to-leading order ("NNLO") indicate increasing numbers of correction terms. Calculating and simulating to higher orders is more accurate, but is often much more computationally expensive. These correction terms can be applied for both electroweak (EW) and quantum-chromodynamical (QCD) effects.

In addition to considering which corrections to apply, we must also consider the flavor scheme of the Monte Carlo generator. Physics processes can be initiated by real quarks or gluons that are present in the initial scattering protons, as well as virtual quarks that are present in the primordial soup of energy that binds each proton together. In the five-flavor Parton Distribution Function (PDF) scheme, we consider massless virtual bottom quarks to be constituents of the proton that are capable of initiating physics processes, while in the four-flavor PDF scheme, we do not. These schemes differ primarily in which diagrams they treat as corrections and which they do not- thus, if we perform QCD corrections at all orders (that is, NN...NNLO), we find that these two schemes produce identical results ???. Diagrams from the five-flavor PDF scheme that are initiated by a bottom quark can be mapped into the four-flavor scheme by adding an additional bottom-quark to the final state; however, for $t\bar{t}H$, $t\bar{t}\gamma\gamma$ and $tW\bar{t}H$ processes, this presents a number of nontrivial modelling issues ???. While the five-flavor scheme offers easier calculation and the ability to model more processes, the four-flavor scheme is better at modelling Higgs kinematics, such as p_T and bottom-quark p_T . For these reasons, we use the four-flavor scheme to model the tHj process only (henceforth called $tHjb$ due to the presence of the additional final-state b-jet) and the five-flavor scheme for all other top-quark-based processes.

1.4.1 The Kappa Framework, STXS, and EFTs

This dissertation features two physics analyses from the ATLAS experiment. The first is a targeted attempt to measure whether the top quark Yukawa coupling is CP-even (respects CP symmetry) or CP-odd (violates it) by closely studying the $t\bar{t}H$ process. In the Standard Model Lagrangian, the top Yukawa coupling is purely CP-even, and can be written in the form $\kappa g_t \bar{t}t h$. Adding a CP-odd component to the coupling adds an additional $\kappa g_t \bar{t}(i\gamma^5)th$ term to the Lagrangian; the total coupling can thus be parameterized as $\kappa g_t \bar{t}(\cos(\alpha) + \sin(\alpha)i\gamma^5)th$ for a "CP-mixing angle" α that governs the relative magnitude of the CP-even and CP-odd terms. Introducing this CP-odd term changes the kinematics of the process, and leads to events with different angular variable values (such as the angle between the two tops) as well as changing the energies and momenta of some decay products. By performing a fit on these variables to determine the relative agreement or disagreement with the standard model, we can extract a value for α . In this dissertation, we refer

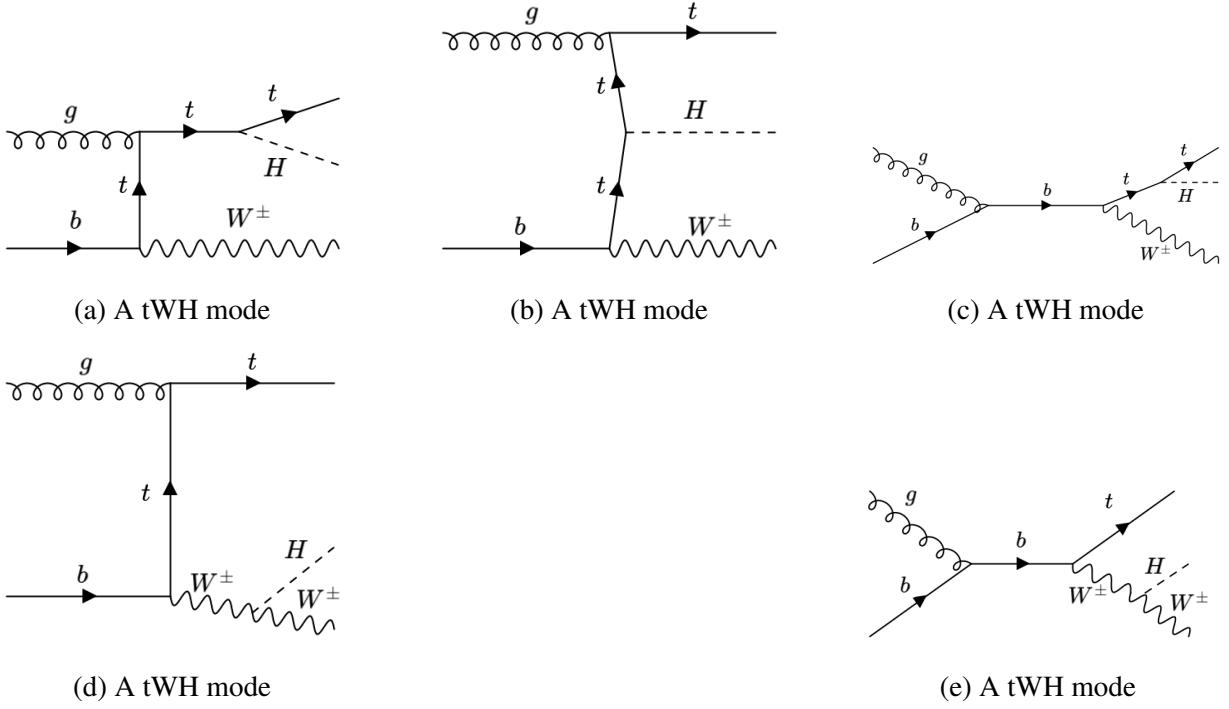


Figure 1.8: Feynman diagrams depicting the leading-order terms for tWH . Because all diagrams contain initial b -quarks, all of these processes can only occur in the five-flavor PDF scheme.

to this analysis as the "ttHCP Analysis".

The second analysis discussed here is a survey of a wide variety of Higgs-based physics processes using the Higgs-to-diphoton decay channel (referred to in this dissertation as the "Higgs Couplings Analysis"). A combined fit is performed in different regions of phase space that correspond to the ggF, VBF, VH, ttH, and tH processes; by doing this, we can extract limits on the individual Higgs couplings κ .

When measuring Higgs couplings in analyses such as these, it is often most useful to note the ratio of the measured coupling to the Standard Model prediction, rather than the measured coupling itself. For some Higgs coupling to a boson or fermion i , we can write

$$\kappa_i = \frac{y_i^{meas}}{y_i^{SM}} \quad (1.16)$$

This κ parameter is called the "coupling strength"—if $\kappa_{\text{particle } i} = 1$, the Higgs couples to particle i exactly as predicted by the Standard Model, while if $\kappa_{\text{particle } i} = 0$, the Higgs does not couple to particle i at all. Measuring the Higgs couplings and interactions in this manner is referred to as the "Kappa Framework" ??.

The parameterization of various production cross-sections and decay widths (divided by their standard model values) in terms of the κ framework is shown in table ?? ???. The "resolved"

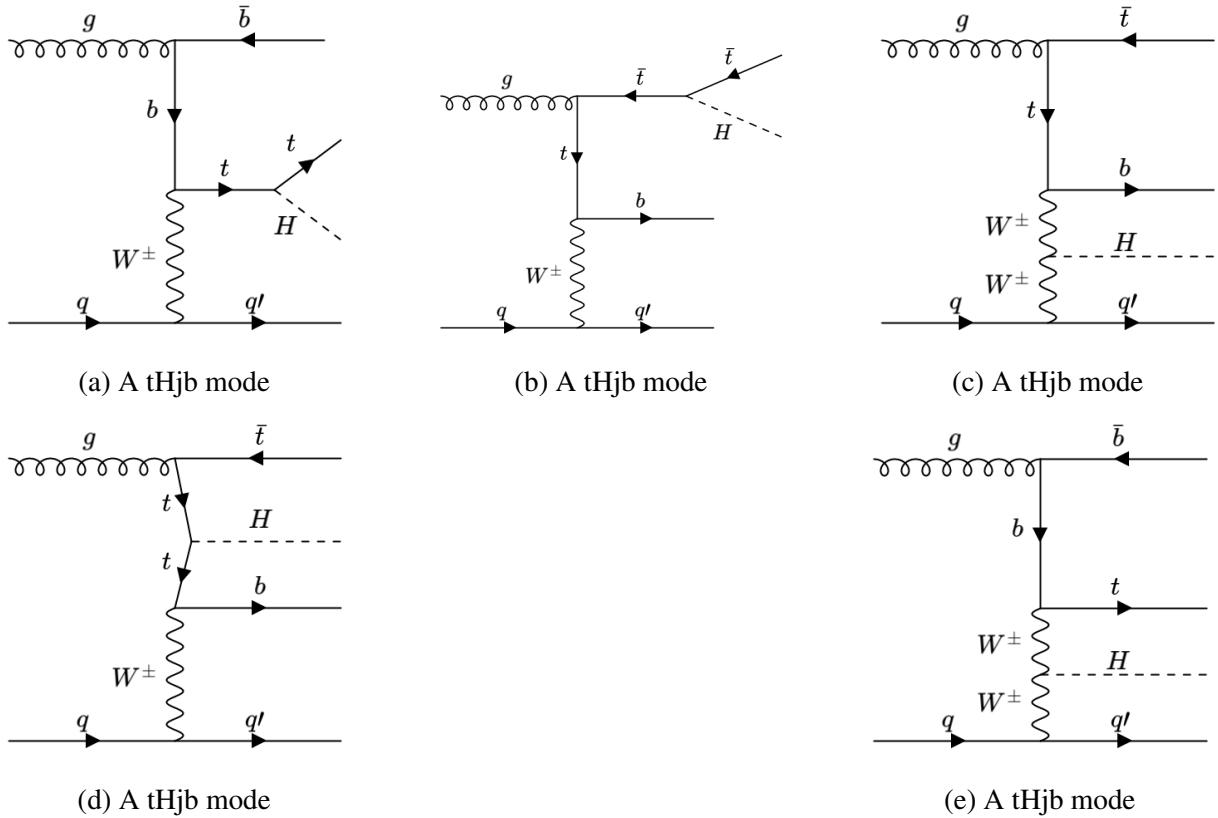


Figure 1.9: Feynman diagrams depicting the leading-order terms for $tHjb$. These diagrams are calculated using the four-flavor PDF scheme. Note that additional diagrams can be created by reversing the direction of the upper fermion "circuit" (the final-state top and bottom must be opposite sign, but $tHj\bar{b}$ and $\bar{t}Hjb$ are equally likely to occur).

column assumes that only the particles of the Standard Model exist. However, it can also be useful to develop "effective" coupling terms that allows for the existence of additional particles that may influence the loop diagrams. These can be modeled by multiplying the resolved coupling terms by the indicated factors:

$$\kappa_g^2 = \frac{\sigma_{ggF}^{meas}}{\sigma_{ggF}^{SM}}, \kappa_\gamma^2 = \frac{\Gamma_{\gamma\gamma}^{meas}}{\Gamma_{\gamma\gamma}^{SM}}, \kappa_{Z\gamma}^2 = \frac{\Gamma_{Z\gamma}^{meas}}{\Gamma_{Z\gamma}^{SM}} \quad (1.17)$$

Production Mode	Effective Modifier	Resolved Modifier
$\sigma(ggF)$	κ_g^2	$1.04\kappa_t^2 + 0.002\kappa_b^2 - 0.04\kappa_t\kappa_b$
$\sigma(VBF)$		$0.73\kappa_W^2 + 0.27\kappa_Z^2$
$\sigma(qq/qg \rightarrow WH)$		κ_W^2
$\sigma(qq/qg \rightarrow ZH)$		κ_Z^2
$\sigma(ggZH)$		$2.46\kappa_Z^2 + 0.46\kappa_t^2 - 1.90\kappa_t\kappa_Z$
$\sigma(t\bar{t}H)$		κ_t^2
$\sigma(b\bar{b}H)$		κ_b^2
$\sigma(tHjb)$		$2.63\kappa_t^2 + 3.58\kappa_W^2 - 5.21\kappa_t\kappa_W$
$\sigma(tWH)$		$2.91\kappa_t^2 + 2.31\kappa_W^2 - 4.22\kappa_t\kappa_W$

Table 1.2: Parameterization of Higgs cross-section dependence on κ coefficients, from ??

Production Mode	Effective Modifier	Resolved Modifier
$H \rightarrow bb$		κ_b^2
$H \rightarrow WW$		κ_W^2
$H \rightarrow gg$	κ_g^2	$1.11\kappa_t^2 + 0.01\kappa_b^2 - 0.12\kappa_t\kappa_b$
$H \rightarrow \tau\tau$		κ_τ^2
$H \rightarrow cc$		$2.46\kappa_c^2$
$H \rightarrow ZZ$		κ_Z^2
$H \rightarrow \gamma\gamma$	κ_γ^2	$1.59\kappa_W^2 + 0.07\kappa_t^2 - 0.67\kappa_t\kappa_W$
$H \rightarrow Z\gamma$	$\kappa_{z\gamma}^2$	$1.12\kappa_W^2 + 0.12\kappa_t\kappa_W$
$H \rightarrow \mu\mu$		κ_μ^2

Table 1.3: Parameterization of Higgs branching ratio dependence on κ coefficients, from ??

The Couplings Analysis is performed within the Simplified Template Cross Sections (STXS 1.2) framework, where different regions of phase space, called "truth bins", are probed according to their production modes and kinematic properties. A visual guide to the different STXS1.2 bins is shown in figure ?? . The categorization is designed such that all events passing the $H \rightarrow \gamma\gamma$ selection criteria will fall into one of the STXS1.2 categories.

In addition to their nominal production and decay modes, the ggH STXS1.2 truth bins include both ggF and hadronically-decaying $gg \rightarrow ZH$ events, while the VBF STXS definition includes hadronically-decaying $qq \rightarrow VH$. Additionally, it is important to note that the leptonic $qq \rightarrow ZH$

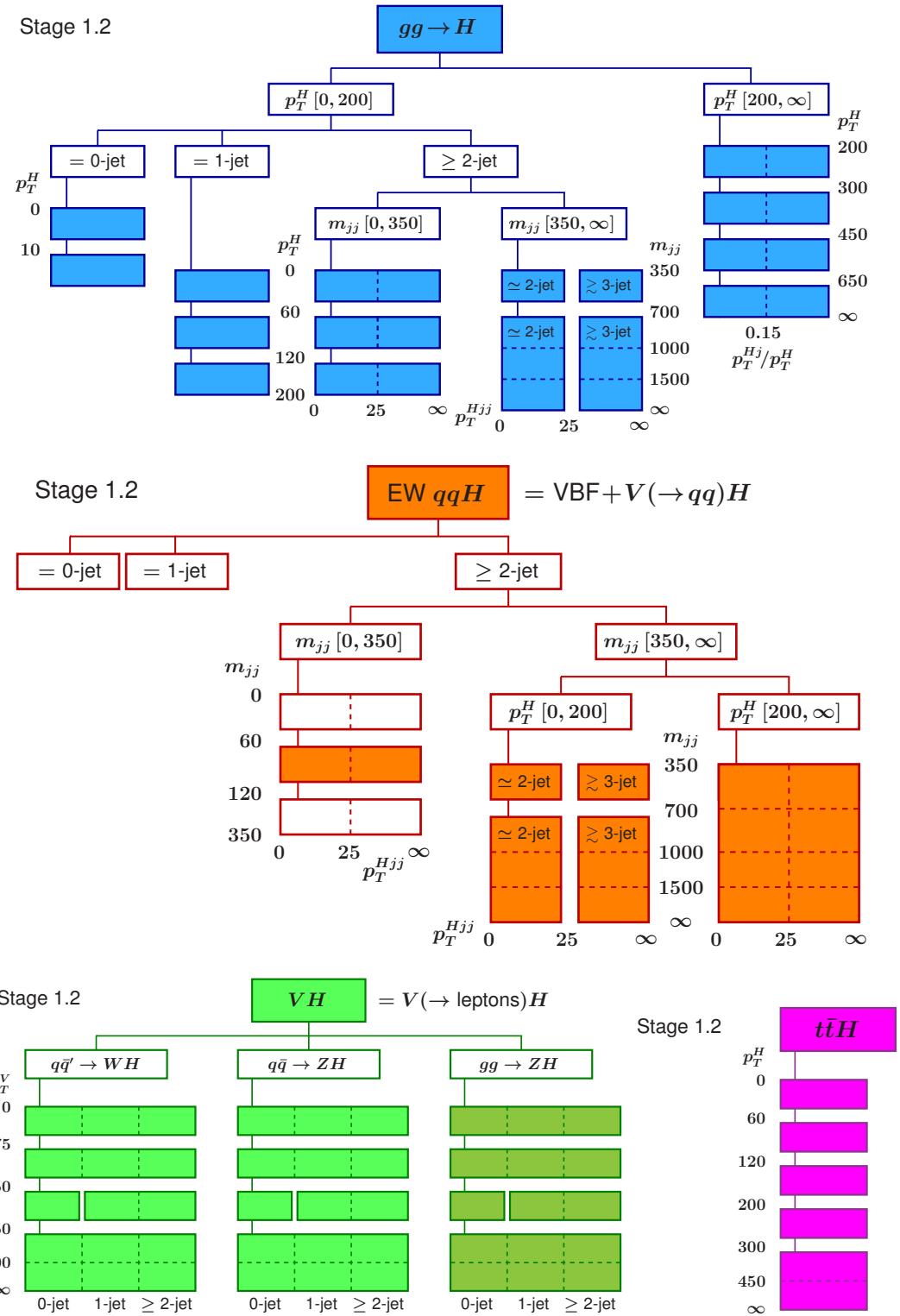


Figure 1.10: Stage 1.2 STXS bin definitions for the main production modes.

and $gg \rightarrow ZH$ truth bins include Z-boson decays both to two charged leptons and to two neutrinos (the latter of which will have a high MET signature).

In the Couplings analysis, the experimental bins targeted are modified somewhat from the truth bins due to the inability to separate some processes: the $b\bar{b}H$ categories are subsumed into the ggH categorization scheme, the leptonic $qq \rightarrow VH$ and $gg \rightarrow VH$ processes are combined into a series of $VH Lep$ regions, tH is split into tWH and $tHjb$, and the $p_T^H > 200 GeV$ bin of the hadronic $qq \rightarrow VH$ truth category is split into two bins based on jet kinematics to aid in statistical combination with other analyses.

Furthermore, truth bins of further production modes like $b\bar{b}H$ and tH are not split into dedicated kinematic regions because of their low production cross sections; in particular, in the Couplings analysis, the $b\bar{b}H$ categories are subsumed into the ggH categorization scheme, while the tH categories are subsumed into the tH categorization scheme.

For all Monte Carlo simulated samples used to construct the STXS truth bins, the absolute value of the rapidity of the Higgs boson is required to satisfy $|y_H| < 2.5$ while jets, built with the anti-kT algorithm using radius $R=0.4$, are required to have $p_T > 30 GeV$ (with the exception of the bins labelled FWDH, which target forward Higgs events with $|y_H| > 2.5$).

Using the Monte Carlo simulated samples described in section ??, we report the cross-section times branching ratio and acceptances (in terms of percentage of events of each production-mode) that fall into each STXS1.2 truth bin in Table ??, and show the acceptance for each production mode in Figure . Several of the STXS truth bins are merged in the Couplings analysis to account for effects such as the inability of the event classifier to discriminate between $qq \rightarrow ZH$ and $gg \rightarrow ZH$; this is detailed in section ??.

In addition to the kappa-framework, the Couplings Analysis also makes use of an Effective Field Theory (EFT) approach to parameterize potential Higgs Couplings beyond the Standard Model. Effective Field Theories offer a model-independent way to parameterize new physics that may not be directly accessible at LHC energies. Using the Standard Model Effective Field Theory parameterization (SMEFT) ??, we can write the Lagrangian for new physics that exists at some energy scale $\Lambda >> v$ (where v denotes the Higgs field vacuum expectation value of 246 GeV) as

$$\mathcal{L}_{SMEFT} = \mathcal{L}_{SM} + \left(\frac{v}{\Lambda}\right)\mathcal{L}_5 + \left(\frac{v}{\Lambda}\right)^2\mathcal{L}_6 \quad (1.18)$$

If we assume that lepton and baryon number are conserved, we find that we cannot have any \mathcal{L}_5 operators, as all allowable dimension-5 operators violate the conservation of lepton number; thus, the leading EFT terms are all proportional to $(\frac{v}{\Lambda})^2$ and have dimension-6.

The dimension-6 Lagrangian term can be expanded as a linear combination of all allowable dimension-6 operators:

STXS 1.2 Truth Bin	$\sigma \times BR_{\gamma\gamma} [\text{fb}]$
gg2H_0J_ptH_0_10	15.038 147
gg2H_0J_ptH_gt10	46.799 419
gg2H_1J_ptH_0_60	14.742 653
gg2H_1J_ptH_60_120	10.223 149
gg2H_1J_ptH_120_200	1.692 748
gg2H_ge2J_MJJ_0_350_ptH_0_60	2.639 954
gg2H_ge2J_MJJ_0_350_ptH_60_120	4.058 768
gg2H_ge2J_MJJ_0_350_ptH_120_200	2.133 361
gg2H_ge2J_MJJ_350_700_ptH_0_200_ptHJJ_0_25	0.569 488
gg2H_ge2J_MJJ_350_700_ptH_0_200_ptHJJ_gt25	0.818 394
gg2H_ge2J_MJJ_gt700_ptH_0_200_ptHJJ_0_25	0.241 653
gg2H_ge2J_MJJ_gt700_ptH_0_200_ptHJJ_gt25	0.363 737
gg2H_ptH_200_300	1.036 991
gg2H_ptH_300_450	0.239 097
gg2H_ptH_450_650	0.035 624
gg2H_ptH_gt650	0.004 947
qq2Hqq_0J	0.816 245
qq2Hqq_1J	3.956 656
qq2Hqq_ge2J_MJJ_0_60	0.224 649
qq2Hqq_ge2J_MJJ_60_120	1.200 933
qq2Hqq_ge2J_MJJ_120_350	1.426 903
qq2Hqq_ge2J_MJJ_350_700_ptH_0_200_ptHJJ_0_25	0.897 833
qq2Hqq_ge2J_MJJ_350_700_ptH_0_200_ptHJJ_gt25	0.337 811
qq2Hqq_ge2J_MJJ_gt700_ptH_0_200_ptHJJ_0_25	1.356 668
qq2Hqq_ge2J_MJJ_gt700_ptH_0_200_ptHJJ_gt25	0.314 044
qq2Hqq_ge2J_MJJ_350_700_ptH_gt200	0.104 740
qq2Hqq_ge2J_MJJ_gt700_ptH_gt200	0.258 838
qq2Hlnu_ptV_0_75	0.466 795
qq2Hlnu_ptV_75_150	0.297 442
qq2Hlnu_ptV_150_250_0J	0.051 545
qq2Hlnu_ptV_150_250_ge1J	0.042 651
qq2Hlnu_ptV_gt250	0.030 152
qq2Hll_ptV_0_75	0.236 572
qq2Hll_ptV_75_150	0.157 167
qq2Hll_ptV_150_250_0J	0.027 476
qq2Hll_ptV_150_250_ge1J	0.022 523
qq2Hll_ptV_gt250	0.015 076
gg2Hll_ptV_0_75	0.014 989
gg2Hll_ptV_75_150	0.038 219
gg2Hll_ptV_150_250_0J	0.007 812
gg2Hll_ptV_150_250_ge1J	0.014 855
gg2Hll_ptV_gt250	0.004 465
ttH_ptH_0_60	0.267 960
ttH_ptH_60_120	0.404 152
ttH_ptH_120_200	0.287 483
ttH_ptH_200_300	0.119 454
ttH_ptH_gt300	0.055 617
bbH	1.044 426
tH	0.187 064

Table 1.4: Simplified template cross sections times diphoton branching ratio for each of the STXS 1.2 truth bins.

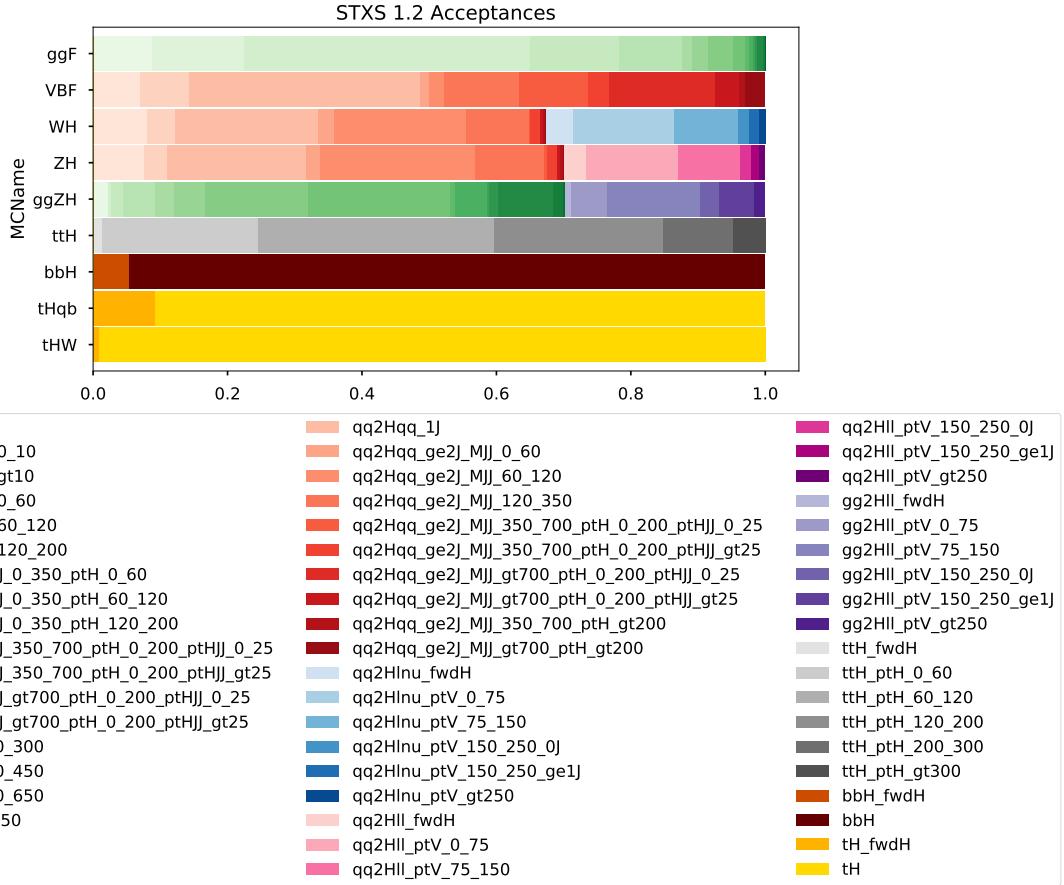


Figure 1.11: Stage 1.2 STXS bin acceptances for all Higgs production modes considered in the Couplings analysis. The FWDH bins target events outside the nominal acceptance, i.e. with $|y_H| > 2.5$.

Wilson coefficient	Operator
c_{Hbox}	$(H^\dagger H) \square (H^\dagger H)$
c_{HDD}	$(H^\dagger D^\mu H)^* (H^\dagger D_\mu H)$
c_{HG}	$H^\dagger H G_{\mu\nu}^A G^{A\mu\nu}$
c_{HB}	$H^\dagger H B_{\mu\nu} B^{\mu\nu}$
c_{HW}	$H^\dagger H W_{\mu\nu}^I W^{I\mu\nu}$
c_{HWB}	$H^\dagger \tau^I H W_{\mu\nu}^I B^{\mu\nu}$
c_{Hl1}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{l}_p \gamma^\mu l_r)$
c_{Hl3}	$(H^\dagger i \overleftrightarrow{D}_\mu^I H)(\bar{l}_p \tau^I \gamma^\mu l_r)$
c_{He}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{e}_p \gamma^\mu e_r)$
c_{Hq1}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{q}_p \gamma^\mu q_r)$
c_{Hq3}	$(H^\dagger i \overleftrightarrow{D}_\mu^I H)(\bar{q}_p \tau^I \gamma^\mu q_r)$
c_{Hu}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{u}_p \gamma^\mu u_r)$
c_{Hd}	$(H^\dagger i \overleftrightarrow{D}_\mu H)(\bar{d}_p \gamma^\mu d_r)$
$ c_{uG} $	$(\bar{q}_p \sigma^{\mu\nu} T^A u_r) \tilde{H} G_{\mu\nu}^A$
c_{ll1}	$(\bar{l}_p \gamma_\mu l_r)(\bar{l}_s \gamma^\mu l_t)$

Table 1.5: Wilson coefficients c_i and corresponding dimension-6 SMEFT operators \mathcal{O}_i used in this analysis.

$$\mathcal{L}_6 = \sum_i c_i^6 O_i^6 \quad (1.19)$$

In the Couplings Analysis, the operators are chosen using the so-called "Warsaw Basis" ??; the c_i terms are the free parameters to be determined by the final likelihood fit, called "Wilson Coefficients". A list of all allowed dimension-6 operators and their corresponding coefficients is given in ??.

Under this model, we can write the cross-section in some STXS bin as

$$\sigma_{bin} = \sigma_{bin}^{SM} + \sigma_{bin}^{interaction} + \sigma_{bin}^{BSM} \quad (1.20)$$

where the first term refers to the Standard-Model cross-section, the second refers to the interaction between the SM and BSM physics processes, and the last refers to the BSM processes themselves.

We can also write:

$$\frac{\sigma_{bin}^{interaction}}{\sigma_{bin}^{SM}} = \sum_i A_i c_i \frac{\sigma_{bin}^{BSM}}{\sigma_{bin}^{SM}} = \sum_i B_{ij} c_i c_j \quad (1.21)$$

So we can parameterize the cross-section as:

$$\frac{\sigma_{bin}}{\sigma_{bin}^{SM}} = 1 + \sum_i A_i c_i + \sum_i B_{ij} c_i c_j \quad (1.22)$$

Similarly, the Higgs partial decay width to diphotons (a measure of how often the Higgs decays to diphotons vs other channels) can be parameterized as follows:

$$\Gamma_{\gamma\gamma} = \Gamma_{\gamma\gamma}^{SM} + \Gamma_{\gamma\gamma}^{interaction} + \Gamma_{\gamma\gamma}^{BSM} \Gamma_{\gamma\gamma} = 1 + \sum_i A_i^{\Gamma_{\gamma\gamma}} c_i + \sum_i B_{ij}^{\Gamma_{\gamma\gamma}} c_i c_j \quad (1.23)$$

Combining these, we can parameterize the event in each STXS bin in terms of the Wilson coefficients, using Monte Carlo simulated events to determine the A_i and B_{ij} coefficients. This is detailed further in Section ??.

1.4.2 CP-Violation in the Top Yukawa Coupling

According to the Standard Model, the Higgs boson is a CP-even scalar boson: it is predicted to have spin-zero, with all its interactions CP-even. Previous analyses from both ATLAS and CMS have placed limits on the existence of anomalous CP-violating Higgs couplings to the gauge bosons (?? ?? and ??); however, any such couplings would be suppressed by a factor of $1/(\Lambda)^2$ (where Λ indicates the energy scale of the new CP-violating physics), so these studies are thus somewhat limited in their sensitivity. Similarly, indirect measurements of the CP-nature of the Higgs coupling to fermions in loop-diagram mediated processes have been performed ??; however, these measurements are highly dependent on the choice of BSM model that induces CP violation.

The CP-analysis detailed in this dissertation marks the first direct measurement of the CP-nature of the Higgs couplings to fermions. Because the top quark Yukawa is the strongest Higgs coupling, it is one of the most useful channels for the performance of this measurement. A CP-violating top Yukawa coupling will influence production rates and kinematics in top-pair associated Higgs production ($t\bar{t}H$) and single-top associated Higgs production (tH , specifically $tHjb$ and tWH). Additionally, CP-violation in the top Yukawa coupling will modify the rates of gluon-gluon fusion Higgs production and Higgs to diphoton decay; however, because these two processes are loop-mediated, they are sensitive to other forms of new physics as well, and thus not enough to directly constrain the CP nature of the top Yukawa coupling on their own. Previous ATLAS analyses have measured the CP properties of the top Yukawa coupling in this manner ??, but such constraints are indirect.

Because of its combined direct and indirect sensitivity to the top Yukawa coupling, we thus find that the $t\bar{t}H$ and tH processes with $H \rightarrow \gamma\gamma$ provide a well-motivated channel for probing the CP-structure of the top-Higgs interaction. In addition, the presence of two photons in the event final state provides a clean signal, further motivating the use of this channel.

Using the Higgs Characterization (HC) model ??, we can parameterize the top-Higgs interaction in the presence of CP-violation as $\mathcal{L} = -frac{m_t v \kappa_t}{\sqrt{2}} (\bar{t}(\cos(\alpha) + \sin(\alpha)i\gamma^5)t)$ for mixing angle $\alpha \in (-180^\circ, 180^\circ)$ and $\kappa_t >= 0$.

With the coupling formulated in such a way, the tH cross-section can be parameterized as

$$\sigma_{ttH} = A\kappa_t^2(\cos(\alpha)^2) + B\kappa_t^2(\sin(\alpha)^2) + E\kappa_t^2\cos(\alpha)\sin(\alpha) \quad (1.24)$$

for some constants A, B, and E (with E expected to be zero for tH), while the tWH and $tHjb$ cross-sections can be parameterized as:

$$\sigma_{tH} = A\kappa_t^2(\cos(\alpha)^2) + B\kappa_t^2(\sin(\alpha)^2) + C\kappa_t(\cos(\alpha)) + D\kappa_t(\sin(\alpha)) + E\kappa_t^2\cos(\alpha)\sin(\alpha) + F \quad (1.25)$$

The C and D terms originate from destructive interference: because the tH process contains diagrams where the Higgs couples to the top as well as diagrams where the Higgs couples to the W-boson, holding the W-Higgs coupling fixed while varying the top-Higgs coupling can alter the total cross-section of the process. Likewise, the F term indicates the contribution only from the diagrams that do not involve the top Yukawa coupling ???. Furthermore, we note that the ttH cross-section is the same for $\alpha = 0^{\text{deg}}$ and $\alpha = 180^{\text{deg}}$ while the tH cross-section is not- thus, measuring the tH interaction allows us to eliminate this degeneracy and determine the sign of the top Yukawa coupling.

Similarly, we can parameterize κ_g and κ_γ according to the parameterizations in ??:

$$\kappa_g^2 = 1.11\kappa_t^2\cos(\alpha)^2 + 2.6\kappa_t^2\sin(\alpha)^2 - 0.11\kappa_t\cos(\alpha)\kappa_\gamma^2 = 0.08\kappa_t^2\cos(\alpha)^2 + 0.18\kappa_t^2\sin(\alpha)^2 - 0.72\kappa_t\cos(\alpha) + 1.64 \quad (1.26)$$

Cross-sections as a function of α as implemented in Monte Carlo simulation are given in ???. We note that, as α increases from $\alpha = 0^{\text{deg}}$ (CP-even) to $\alpha = 90^{\text{deg}}$ (CP-odd), the ttH cross-section decreases while the tH and ggF cross-sections increase. Similarly, the $H \rightarrow \gamma\gamma$ branching ratio increases as α increases.

By developing targeted analysis categories and parameterizing the yield in each category as a function of α , we can thus perform a likelihood fit to the data and set limits on the true value of α .

CHAPTER 2

The ATLAS Detector

2.1 The Large Hadron Collider

At roughly 27 kilometers in circumference, the Large Hadron Collider (LHC) at the Center for European Nuclear Research (CERN) is the largest machine ever built, slicing through the solid rock of the Jura mountains and running more than 100 meters beneath the Swiss-French border ???. It is a proton-proton collider operating at a center-of-mass energy of 13 TeV, circulating two beams of protons in opposite directions, each at more than 99% the speed of light ???. By utilizing a system of radiofrequency (RF) cavities and dipole magnets, protons are delivered to four primary collision points in bunches spaced 25 nanoseconds apart ??.

The LHC achieves such a powerful center-of-mass energy by repurposing a significant fraction of older collider physics infrastructure, including the Super Proton Synchrotron (SPS), which was itself one of the most powerful particle accelerators in the world at one time. In order to achieve collision, hydrogen atoms are first stripped of their electrons using an ionizing cathode filament in a device called a duoplasmatron; this produces a plasma that is filtered to produce beams of protons ???. The LINAC2, a linear accelerator, then accelerates the resultant proton beams to a collision energy of 50 MeV; the beams then move to the Proton Synchrotron Booster (PSB), a circular accelerator that accelerates them to an energy of 1.4 GeV, followed by the Proton Synchrotron (PS), a second circular accelerator that accelerates them to 25 GeV. Following this, beams enter the aforementioned Super Proton Synchrotron, where they are accelerated to a collision energy of 450 GeV before being injected into the Large Hadron Collider ring. This infrastructure is depicted in Figure ??.

The tight bunching of protons by the LHC ring allows the collider to deliver collisions at a high luminosity, a measure of the number of expected collision events per unit of beam area. Luminosity is measured in both instantaneous and integrated form (i.e., 'luminosity per unit time' and total luminosity delivered to date'); as of the end of the second major LHC run, the collider

CERN's Accelerator Complex

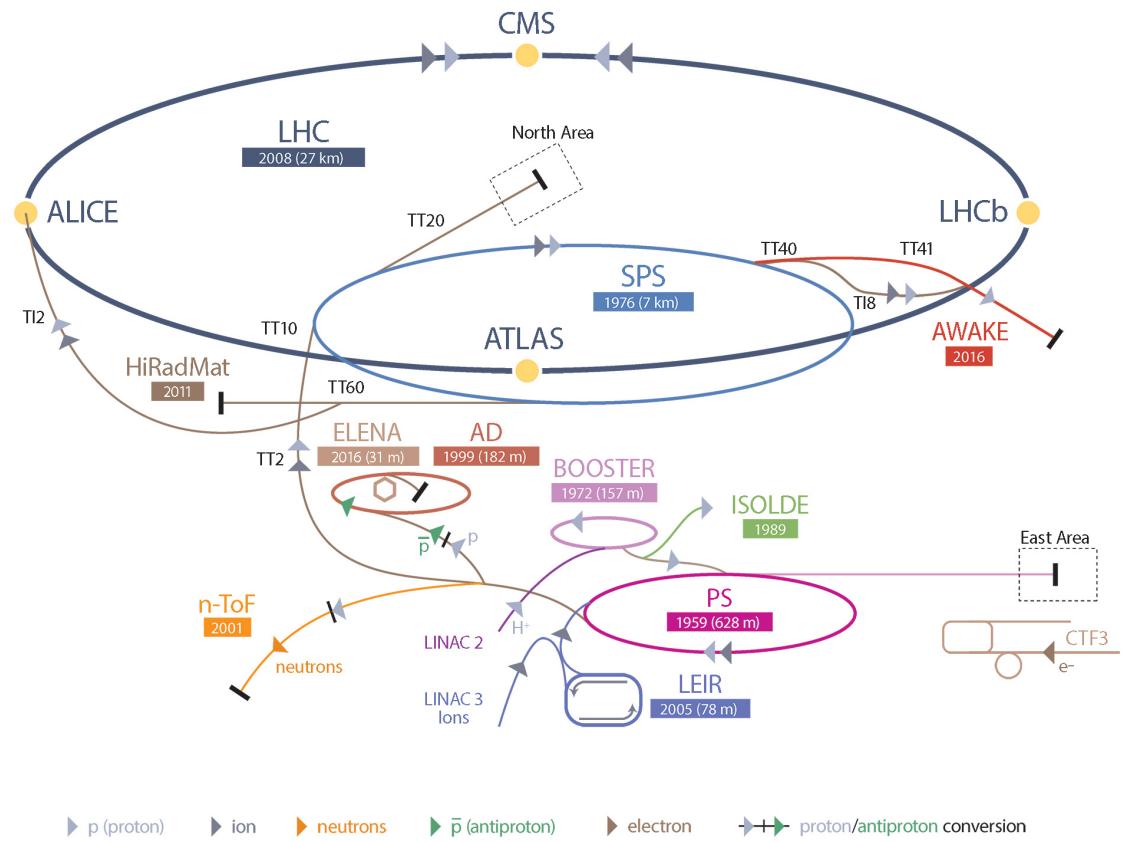


Figure 2.1: The infrastructure of the LHC accelerator ring, including the SPS and LINAC2. ??

has delivered a total integrated luminosity of 139 fb^{-1} , equivalent to approximately $1.39 * 10^{41}$ collisions per square centimeter ??.

At the time of this writing, the LHC is nearing the end of 'Long Shutdown 2' (LS2), a two-year long upgrade period designed to increase detector performance and lay the groundwork for the upcoming high-luminosity LHC (HL-LHC) upgrade, which is slated to begin in approximately 2027 and aims to increase the design luminosity of the LHC by at least a factor of 5 ??HL-LHC. Higher luminosity allows for the production of more rare physics events, but also dramatically increases the incidence of unwanted "pileup" events; mitigating this is one of the major efforts of the CERN physics community during the lead-up to the HL-LHC run.

The LHC ring accelerates beams to a center-of-mass collision energy of 13 TeV. The ring consists of two primary beam pipes, one containing a "clockwise" beam and the other containing a "counterclockwise" beam, which overlap at the four primary collision points. These correspond to the four major LHC physics experiments: ATLAS ?? and CMS ??, which are "general purpose" physics detectors, LHCb ??, which is specialized to study the physics of hadrons containing bottom (or "beauty") quarks, and ALICE ??, which is primarily designed for heavy-ion physics. Due to the highly sensitive nature of their instrumentation, these detectors are buried approximately 100 meters underground to avoid interference from high-energy cosmic rays.

2.2 The ATLAS Detector

The work contained in this dissertation was performed using the ATLAS (A Large Toroidal LHC ApparatuS) detector. Together with its "sibling detector" CMS (the Compact Muon Solenoid), the ATLAS detector is one of CERN's two "general purpose" detectors, so-called because it is designed to observe a wide variety of high-energy physics phenomena. The ATLAS detector is cylindrical, with a diameter of 25 meters and a length of 44 meters ???. Upon colliding at the interaction point in the center of the detector, the constituent particles of each proton will interact with one another through a host of different physical processes, producing a variety of new particles traveling at high velocities. The ATLAS detector is comprised of several specialized subsystems, each of which is responsible for logging different properties of these decay products. Each of these subsystems is optimized to detect specific properties of a collision event; taken together, they provide a detailed snapshot of the moment of interaction. The Inner Detector (ID), a high-resolution detector composed of primarily of silicon, measures the tracks of particles close to the primary vertex. It is encased within a 2T solenoidal magnetic field that bends the tracks of charged particles produced at the collision point, allowing particle momenta to be measured as a function of the curvature of their trajectories. Enclosing the Inner Detector are two nested calorimeter systems: a liquid-argon-and-lead electronic calorimeter (ECAL) that measures the energy of incident photons

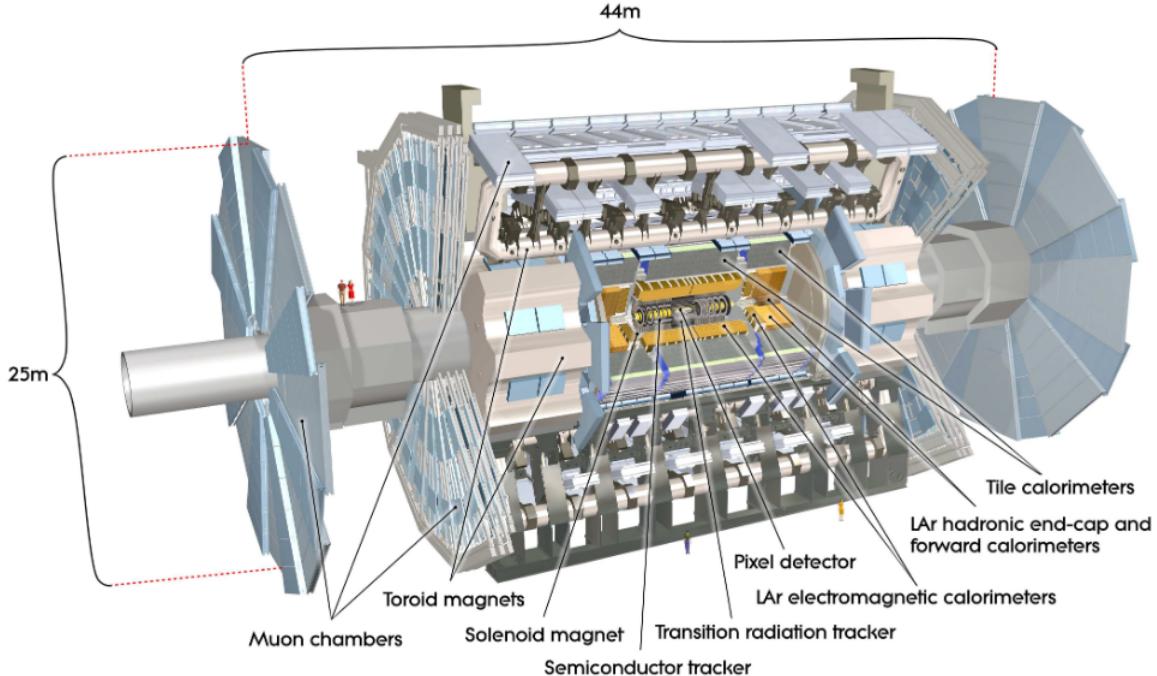


Figure 2.2: A diagram of the various subsystems of the ATLAS detector. ??

and electrons, and a primarily scintillator-tile-and-steel hadronic calorimeter (HCAL) that measures the energy of hadronic showers. Finally, a muon spectrometer composed of gas-filled tubes and chambers records the tracks of muons, which do not decay in the calorimeters like most other particles ???. The entirety of the muon spectrometer is suffused with magnetic field generated by 24 toroidal magnets, eight of which lie in the barrel, generating a field of 1T and eight of which lie in each endcap, generating a field of 0.5T ???. These magnets bend the trajectories of muons in a manner similar to the inner solenoidal magnet system, allowing the momenta of particles to be measured based on the track curvature. The ATLAS detector is also equipped with a highly sensitive trigger system, which uses a wide variety of physics signatures in each of these detector subsystems to determine which collision events to store for later analysis ???. A diagram of the ATLAS detector is shown in figure ??.

The ATLAS detector uses a right-handed coordinate system, with the central interaction point at $z = 0$. The x -axis points from the interaction point to the center of the LHC ring, while the y -axis points upward. However, particle trajectories are more commonly measured not in (x,y,z) coordinate space but utilizing angular variables (r and ϕ) in the transverse plane and pseudorapidity η in the longitudinal plane, where $\eta = -\ln(\tan(\theta/2))$. This is illustrated in figure ??.

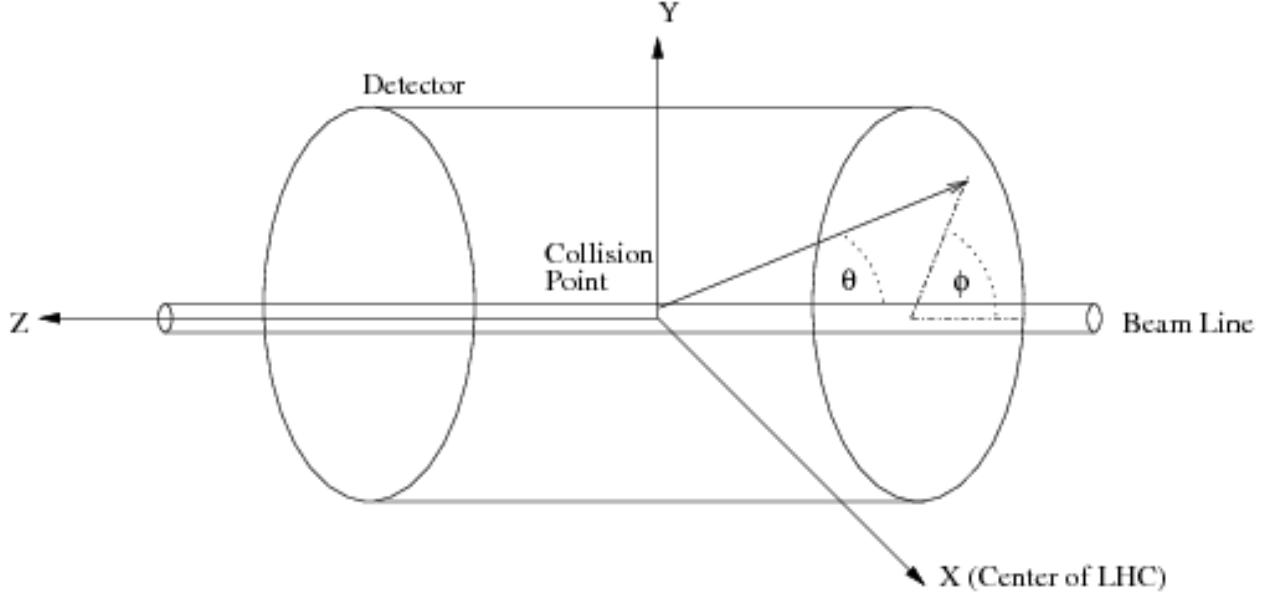


Figure 2.3: The coordinate system used to define the ATLAS detector geometry. ??

2.2.1 Inner Detector

The Inner Detector is composed of five highly granular subdetectors. The closest to the beam, the silicon-based Insertable B-Layer, was added between the LHC’s first and second runs in order to improve precision track-vertex measurement for tasks such as b-hadron tagging and tau lepton identification ??, methods discussed more in section ?? . Proceeding outward from the IBL are the silicon-based Pixel detector and Semiconductor Tracker (SCT), followed by the gas-based Transition Radiation Tracker (TRT) ?? . The inner detector is encased within a solenoidal magnetic field which causes charged particle tracks to bend according to their charge, as detailed in ?? .

The IBL is composed of 26,880 $250 \times 50 \mu\text{m}^2$ silicon pixels arranged in an 80 column by 336 row geometry. The pixel modules are supported by 14 support staves, tilted at approximately 14 degrees from the nominal to achieve near-complete cylindrical coverage of the barrel region. The IBL is approximately 33.25 mm from the beam axis, necessitating a replacement of the beryllium beam-pipe with a smaller-radius one for detector reintegration ?? .

Silicon is an optimal material for precision trackers due to its status as a semiconductor. Silicon wafers can be easily ”doped” with atoms of other elements, leading them to carry either an excess of electrons (n-type semiconductors) or a deficit of electrons (p-type semiconductors). By putting the two together, an n-p junction is formed, leading to a one-way current gate known as a diode. A voltage is then applied to the diode to halt current flow (”reverse biasing”). When ionizing particles such as those produced in collisions pass through the silicon, they will produce a cascade of charge carriers that will induce a current in the diode, which can then be measured by the precision front-

end electronics attached to the silicon wafers ??.

The pixel detector is similarly composed of silicon pixels, each of which is either $50 \times 400 \mu\text{m}^2$ (nominal) or $50 \times 600 \mu\text{m}^2$ (near the front-end readout electronics due to spatial constraints). The detector contains 1744 identical pixel sensors, each of which contains 47232 pixels. Three layers of pixels in the barrel region sit at $r = 50.5, 88.5$ and 122.5 mm from the beam axis, while three layers of pixels in the endcap region sit at $r = 495, 580$, and 650 mm from the beam axis. The silicon wafers in the pixel detector are "n+ type" strips (n-type semiconductors with extra doping) embedded in an n-type bulk, rather than the traditional n-type and p-type semiconductors due to the fact that excessive radiation from the beam can cause standard n-and-p-type semiconductors to change into one another ??.

The SCT is composed of longer silicon microstrip wafers made out of standard n-p type semiconductors. It consists of four layers of 2112 total silicon microstrip modules in the barrel region and 1976 microstrip modules in the endcap regions, with 770 microstrips per sensor module. The strips are 12cm long and $80 \mu\text{m}$ wide; the distance of modules from the beam pipe ranges from radii of $r = 299$ mm to $r = 514$ mm in the barrel and $r = 852.8$ mm to $r = 2720.2$ mm in the endcaps. The SCT is designed using less-costly materials than the TRT and IBL, but serves a valuable complementary function ??.

Finally, the TRT consists of polyimide "straws" filled with a gas mixture (70% Xenon, 27% Carbon Dioxide, and 3% Carbon Tetrafluoride), each 4mm in diameter and 144 cm long. In the center of each straw tube is a $31 \mu\text{m}$ in diameter grounded gold-plated tungsten wire filament that serves as an anode. Charged particles travelling through the gas medium will ionize it, producing a cascade of ions that will drift to the anode wire. The TRT is uniquely designed to detect low-energy X-ray transition radiation photons, which allows for better identification of electron candidates ?? . The barrel region consists of 73 layers of sensor modules ranging from $r = 554$ mm to $r = 1082$ mm from the beam pipe, while the endcaps straws, arranged in a "wheel" shape, stretch from $r = 852.8$ mm to $r = 2720.2$ mm from the beam pipe.

Additionally, two Minimum Bias Trigger Scintillators, composed of plastic scintillator tiles, sit on the circular endcaps of the ID in order to provide trigger capabilities in the forward region $|\eta| > 2.5$. The MBTS endcap scintillator consists of 2 cm thick polystyrene scintillator disks mounted on each endcap at a distance of 3.6m from the collision point at .

Together, these subsystems form the ATLAS detector's Inner Detector. Each subdetector has a different spatial resolution- the IBL has a resolution of approximately $10 \mu\text{m}$, the pixel detector has a resolution of $12 \mu\text{m}$, the SCT has a resolution of $16 \mu\text{m}$, and the TRT has a resolution of $130 \mu\text{m}$. The combined effect of these subdetectors is thus to record particle tracks produced in collisions with very high granularity within the range $|\eta| < 2.5$, allowing for sophisticated physics analyses to be performed.

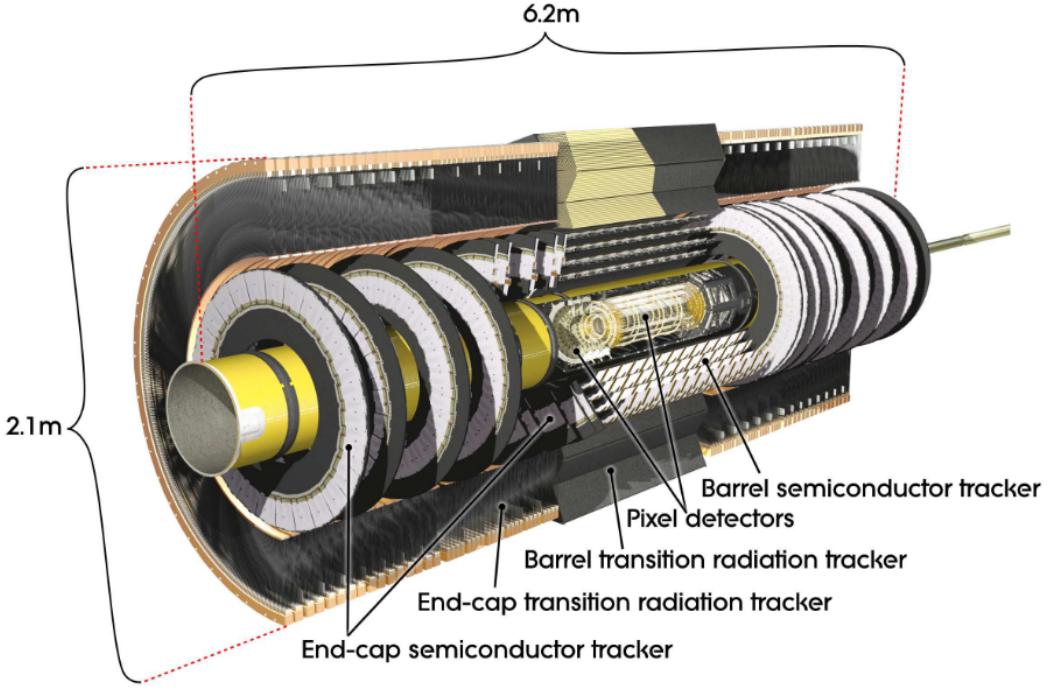


Figure 2.4: An illustration of the Inner Detector. ??

2.2.2 Solenoid Magnet

Track-bending in the Inner Detector is performed by a solenoid magnet providing a field of 2T throughout the ID bulk. The solenoid magnet is constructed of a single layer of high-strength aluminum-stabilized Nb/Ti conductor. The magnetic field it provides is axial (in the z direction), and thus the direction of bending of charged tracks in the ID is in the ϕ direction.

2.2.3 Calorimeters

Measurement of particle energies is performed by the ATLAS detector's two calorimeters. Both calorimeters provide a wider η coverage than the Inner Detector, allowing for measurement of hadronic jets and electromagnetic showers that fall outside the $|\eta| < 2.5$ precision tracking range. The liquid-argon based calorimeters function differently than the silicon-based tracking Inner Detector: they do not require the incident particles to be charged, and record energies of particles as they shower (decay) in the calorimeter material rather than simply recording their tracks as they pass through. Broadly speaking, there are two types of calorimeters: homogeneous calorimeters, which are composed entirely of active "measuring" material such as liquid argon or polystyrene scintillator tile, and sampling calorimeters, which interleave layers of this active material with slabs of material designed to cause the incident particles to shower ?? . Both the electronic and hadronic

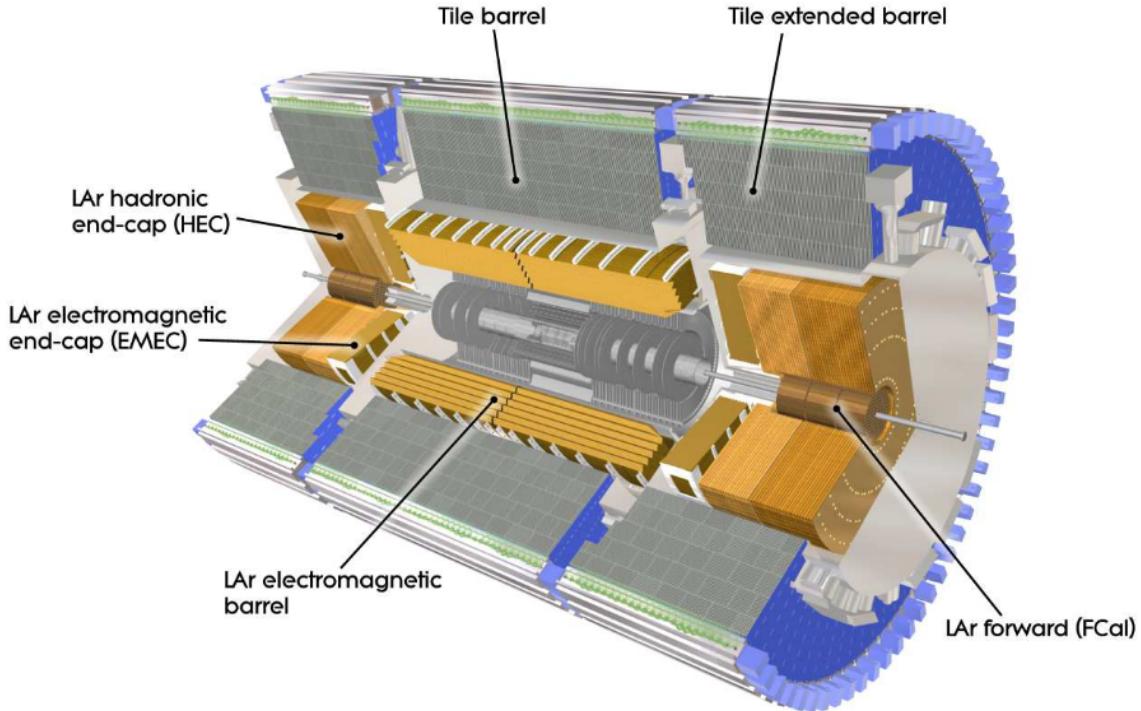


Figure 2.5: An illustration of the ATLAS calorimeter systems. ??

calorimeters are sampling calorimeters; however, they utilize different materials in both the active and non-active layers of the calorimeter. A diagram of the calorimeter system is shown in figure ??.

2.2.3.1 ECAL

When electrons and photons enter the non-active layer of a sampling calorimeter, they interact with the electric fields produced by the atoms of the layer material and produce showers of bremsstrahlung photons and electron-positron pairs, which then produce electrical signals in the Liquid Argon active layers that are read out by specialized electrodes. The depth of the electronic calorimeter thus must be chosen carefully in order to capture the entirety of a typical EM shower (measured in "radiation lengths" X_0 , this varies depending on the absorber material). All ECAL subsystems are made of the same materials: liquid Argon (LAr) sampling layer with lead-plate absorber layers, arranged in a unique "accordion" geometry that gives the ECAL complete coverage in ϕ . Due to the high atomic number of lead, the electron clouds of the atoms of the absorber plates are well-populated, thus leading to a high likelihood that photons or electrons will shower within the calorimeter bulk.

The ATLAS electronic calorimeter consists of two primary subsystems: one barrel calorime-

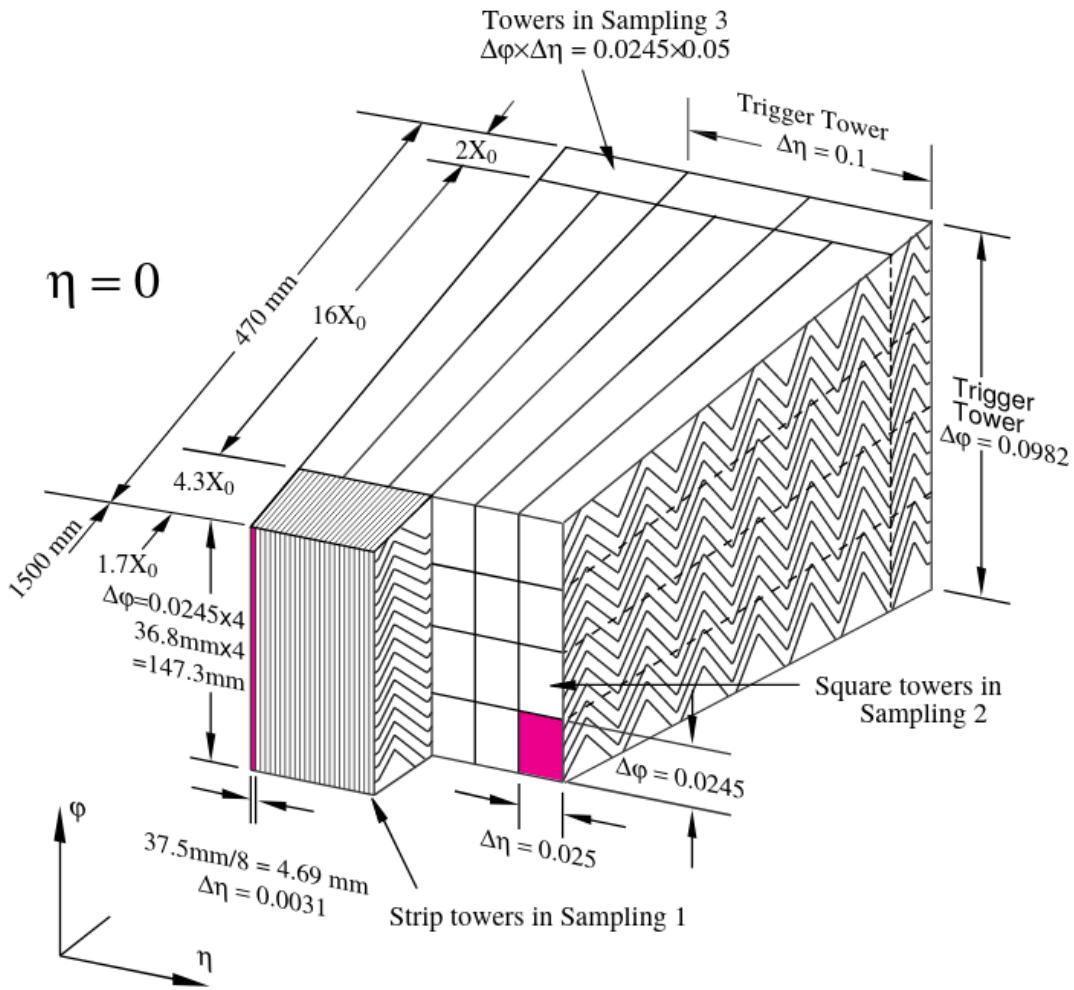


Figure 2.6: A cutaway diagram of the barrel ECAL depicting the "accordion" absorber geometry ??

ter, covering $|\eta| < 1.475$, and two end-cap calorimeters, covering $1.375 < |\eta| < 3.2$. The barrel calorimeter is composed of two identical half-barrels, separated by a 4mm horizontal gap at $z=0$, along the plane of the beam pipe. The lead absorber plates, readout electrodes, and "honeycomb" spacers that form the cavities for the liquid argon layer are all zigzag-shaped and arranged in a circular "starburst" pattern around the barrel ??, as depicted in figures ?? and ???. Each half-barrel consists of 1024 accordion absorbers with their associated interleaved spacers and electrodes, weighs 57 tons, and is 3.2 m long, with an inner radius of 2.8 m and an outer radius of 4.0 m. There are three layers of absorbers in the inner precision-measurement region ($0 < |\eta| < 1.35$) of the barrel and two in the higher- η region ($1.35 < |\eta| < 1.475$), each with a different segmentation; the angle, lateral orientation and length of the accordion waves varies with η .

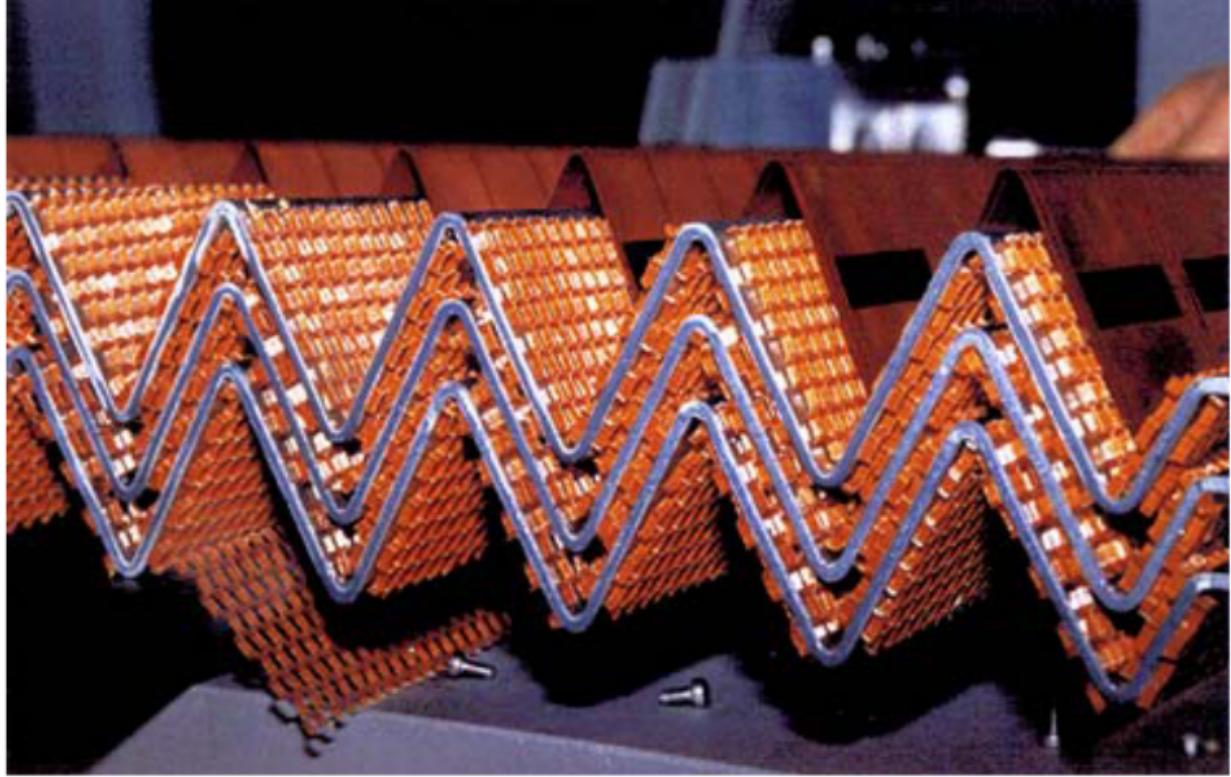


Figure 2.7: A photograph of an ECAL absorber. ??

Each end-cap calorimeter is composed of two wheels, one large outer wheel covering $1.375 < |\eta| < 2.5$, and one smaller inner wheel covering the range $2.5 < |\eta| < 3.2$. Each endcap is 63 cm thick, weighs 27 tons, and contains 768 accordion absorbers in the outer wheel and 256 absorbers in the inner wheel. Like the barrel, the calorimeter structure in the endcaps consists of three layers of absorber modules in the region nearest the beam pipe in the outer wheel ($1.5 < |\eta| < 2.5$) and only two layers both nearest the edges of the outer wheel ($1.375 < |\eta| < 1.5$) and in the entirety of the inner wheel. In the endcaps, the "zigzag" accordion folds are oriented longways, in the z-direction ??.

In the barrel, each absorber is at minimum 22 radiation lengths thick (though this varies up to 33 radiation lengths as a function of pseudorapidity η), thus ensuring that electromagnetic showers are contained within the calorimeter bulk. In the endcaps (except for the outer wheel edge region $1.375 < |\eta| < 1.475$), the calorimeter systems are at minimum 24 radiation lengths thick.

In addition to the accordion sampling layer folds of the ECAL bulk, a presampler detector composed of 64 modules in the barrel and 32 in each endcap provides information about electromagnetic objects before they enter the sampling calorimeter. The presampler modules consist of interleaved electrodes between glass-fiber composite plates, filled with liquid argon in the gaps. In the barrel, the presampler is 11mm thick and covers the entire barrel range; in the endcaps, the

presampler is 4mm thick and covers the range ($1.5 < |\eta| < 1.8$) ??.

2.2.3.2 HCAL

Hadronic showers typically have both electromagnetic and non-electromagnetic components, and, as such, are typically much larger than electromagnetic showers. As such, the hadronic calorimeter is required to be much deeper than the electronic calorimeter, in order to longitudinally contain hadronic showers in their entirety. The characteristic length scale for hadronic showers is the nuclear interaction length (λ_{int}), which can be up to 30 times larger than the radiation length X_0 . The ratio λ_{int}/X_0 is proportional to atomic number; thus, placing at least one layer at the front of a hadronic calorimeter that uses a high-atomic number material such as lead in its absorbers is essential in order to distinguish between purely electromagnetic showers and the electromagnetic components of hadronic showers. For the ATLAS detector, this purpose is served by the ECAL: a significant fraction of the energy of hadronic jets is deposited in the ECAL as electromagnetic activity due to the presence in jets of charged hadrons such as pions. In deeper layers, calorimeter absorbers should be constructed of a dense material such as iron or steel, so that hadrons are able to interact with the nuclei of absorber-layer atoms and produce signals in the active layers ??.

The ATLAS HCAL is composed of three major subsystems: the TileCal, which uses polystyrene scintillator tiles in the active layer and steel in the absorber layer; the LAr endcap calorimeters, which use liquid argon as the active layer and copper as the absorbers; and the LAr forward calorimeters, which use liquid argon as the active layer and contains both copper and tungsten absorber layers.

The TileCal is located immediately behind the electronic calorimeter, and covers the pseudorapidity range $|\eta| < 1.7$. It consists of a central barrel region and two extended barrel regions, one on each side of the detector cylinder, as depicted in figure ?? . The central barrel is 5.8 m long and each extended barrel is 2.6m long; all three have an inner radius of 2.28m and an outer radius of 4.25m. This corresponds to a total radial depth of approximately $7.4 \lambda_{int}$. Additionally, each subsystem is composed of 64 wedge modules, each spanning 5.625° in ϕ . Each wedge consists of a series of trapezoidal scintillator tiles (3mm thick, but varying in width and height depending on position in the wedge) interleaved with steel absorber; two fiber-optic cables connect each scintillator tile to two photomultiplier tubes in order to enable accurate readout. Each tube connects to multiple tiles; these readout channels are structured in order to form a system of calorimeter cells.

When ionizing particles produced by interactions in the steel absorber reach the scintillator tiles, they produce ultraviolet light. The scintillator tiles are composed of polystyrene doped with 1.5% p-terphenyl (PTP) and 0.044% 1,4-bis(5-phenyloxazol-2-yl) benzene (POPOP), which serve to shift the ultraviolet photons into visible light. These are then read out through the photomultiplier tubes ??.

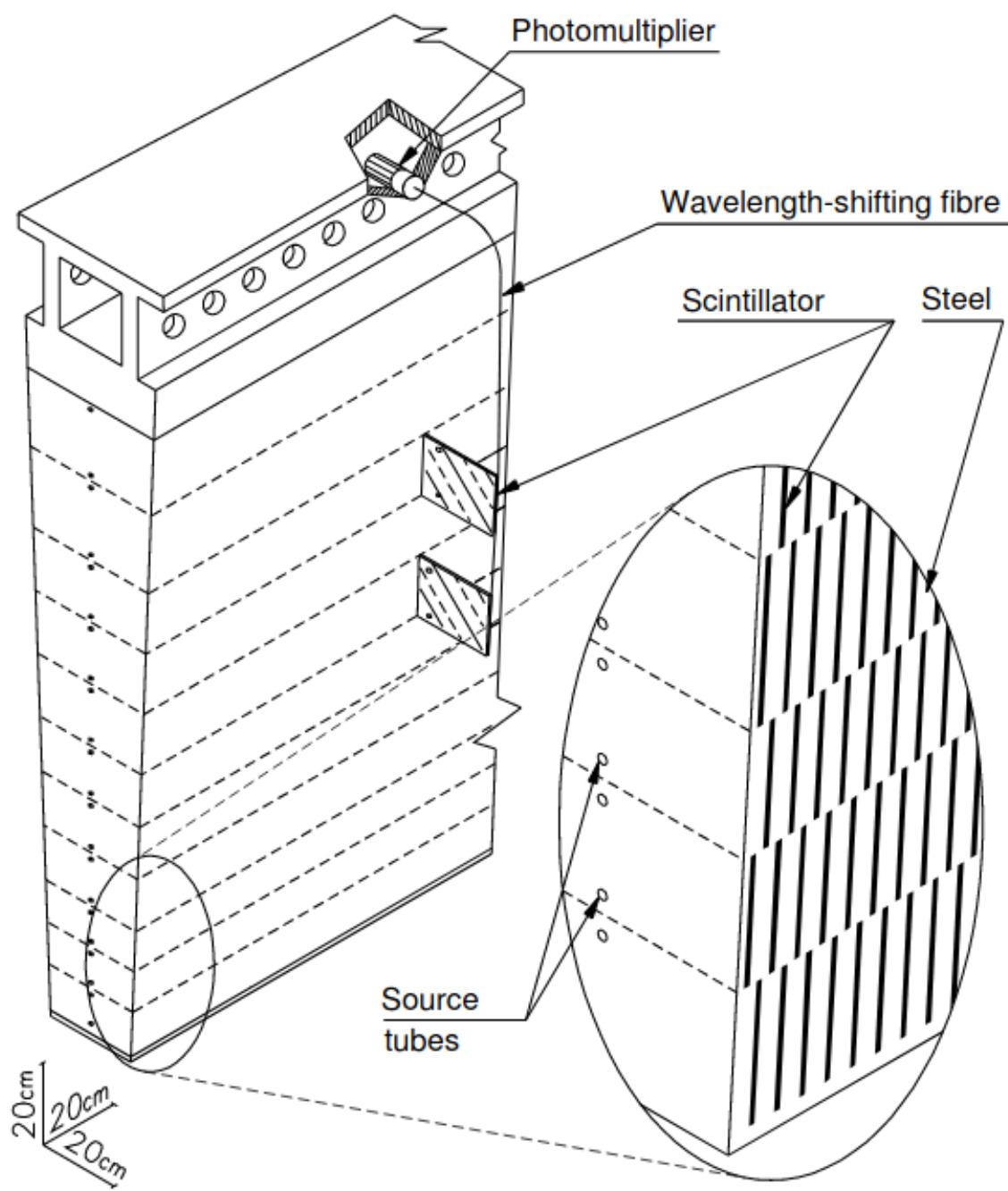


Figure 2.8: A diagram of the TileCal geometry ??

The Hadronic Endcap Calorimeter (HEC) uses flat copper plates as the absorber layer and liquid argon as the active material, and covers the range $1.5 < |\eta| < 3.2$. It consists of two wheels in each endcap, each wheel containing two longitudinal sections. Each wheel is cylindrical with an outer radius of 2030 mm, and consists of 32 identical wedge modules. The front wheels contain 24 copper plates each; the rear wheels contain 16 each.

Finally, the LAr forward calorimeters (FCAL) sit in the very forward region closest to the beamline, $3.1 < |\eta| < 4.9$, and are designed to measure the energy of very forward, often low-energy hadronic jets. The FCAL is unique in that, due to its forward position, it experiences a very large particle flux; it uses very thin active layers in order to mitigate the potential ion buildup this may cause. The FCAL contains three subdetectors, labelled FCAL1, FCAL2 and FCAL3 proceeding outward from the beamline.

Because it lies outside the eta range of the electromagnetic calorimeter, the FCAL must also contain an electronic calorimeter component in order to capture the electronic components of hadronic showers. FCAL1 is this electronic calorimeter, and uses copper as its active material. It is composed of a matrix of copper plates with circular holes in them, through which are pushed copper electrode rods utilized for readout. The gap between these rods and the plates is then filled with LAr; this setup offers very fine control over the active layer thickness. This can be seen in figure ??.

FCAL2 and 3, however, are optimized for the much longer characteristic shower scales of hadronic interactions. They are set up identically to FCAL1, but use tungsten readout electrode rods instead of copper ??.

2.2.4 Toroid Magnets

A series of three large toroid magnets (one in the barrel and one in each endcap) are designed to curve the tracks of muons in the muon spectrometer, allowing for their momentum to be measured using their curvature. Each toroid consists of eight rectangular air-core coils, wound in the $R - z$ plane. Barrel and endcap toroids are rotated at an angle of 22.5° with respect to each other in order to provide maximal coverage. Because the toroidal magnets are toroids and lie in the $R - z$ plane, they curve tracks in this plane as well; this is distinct from the solenoid, which curves tracks in ϕ .

Each barrel toroid magnet consists of a "racetrack-style" Aluminum-stabilized Nb/Ti/Cu conductor wound into two double pancakes; together, these supply a field bending-power ranging from 1.5 to 5.5 Tm in the pseudorapidity range $0 < |\eta| < 1.4$?. Endcap toroids are similarly constructed racetrack-style magnets, but are bolted and glued together with eight keystone wedges in order to withstand the Lorentz forces and hold the magnets together. Endcap toroids supply a bending-power of 1 Tm to 7.5 Tm in the range $1.6 < |\eta| < 2.7$??.

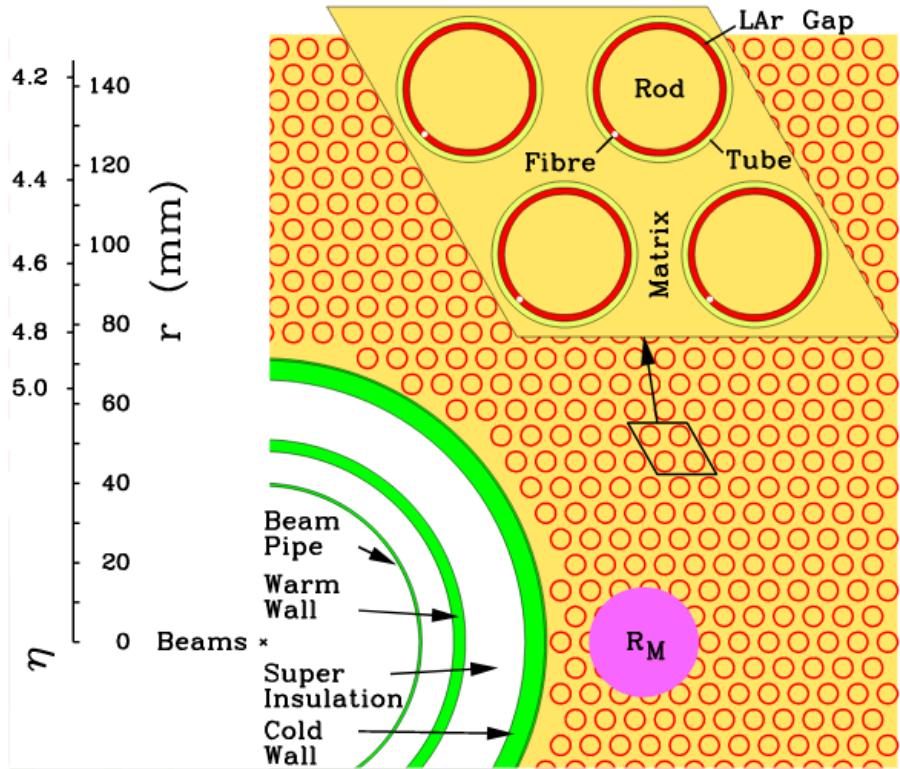


Figure 2.9: A diagram of the FCAL geometry ??

2.2.5 Muon Spectrometer

Unlike electrons and photons, muons are not stopped by the electronic calorimeter. This is due to the fact that a charged particle's energy loss as a function of unit distance due to bremsstrahlung is proportional to its energy and inversely proportional to the square of its mass; because muons are much heavier than electrons, this energy loss drops off substantially at the energies seen at the LHC. While at very high energies, muons again begin to lose most of their energy due to radiation, this occurs at muon energies exceeding a few hundred GeV. Thus, while muons may lose some energy in the electronic calorimeter, and at high energies may be slowed substantially, it is typically not enough to cause them to come to a stop ??.

Similarly, because muons do not interact through the strong nuclear interaction, like pions and other hadrons, they are not stopped by the high-Z material of the hadronic calorimeter. Thus, muons in ATLAS are "minimum ionizing particles", that is, particles that lose energy when travelling through material primarily through the process of ionization ?? . Therefore, a specialized detector subsystem must be built to measure the energies of muons exiting the detector.

The Muon Spectrometer, is, as its name suggests, not a calorimeter, but a device designed to measure the momentum spectra of outgoing muons (from which their energies can be calculated). It consists of four major types of detector chambers, two designed for muon tracking and two designed for triggering. A cutaway diagram of the full Muon Spectrometer is shown in figure ??.

The primary tracking detectors are the Monitored Drift Tube (MDT) chambers. A single MDT is an aluminum tube 29.970 mm in diameter, threaded through with a $50 \mu\text{m}$ radius tungsten-rhenium wire that is affixed with a stopper at each ends. The tube is pumped through with a mixture of 93% argon gas and 7% carbon dioxide, held at a pressure of three bars. Each anode wire is held at a potential of 3080 V, while each tube exterior serves as a cathode, and is grounded. As a muon passes through the tube, it will ionize the gas within; the electrons thus produced will drift to the high-voltage wire and produce electrical signals that can be read out by the specialized boards to which the tubes are attached. Because a muon passing through the tube will leave a trail of ions in its wake, each tube is able to record the minimum radius at which the track passes through the wire, as depicted in figure ?? . However, because this in itself is not enough to reconstruct a track, multiple layers of tubes must be placed together to form a chamber module.

There are 18 varieties of MDT chamber modules, each containing a different number of MDT tubes. Each chamber module consists of one or more "multilayers" of tubes, which consist of either three or four layers of tubes vertically separated by 6.5, 170, or 317 mm, depending on their position in the muon spectrometer. In the barrel, chambers are arranged in three concentric rings around the beam axis, at radii approximately 5m, 7.5m, and 10m from the beamline. In the end-cap region, chambers are arranged in large wheels perpendicular to the beam at approximately 7.4m,

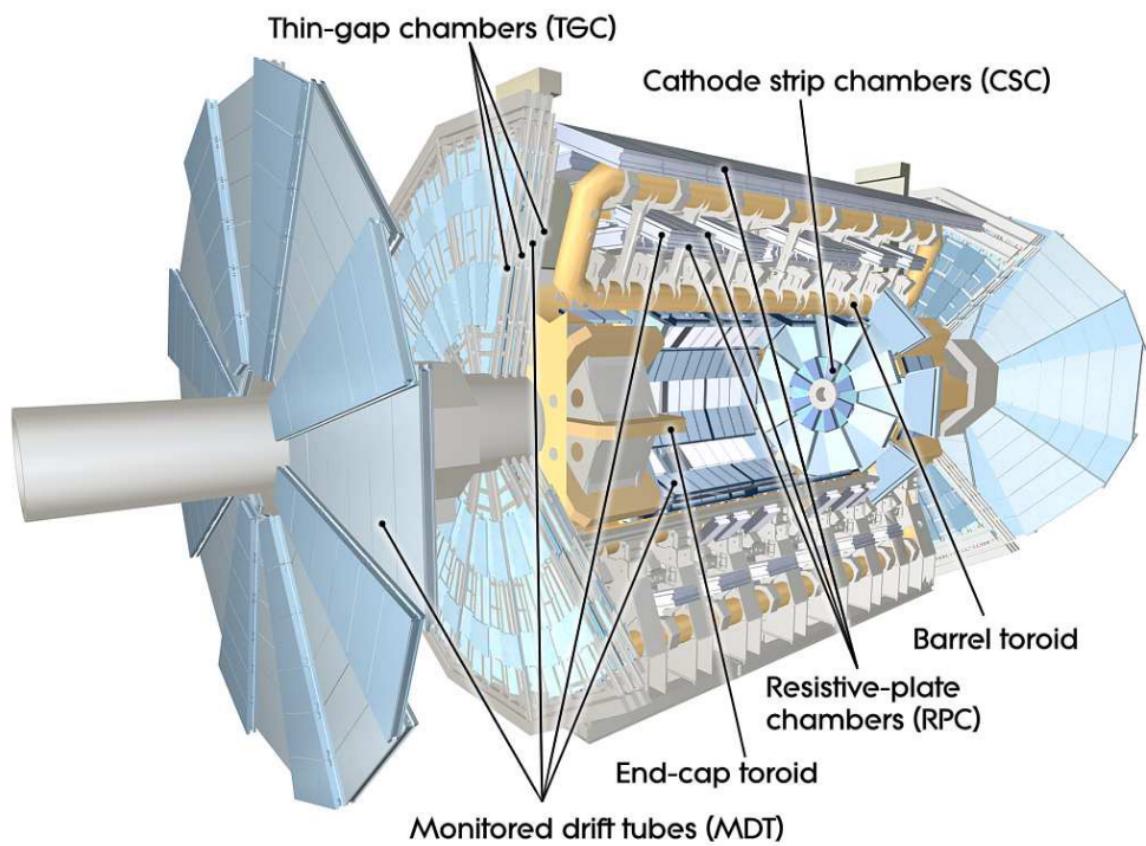


Figure 2.10: A diagram of the Muon System ??

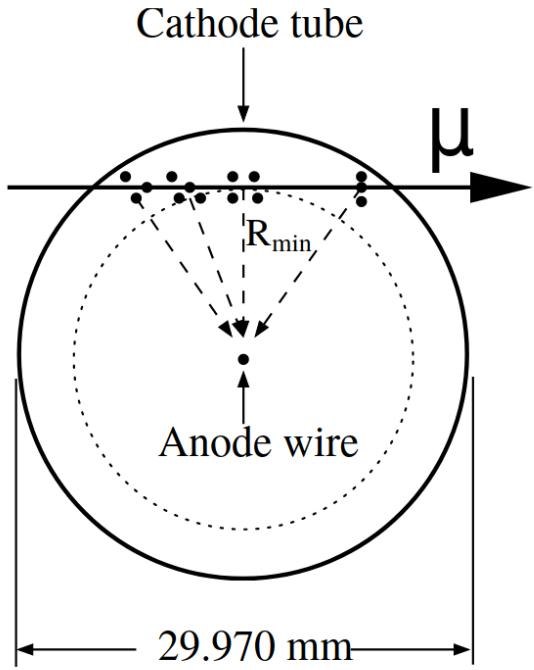


Figure 2.11: A diagram of a Monitored Drift Tube ??

10.8m, 14m, and 21.5m from the interaction point. Each MDT chamber has a spatial resolution of approximately $35 \mu\text{m}$. Overall, MDT chambers cover the region $|\eta| < 2.7$, though the number of chambers intersected varies as a function of η ??.

The MDT system geometry is shown in figure reffig:MDTChamber.

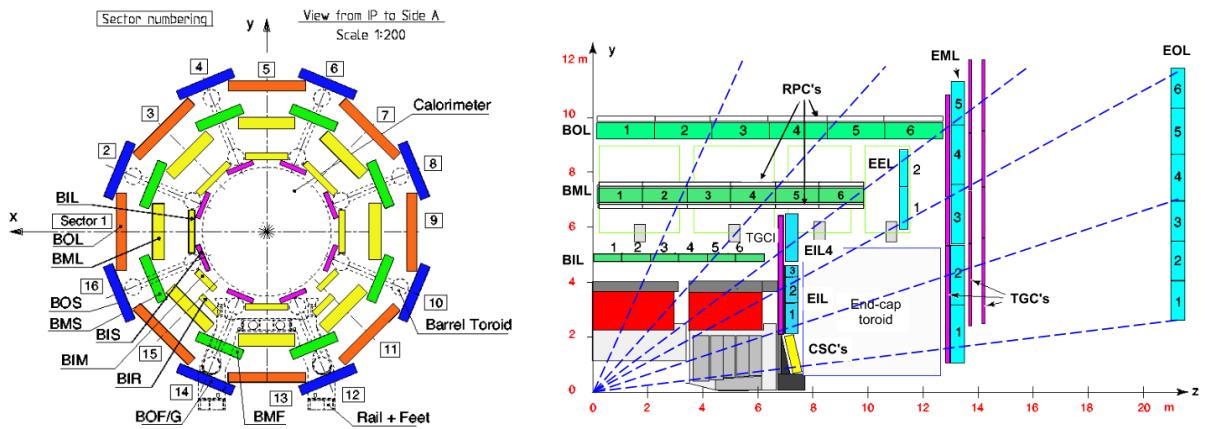


Figure 2.12: A diagram of the MDT chamber geometry from two positions, one looking down the beam pipe and one alongside the detector. ??

Closer to the beam, the event rate is higher due to beam halo effects caused by protons decaying

into pions (and in turn, muons) in the beam pipe???. Because the expected rate exceeds the tolerance of the MDT chambers, a different type of muon chamber is situated closer to the beam pipe in order to safely deal with the larger event rate. Cathode Strip Chambers (CSCs), are used in the inner endcap chambers in the range $2 < |\eta| < 2.7$.

Each CSC endcap consist of two circular disks, one containing eight large chambers and one containing eight small chambers. A CSC is a type of multiwire proportional chamber that functions similarly to an MDT: muons passing through the CSC will ionize the $Ar - CO_2$ gas within; electrons will then drift to anode wires and produce a current. Unlike MDTs, however, CSCs are flat, and threaded through with multiple anode wires. Unlike an MDT, a CSC utilizes cathode strips for readout, rather than reading from the anode wire directly: electrons produced during ionization will float to the anode wire; this will produce a charge screening effect that induces a current in the cathode strips. These are then read out; the strength of the current on each strip can be used to determine the path of the muon track.

Each chamber contains eight layers of cathode strips crisscrossed with anode wires and filled with a mixture of 80% argon gas and 20% CO_2 . Four of these layers are segmented in the η direction and four segmented in the ϕ direction, enabling spatial determination of both coordinates of a muon track. Each CSC chamber has a spatial resolution of approximately $40 \mu m$ in the "bending" η plane and 5 mm in the "non-bending" ϕ plane. The CSC endcaps are angled slightly toward the interaction point in order to improve resolution; the readout pitch is 5.31mm for the large chambers and 5.66mm for the small ones ??.

In addition to the precision MDT and CSC chambers, the muon system contains a series of chambers designed explicitly for quickly triggering on muon events. These are the Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs). These provide cursory information about the tracks of muons passing through the detector in order to rapidly determine whether an event should be stored for later analysis.

In the barrel region ($|\eta| < 1.05$), the trigger chambers are Resistive Plate Chambers (RPCs). RPCs are gas-filled detectors, like MDTs and CSCs, but do not contain a wire. Instead, in each RPC gas-gap module, two resistive plates made of plastic laminate are held 2mm from one another and suffused with a large electric field of 4.9 kV/mm. Incoming muons will then produce ionization avalanches that will induce a current on copper anode strips on the outside of the module through capacitive coupling. RPCs are filled with a gas mixture of 94.7% 1,1,2,2-Tetrafluoroethane ($C_2H_2F_4$), 5 % Isobutane (Iso-C₄H₁₀) and 0.3% Sulfur Hexafluoride (SF₆). Each RPC chamber consists of four gas-gap modules, segmented in ϕ in a manner similar to the CSCs. RPCs follow the same naming scheme as the MDT chambers, and are arranged in three concentric layers around the detector, one small at radii of 7.820 m, 8.365 m, and 10.229m, and one large at radii of 6.800m, 7.478m, and 9.832m.

In the endcaps ($1.05 < |\eta| < 2.4$) , the trigger chambers are Thin Gap Chambers (TGCs). TGCs serve two purposes: first, to provide trigger capabilities, and second, to complement the radial measurement of the endcap MDT chambers by providing a precision measurement of the ϕ coordinate of incoming muon tracks. TGCs are structured very similarly to CSCs, as they are also multiwire proportional chambers. However, TGCs are designed so the wire-to-cathode distance (1.4mm) is smaller than the wire-to-wire distance (1.8mm). TGCs are filled with 50% CO_2 and 45% n-pentane (n-C5H12); the choice of this gas in addition to the chamber geometry substantially increases the ratio of charges reaching the wire to ionizations produced by the incident particle. Two layers of copper strips in each chamber are segmented in ϕ and used as readout channels.

One set of doublet TGC chambers is placed just inside the inner MDT endcap, covering the range $1.05 < |\eta| < 1.92$ (divided radially into two concentric non-overlapping "Endcap" and "Small Wheel" regions); three others (one triplet and two doublets) are placed with the middle MDT endcap to form the "Big Wheel" (covering the range $1.92 < |\eta| < 2.4$). On the Big Wheel, TGC chambers are arranged in two circular wheels, each split into 12 segments. The outer wheel (further from the interaction point) contains 48 modules, each subtending 7.5° , while the inner wheel (closer to the interaction point) contains 24 modules each subtending 15° . Each module contains between 3 and 12 TGCs, depending on its position in the wheel. Similarly, the inner TGC layer is also arranged in a wheel shape, and contains a total of 180 TGCs refTGCs. The TGC wheels are depicted in figure ??.

2.2.6 Trigger

Due to the high rate of collision events delivered by the accelerator, it is impossible to save the data from every single collision produced in the detector. Luckily, only a subset of these are relevant for physics analysis; indeed, the majority of events produced in the detector are "minimum bias" events caused by low-energy interactions between beam protons that are useful only for performance studies ?? . A three-leveled system of triggers is therefore implemented in order to reduce the event rate from 40 MHz to a more manageable 200 Hz. The trigger system is a mix of both hardware and software based subsystems, and is optimized to balance the demands of complexity and speed: more sophisticated algorithms take more time to run, but this cannot get in the way of more data-taking. Each successive trigger is thus more complex than the last, with the lower-level trigger systems applying coarser selections and the higher-level triggers applying finer ones ??.

The first tier of trigger systems, called the L1 trigger, is hardware-based. It utilizes information directly received from the detector subsystems to search for a variety of noteworthy physics objects: high-transverse-momentum muons (transverse momentum here meaning the component of the momentum in the radial coordinate, denoted by p_T) are identified using the RPC and TGC

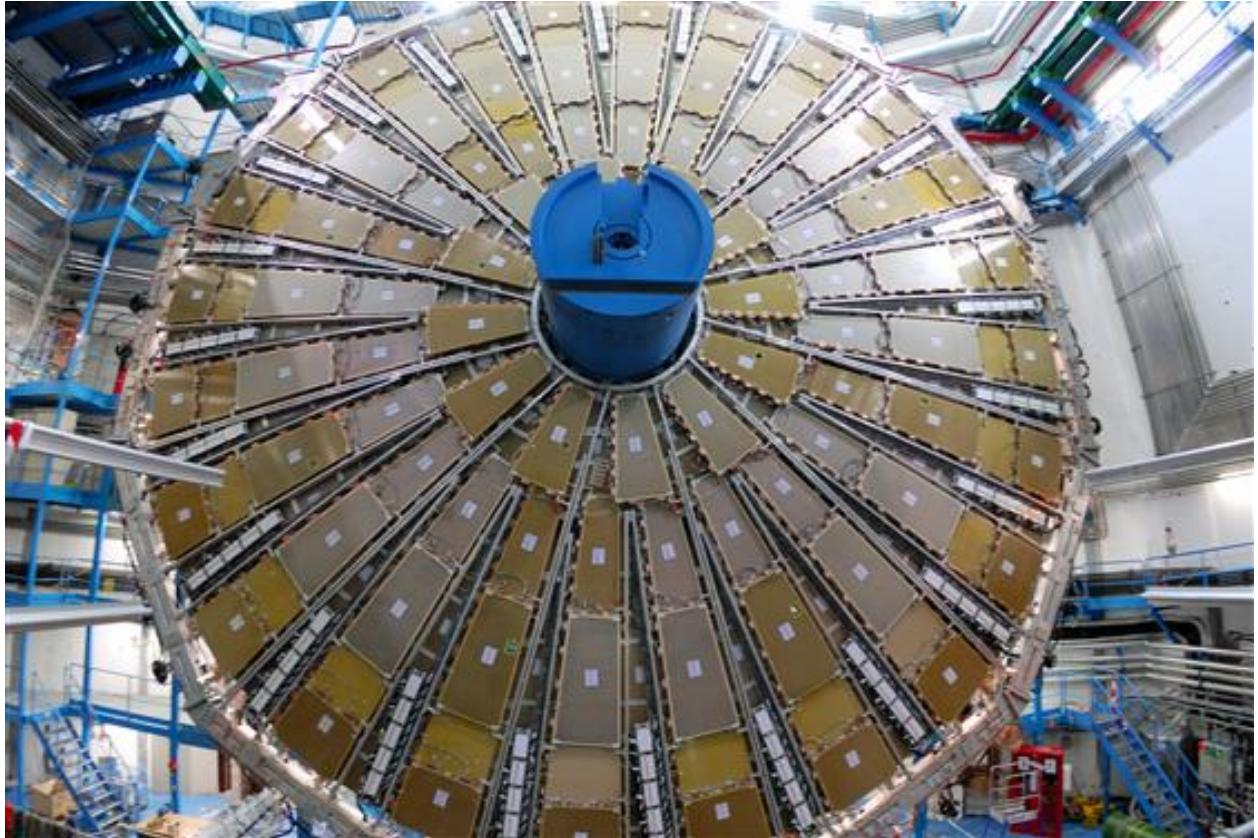


Figure 2.13: A photograph of the TGC wheels. ??

chambers discussed in section ??, while photons, electrons, tau leptons, and hadronic jets are identified using coarse calorimeter information. These triggers form the "L1Muon" and "L1Calo" trigger menus; this information in addition to that of other subsystems such as the MBTS are then sent to the Central Trigger Processor (CTP). The CTP then combines the information from these trigger systems and sends it to the L2 trigger; it also indicates spatial regions of interest using a third trigger called "L1Topo" and calculates quantities of interest such as missing transverse energy (here denoted E_T^{miss} , defined as the negative vector sum of the p_T of all objects in the event. E_T^{miss} serves as an indicator of the presence of unobserved neutrinos or other "invisible" decay products in the event). The combined activity of the L1 triggers serve to reduce the event rate to less than 1kHz.

After passing the hardware-based L1 trigger, events then proceed to the software based L2 trigger and Event Filter, which are known together as the High Level Trigger (HLT). HLT triggers utilize a full-granularity event display in order to perform trigger calculations, analyzing the regions of interest marked by the L1 trigger. It is at the HLT stage that reconstruction and matching of tracks using Inner Detector information, as well as jet, photon and electron object identification and tagging of objects such as b-jets and tau leptons is performed. All events that pass one or more

triggers in the HLT menu are recorded, reducing the event rate to the aforementioned 200 Hz. Triggers may also be 'prescaled' (for which only a set fraction of the events that pass are recorded); this allows analyzers to explore properties of events that normally occur at a rate much higher than can be handled by the trigger system. The triggers utilized in the analyses described in this dissertation are detailed in section ?? and ??.

CHAPTER 3

Experimental Methods

3.1 Experimental Methods

It is not possible to induce a particular interaction in a particle collider: every physics process that *can* occur from the collision of two protons at 13TeV *will* occur in the ATLAS detector. Thus, the process of studying rare Higgs interactions is not one of determining how to produce them in a laboratory setting, but of determining whether or not they are being produced in our detector at all. Because physics 'signal' processes are fundamentally indistinguishable from unwanted 'background' processes, ATLAS studies are statistical counting experiments which test the compatibility of the observed numbers of events with different detector signatures against the event counts that are predicted to occur under different physics hypotheses.

Performing these experiments effectively in the data-rich environment of the LHC requires many clever statistical and computational tools, which vary from analysis to analysis. However, in all analyses, it is vitally important both to accurately model what the physics processes of interest may look like in the ATLAS detector and to translate observed detector signals into high-level physics objects. The Monte Carlo simulation method is used for the former, while the latter is performed by the variety of identification and reconstruction techniques detailed in this chapter.

3.1.1 Monte Carlo

The Monte Carlo simulation method is a way of modelling how a given physics process that may occur at the point of collision may translates into signatures in the ATLAS detector. For a given physics process, say, $gg \rightarrow t\bar{t}H$, we must model both the initial process (i.e., what is the likelihood two gluons will interact to make a $t\bar{t}H$ event? What properties of the gluons make this more or less likely to occur? How will the top quarks and Higgs boson decay, and what might their decay products look like?), and the detector stage (how will the hadronic jets produced by decaying top quarks and the photons produced by a decaying Higgs look in the ATLAS detector?).

In order to make this problem computationally tractable, each stage is performed separately. In the first stage, the *generator* stage, physics events for a given process are generated according to their cross-sections σ (a measure of how likely they are to occur that is calculated using perturbation theory). A variety of different generator packages are used to produce events for different processes; common generator packages include PYTHIA ??, HERWIG ??, and aMCatNLO ?? . In many cases, intermediate decays are also performed by the generator software. Meanwhile, detector-level simulation is performed using the GEANT4 (GEometry ANd Tracking 4) package ??, which models the signatures of different particles produced in a collision as they pass through detector material.

Simulated events are then reconstructed in the same manner as real events. Because, unlike data events, Monte Carlo events include a "truth record" of what particles they contain, Monte Carlo simulations enable us to perform a number of useful tasks, including training machine learning algorithms to search for specific processes, model the relative contributions of various signal and background processes, and calculate various statistical and systematic uncertainties on our measurements ??.

3.1.2 Reconstruction and Tagging

Objects produced in the ATLAS detector produce a wide variety of signatures, many of which can be difficult to distinguish from one another. For this reason, dedicated algorithms and criteria are employed to reconstruct and identify the particles used in ATLAS analyses. A sample illustration of the ATLAS detector showing typical shapes of a variety of physics objects is shown in Figure .

3.1.3 Tracks

Track reconstruction is essential for identifying and measuring the properties of a wide variety of physics objects, including both muons and hadronic jets.

Initially, the "space points" (spatial coordinates) of potential particle hits are reconstructed using clusters of energy deposits in both the Pixel and SCT detectors. Track candidates are then identified combinatorially using this space-point information (with a minimum of three space-points per track candidate); these track candidates are then passed through an analytic weight-based "ambiguity solver" designed to weed out unphysical track candidates. To aid the ambiguity solver, a neural network is implemented in order to distinguish between isolated and merged clusters of energy deposits. After this, a neural network-based procedure is used to fit the identified tracks. ??.

Fitted tracks from the Pixel detector and SCT can be extended into the TRT if nearby TRT hits are identified. In addition, tracks that are seeded by TRT hits can also be identified and recon-

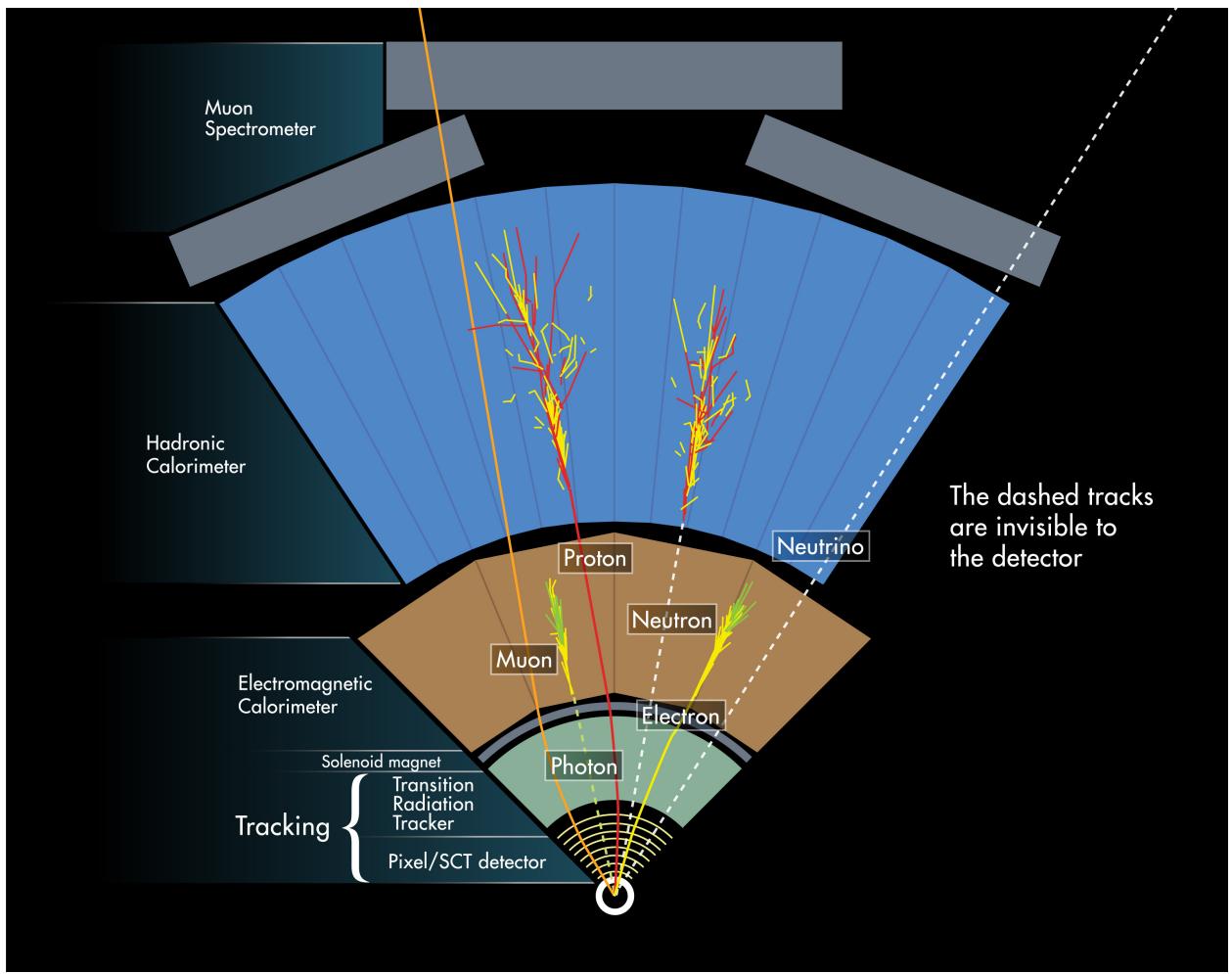


Figure 3.1: Shapes and signatures of a variety of objects in the detector ??

structed using an "outside-in" approach; this helps mitigate the potential loss of real tracks at the ambiguity-solver stage ??.

After tracks are identified and reconstructed, they are used to identify interaction vertices in the event. Each vertex candidate must contain at least two tracks, each with transverse momentum $p_T > 400$ MeV and $|\eta| < 2.5$, at least nine hits in the Pixel or SCT detector for tracks with $|\eta| < 1.65$ or at least eleven hits for tracks with $|\eta| > 1.65$, at least one hit in the first two pixel layers (i.e., the IBL and the inner b-layer), no more than one shared module (i.e. one shared pixel hit or two shared SCT hits), no holes in the Pixel detector, and no more than one hole in the SCT. An iterative combinatorial procedure is then used to fit all compatible tracks to vertex candidates; the primary vertex (PV) representing the initial protom-proton collision is chosen as the vertex with the greatest Σp_T^2 ??.

However, we note that, for $H \rightarrow \gamma\gamma$ events, a potential lack of charged particles in the final state means that a different primary-vertexing method utilizing ECAL clusters must be considered. This is detailed at length in section ??.

3.1.4 Clusters

Photons, electrons and jets are reconstructed using topological clusters of hits ("topo-clusters") in the electromagnetic and hadronic calorimeters. Clusters are reconstructed using a "seed-and-collect" method, where cluster-initiating "seeds" are defined as calorimeter cells with an energy four times greater than the expected noise threshold for that cell ??, ??.

After the identification of seed cells, calorimeter cells neighboring the seed cell with recorded energy greater than twice their noise threshold are added to the proto-cluster. If proto-clusters contain more than two local maxima with energy greater than 500 GeV in a single cell, the proto-cluster is split into two in order to account for potential overlap. Initially, seed cells in these split clusters may reside only in the ECAL sampling layers EMB2, EMB3, EME2 and EME3, or the forward calorimeter module FCAL0, after this splitting, clusters are then iteratively split again, with maxima now allowed in the other HCAL and FCAL layers (The first layer of the ECAL is not used to seed topo-clusters in order to reduce the likelihood of producing clusters of noise). In the case of overlap between multiple clusters, cells are assigned to the two cluster candidates with the largest maxima.

3.1.4.1 Electrons and Photons

Electrons and photons are defined using tracks that are matched to topo-clusters in the ECAL. Topo-clusters compatible with EM showers are selected and used to define Regions of Interest; those Regions of Interest are then matched to tracks in the ID. Track candidates are extrapolated

into the Regions of Interest using both the measured track momentum and the rescaled momentum measured in the relevant topo-cluster (the latter of which allows for accounting of radiative energy loss in the calorimeter).

For a track to be considered matched to a topo-cluster, under either momentum-definition-extrapolation, it must fall within $|\Delta\eta| < 0.05$ of its relevant topo-cluster and satisfy $-0.10 < q \times (\phi_{track} - \phi_{cluster}) < 0.05$, where q is the charge of the incident particle.

Topo-clusters matched to a single charged track are considered to be electron candidates, topo-clusters matched to two oppositely-charged tracks forming a vertex consistent with a photon are considered to be "converted" photon candidates (i.e., clusters resulting from a photon converting into an electron-positron pair in the ID), and topo-clusters matched to no tracks are considered to be "unconverted" photon candidates. In addition, single tracks that have no hits in the innermost layers of the ID are also considered as potential converted photon candidates??.

Since the start of Run 2, combined topo-clusters called "superclusters" have been implemented in the EM clustering process in order to capture bremsstrahlung photons and other energy lost in the calorimeter ???. For electrons, a supercluster seed must be a cluster with momentum $p_T > 1\text{GeV}$ matched to a track with at least four hits, while for photons, a supercluster seed must be a cluster with momentum $p_T > 1.5\text{GeV}$. Satellite clusters are then added to the seed to form a supercluster if they fall into a window $\Delta|\eta| \times \Delta\phi = 0.075 \times 0.125$ around the center of the seed cluster. For electrons, an additional satellite cluster search is performed, adding clusters that fall into the window $\Delta|\eta| \times \Delta\phi = 0.075 \times 0.125$ that are matched to the same track as the seed. For converted photons, the $\eta - \phi$ window is not used: all satellite clusters matched to one of the tracks of the converted photon vertex are added to the supercluster.

Following the identification of superclusters and tracks, energy calibration corrections are applied (determined using Monte Carlo simulation for photons and $Z \rightarrow ee$ decays for electrons)??, and the photon and electron candidate objects are passed to cluster-shape-based identification algorithms. For both electrons and photons, three identification working points are defined using cut-based multivariate discriminants based on the shower-shape variables.

For photons, the *loose* ID threshold is determined based on the following variables:

- Acceptance: $|\eta| < 2.37$, excluding the calorimeter crack at $1.37 \leq |\eta| < 1.52$
- R_{had} : the ratio of transverse energy deposited in the HCAL to transverse energy deposited in the ECAL for clusters with $0.8 \leq |\eta| < 1.37$
- R_{had1} : the ratio of transverse energy deposited in the first layer of the HCAL to transverse energy deposited in the ECAL for clusters with $0.8 \leq |\eta| < 1.37$
- R_{eta} : the ratio of the energy deposited in the ECAL in a 3×7 rectangle in the $\eta \times \phi$ plane

to the energy deposited in the ECAL in a 7×7 rectangle in the $\eta \times \phi$ plane, both centered on the calorimeter cell with the most deposited energy.

- Lateral shower width w_{eta2} : $\sqrt{\frac{\sigma E_i \eta_i^2}{\sigma E_i} - \frac{\sigma E_i \eta_i^2}{\sigma E_i}}$ (where E is the energy and η is the pseudo-rapidity of cell 'i', summed over all cells in a 3×5 rectangle centered around the most energetic calorimeter cell).

The *medium* photon ID threshold is determined based on both passage of the loose threshold and the variable $E_{ratio} = \frac{E_1 - E_2}{E_1 + E_2}$, where E_1 and E_2 are the leading and subleading energies deposited in calorimeter cells, respectively.

The *tight* photon ID threshold is determined based on passage of the medium ID threshold and the following shape variables in the strip layer of the ECAL

- Lateral Shower Width: $w_{s3} = \sqrt{\frac{\sigma E_i (i - i_{max})^2}{\sigma E_i}}$ (where E is the energy of a strip, ' i_{max} ' is the index of the highest-energy strip, and 'i' is the index of each strip with respect to i_{max} calculated in a $3 \times 2\eta - \phi$ rectangle centered around the strip with the maximum energy deposit)
- Total Lateral Shower Width: $w_{stot} = \sqrt{\frac{\sigma E_i (i - i_{max})^2}{\sigma E_i}}$ (where E is the energy of a strip, ' i_{max} ' is the index of the highest-energy strip, and 'i' is the index of each strip with respect to i_{max} calculated in a $20 \times 2\eta - \phi$ rectangle centered around the strip with the maximum energy deposit)
- ΔE_s : the difference between the second-largest strip energy and the minimum energy in the strips that lie between the largest- and second-largest strip energies
- f_1 : The ratio of the energy in the first layer to the energy in the whole EM cluster.
- f_{side} : the total energy outside the three central strips but within seven strips, divided by the energy of the three central strips

Each working-point threshold varies in bins of η . For loose and medium working points, converted and unconverted photons are not treated differently, but for tight working points, they are determined separately. ??

For electrons, the identification process proceeds similarly: working-points are defined using shower-shape variables $f_1, E_{ratio}, w_{stot}, R_{eta}, w_{eta2}, f_3, R_{had}$, and R_{had1} , as well as R_{phi} (the ratio of the energy deposited in the ECAL in a 3×3 rectangle in the $\eta \times \phi$ plane to the energy deposited in the ECAL in a 3×7 rectangle in the $\eta \times \phi$ plane, both centered on the calorimeter cell with the most deposited energy). Electron ID working-points also include the following track variables:

- n_{Blayer} : Number of hits in the B-layer

- n_{Pixel} : Number of hits in the Pixel
- n_{Si} : Number of hits in the silicon detectors (ID and SCT)
- d_0 : the transverse impact parameter relative to the beamline
- $|d_0/\sigma(d_0)|$: the impact parameter significance relative to its uncertainty
- $\Delta(p)/p = (p - p_{last})/p$: the momentum difference between the track perigee and its endpoint, divided by the momentum at perigee
- $eProbabilityHT$: the electron probability based on TRT radiation
- $\Delta\eta_1$: the difference in pseudorapidity between the cluster position in the first layer and the matched track
- $q \times (\phi_{track} - \phi_{cluster})$: where ϕ_{track} is the momentum-rescaled track extrapolated from the perigee and $\phi_{cluster}$ is the cluster position in the second ECAL layer.
- E/p : ratio of the cluster energy to the track momentum (used for $E_T > 150GeV$ only)

However, the photon identification is a cut-based process, while the electron identification process relies on a likelihood-based discriminant. For the analyses discussed in subsequent chapters, we utilize the *Medium* electron working point, with an efficiency of approximately 88% ???. Electrons are also required to satisfy $p_T > 10GeV$, $|\eta| < 2.47$ (excluding the calorimeter crack at $(1.37 < |\eta| < 1.52)$), have track impact parameter significance $|d_0/\sigma(d_0)| < 5$, and satisfy $z_0 \times \sin(\theta) < 0.5mm$ with respect to the primary vertex ??.

In addition, for the analyses discussed in this dissertation, all photon candidates must have $p_T > 25GeV$ and $|\eta| < 2.37$, with the region $1.37 < |\eta| < 1.52$ vetoed.

To aid in background modelling, additional photon ID working points are created. These "LoosePrimeN" working points involve photons that pass the 'loose' identification criteria, but fail one or more of N 'tight' identification criteria. The CP Analysis uses the LoosePrime4 working point, which is defined as photons passing the Loose identification but failing one or more of the w_{s3} , f_{side} , ΔE_s , and E_{ratio} criteria, while the Couplings analysis uses the LoosePrime3 working point (defined as photons passing the Loose identification but failing one or more of the w_{s3} , f_{side} , ΔE_s criteria) ??.

To distinguish between photons originating from the hard scatter (i.e., the process of interest in an event) and photons radiated off of other final-state objects, we consider the relative isolation of identified photons. Photons near large amounts of calorimeter activity are likely to be radiative photons ('non-prompt') radiated from final-state particles after the hard-scatter event, while photons that are isolated are more likely to originate from the hard-scatter.

Two types of isolation variables are considered: calorimetric and track-based.

The calorimetric isolation variable employed in the analyses discussed here is $E_T^{coneXX} = E_{T,raw}^{isolXX} - E_T^{core} - C$, where $E_{T,raw}^{isolXX}$ is the total calorimeter energy in a cone of $\Delta R = XX/100$ around the electron or photon of interest, E_T^{core} is the total calorimeter energy in a 5×7 rectangle in $\Delta\eta \times \Delta\phi$ around the barycenter of the electron or photon of interest, and C is a correction for pileup and leakage.

The photon track isolation variable employed in the analyses discussed here is $p_T^{cone20} = p_T^{cone} - p_T^{core}$, where p_T^{cone} is the total track momentum of all tracks with $p_T > 1\text{GeV}$ in a cone of $\Delta R = 0.2$ around the photon of interest and p_T^{core} is the total track momentum of all tracks with $p_T > 1\text{GeV}$ matched to the photon candidate. In addition to satisfying $p_T > 1\text{GeV}$, all tracks considered for this metric must also fall within $z = 3\text{mm}$ of the diphoton vertex and have $|\eta| < 2.5$.

The electron track isolation variable considered is $p_T^{varcone20}$, identical to p_T^{cone20} except for the fact that, rather than consider a constant-radius cone of $\Delta R = 0.2$, we consider a cone of radius $\Delta R = \max(\frac{10}{p_T[\text{GeV}]}, 0.2)$??.

Unlike in Run 1, photon isolation cut thresholds are defined as a function of photon energy rather than being fixed. The isolation thresholds used are *FixedCutLoose* ($E_T^{cone20} < 0.065E_T^\gamma$ and $p_T^{cone20} < 0.05E_T^\gamma$) and *FixedCutTight* ($E_T^{cone40} < 0.022E_T^\gamma + 2.45\text{GeV}$ and $p_T^{cone20} < 0.05E_T^\gamma$) for photons and *FCLoose* ($E_T^{cone20} < 0.2p_T$ and $p_T^{cone20} < 0.15p_T$) for electrons ??.

In analyses discussed in this dissertation, the diphoton vertex originating from the Higgs decay is often not the hardest vertex. This is because, in processes such as gluon-gluon fusion (ggF), many events contain a low final-state track multiplicity ???. The diphoton vertex is therefore identified using a Neural Network, trained on variables including the "photon pointing" (that is, the vertex position in z most compatible with the shower-shapes observed in the ECAL), the $\Delta\phi$ between the vector sum of the track momenta and the diphoton system (as determined by the ECAL), and the scalar momentum sums p_T and p_T^2 for the tracks in each diphoton vertex candidate ???. Photon energies and pseudorapidities are corrected to reflect this new vertexing procedure.

The efficiency of the Neural Network diphoton vertex compared to the hardest vertex for a number of targeted physics processes ("STXS bins") for the full Run-2 Couplings analysis is shown in Figure .

3.1.4.2 Jets

Hadronic jets are identified from topo-clusters using the $anti-k_t$ algorithm. Clustering algorithms iteratively follow the following procedure: for each topo-cluster i , calculate the distance between both the topo-cluster and every other topo-cluster j and between the topo-cluster and the beam B :

$$d_{i,j} = \min(p_{T,i}^{2k}, p_{T,j}^{2k}) \frac{((y_i - y_j)^2 - (\phi_i - \phi_j)^2)}{R^2} d_{i,B} = p_{T,i}^{2k} \quad (3.1)$$

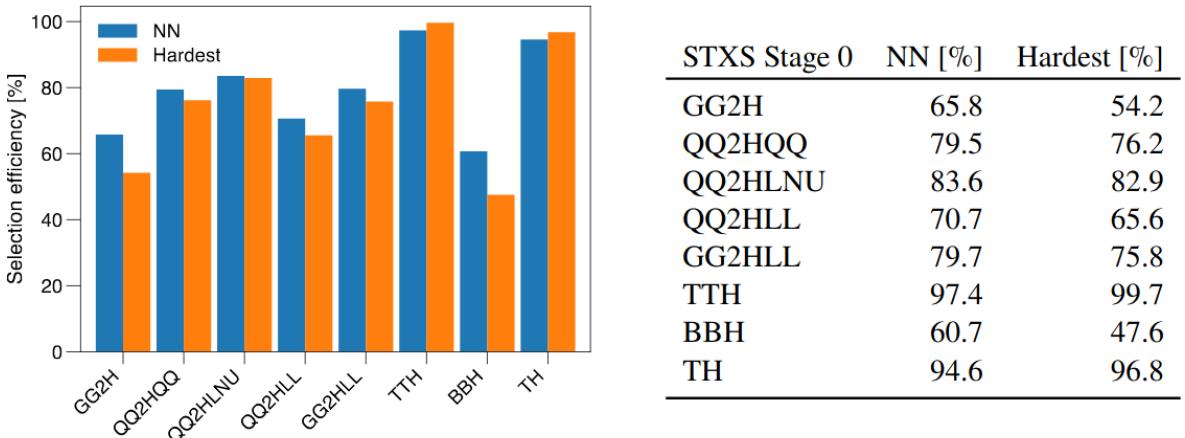


Figure 3.2: Shapes and signatures of a variety of objects in the detector ??

Where y_i is the rapidity of particle i , defined as $y = \frac{1}{2} \ln(\frac{E+p_L}{E-p_L})$ (where p_L is the particle's momentum in the beam direction), $R = 0.4$ (the jet radius), and k is some constant. For the anti- k_T algorithm, $k = -1$; other common clustering algorithms use $k = 1$ (k_t) or $k = 0$ (Cambridge-Aachen).

After calculating these metrics, combine the two cluster objects that are the smallest distance apart and re-calculate the distance metrics for the combined object. If a cluster object is closer to the beam than to another object, it is considered a jet and is removed from the list of objects for potential combination.

The algorithm is both infrared safe, meaning that very-low- p_T objects do not appreciably impact the jet-finding algorithm results, and collinear safe, meaning that splitting one high- p_T object into two collinear ones does not impact the results. Compared to other jet-finding algorithms, the anti- k_T algorithm is more robust to the effects of low-energy radiation ??.

For the Run-2 Couplings analysis detailed in section ??, the ParticleFlow algorithm detailed in ref ?? is implemented. This helps to provide clearer pictures of the individual constituents of a jet. The ParticleFlow algorithm is an intermediate step in jet reconstruction, which subtracts off energy deposits not associated with the primary vertex and replaces energy deposits that are associated with the primary vertex with the momenta of the relevant tracks. This is done using the modified diphoton vertex discussed previously.

The collection of jets is "cleaned" to remove jets that have signatures consistent with calorimeter noise, and jets that do not satisfy $|y| < 4.4$ and $p_T > 25\text{GeV}$ are rejected as well.

A Jet Vertex Tagger (JVT) multivariate discriminant is also used to reject pileup jets. This utilizes variables such as Jet Vertex Fraction (the p_T in a given jet that is track-associated to the primary vertex divided by the p_T in a given jet that is track-associated to all vertices) and R_{pT} (the

fraction of the p_T in a given jet that is track-associated to the primary vertex divided by the total jet p_T). Different JVT thresholds are used for PFlow Jets and the standard "EMTopo" jets that do not use the ParticleFlow algorithm ??.

3.1.4.3 b-jets

While in general it is very difficult to identify the type of particle that seeded a hadronic jet, jets originating from hadrons containing bottom quarks are often identifiable with some degree of accuracy. This is due to their long lifetimes compared to other hadrons, meaning that they will travel a short distance in the inner detector ($450\mu m$) before hadronizing into jets.

For EMTopo jets, the Boosted Decision Tree-based Mv2c10 algorithm is used, while for ParticleFlow jets, the Neural Network-based DL1r algorithm is used. The Mv2c10 algorithm is based on a Boosted Decision Tree trained on variables such as impact parameter significance, the presence and properties of a secondary vertex, and jet kinematic variables such as p_T and η ??; the DL1r algorithm is based on a Deep Neural Network ?? trained on the same input variables as the Mv2c10 tagger, with additional vertex variables included to allow for charm-quark jet discrimination. The DL1r tagger also features a recurrent neural network trained to take advantage of correlations between track impact parameters in discrimination. The Mv2c10 tagger working point chosen for the CP analysis has an efficiency of approximately 77%, while the DL1r working point chosen for the Couplings analysis has an efficiency of approximately 70%.

3.1.4.4 Muons

Muons are reconstructed using tracks from either the muon spectrometer and inner detector or the muon spectrometer only.

In the muon spectrometer, segments of hits along the bending plane in the MDT chambers are identified using a Hough transform ??, an image-processing technique from the field of edge-detection that maps observed line segments into a polar-coordinate feature space. The coordinate orthogonal to the bending plane is found using TGC and RPC hits, while segments in the CSC detectors are reconstructed using a separate combinatorial algorithm. Muon tracks are then reconstructed from the segments using a χ^2 algorithm- a track must contain at least two segments, except in the barrel-endcap transition region, where one segment is sufficient to build a track.

Tracks are used to define four types of muons: Combined Muons, for which a track is fitted using both the Inner Detector and Muon Spectrometer, Segment-Tagged, for which a track in the ID is matched to a segment (but not a full track) in the Muon Spectrometer, Calorimeter-Tagged for which a track in the ID is matched to a calorimeter deposit compatible with a muon, and Extrapolated Muons, reconstructed using segments in the Muon Spectrometer that are not

associated with any ID tracks but are compatible with the primary vertex. For all muons using the ID, criteria on the number of hits in each ID subsystem are employed to ensure track quality.

Muons must also pass identification criteria to distinguish them from potential fake muons (usually pions or kaons). Three main criteria are used:

- The charge-to-momentum ratio significance, defined as the absolute value of the difference between the ratio of the charge and momentum of the muons measured in the ID and MS divided by the sum in quadrature of the corresponding uncertainties
- ρ' , defined as the absolute value of the difference between the transverse momentum measurements in the ID and MS divided by the p_T of the combined track
- the normalized χ^2 of the combined track fit

The analyses detailed in this thesis use the Medium identification working point, which allows only Combined and Extrapolated muons. Medium Combined Muons require at least 3 hits in at least two MDT layers, except for tracks in the $|\eta| < 0.1$ region, where tracks containing at least one MDT layer but no more than one MDT hole layer are allowed. Medium Extrapolated muons require at least three MDT or CSC hits, and are only used in the region $2.5 < |\eta| < 2.7$. Additionally, the q/p significance is required to be less than 7 ??.

Like other objects, we require isolated muons. Muon isolation variables are defined identically to those of electrons for the ttH-CP analysis, while for the Couplings analysis, the ParticleFlow algorithm is used to augment the isolation procedure with a new isolation variable called $neflowiso$, corresponding to the energy deposited in neutral reconstructed ParticleFlow objects rather than raw calorimeter clusters ??.

For the standard isolation, we use the *FixedCutLoose* working point, corresponding to $E_T^{cone20} < 0.30 \times p_T^\mu$ and $p_T^{varcone30} < 0.15 \times p_T^\mu$. For the ParticleFlow isolation, we use $(ptvarcone30 + 0.4 \times neflowisol20) < 0.16 \times p_T$ for muons with $p_T < 50\text{GeV}$ and $(ptvarcone20 + 0.4 \times neflowisol20) < 0.16 \times p_T$ for muons with $p_T > 50\text{GeV}$.

Additionally, muons are required to satisfy $p_T > 10\text{GeV}$, $|\eta| < 2.7$, have track impact parameter significance $|d_0/\sigma(d_0)| < 3$, and satisfy $z_0 \times \sin(\theta) < 0.5\text{mm}$ with respect to the primary vertex ???. Muon momentum and energy scale and resolution calibrations are determined using $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ decays ???. A charge-dependent sagitta bias calibration is also applied to correct for a slight observed misalignment of the muon system and inner detector.

For the analyses discussed in this dissertation, all muon candidates must have $p_T > 15\text{GeV}$ and $|\eta| < 2.37$.

3.1.4.5 Overlap Removal

An overlap removal procedure is employed to avoid double-counting objects. It proceeds in the following order:

- Remove electrons within $\Delta R = 0.4$ of any photon
- Remove muons within $\Delta R = 0.4$ of any photon
- Remove jets within $\Delta R = 0.4$ of any photon
- Remove jets within $\Delta R = 0.2$ of any electron
- Remove electrons within $\Delta R = 0.4$ of any jet
- Remove muons within $\Delta R = 0.4$ of any jet

3.1.4.6 MET

The initial momentum of protons approaching the collision point is almost entirely in the beam direction, that is, there is no appreciable momentum component in the transverse direction. Thus, by conservation of momentum, we expect that, when we sum the momenta of all particles produced in the event, we should not see an appreciable momentum excess in the transverse direction. However, often such an excess does occur, indicating the presence of undetected particles that have caused the visible transverse momentum imbalance. Missing Transverse Energy, or MET, is something of a misnomer, as it is missing *momentum*, not energy, but it is one of the most valuable tools for studying rarely-interacting particles such as neutrinos in the ATLAS detector.

The MET is defined as the negative vector sum of the transverse momenta of all objects associated with the diphoton vertex, including electrons, photons, muons, jets, and any additional low- p_T tracks. For the purpose of MET reconstruction in $H \rightarrow \gamma\gamma$ events, τ leptons are treated as jets; this is allowable because they occur rarely in the parameter spaces studied and decay dominantly to hadrons ??.

3.1.4.7 Tau Leptons

We note that, though many physics analyses use τ leptons, they are not considered in any of the event signatures discussed in this thesis, so we refrain from discussing their reconstruction at length. However, we note that tau leptons can decay both hadronically and leptonically; as a result of this, their reconstruction depends on the reconstruction of electrons, muons, and jets.

3.1.4.8 Top reconstruction

In order to properly discriminate between CP-even and CP-odd $t\bar{t}H$ and tH processes, it is essential to reconstruct top quark kinematics from their decay products with some degree of accuracy.

We assume that each quark and gluon corresponds approximately to one jet per event in $t\bar{t}H(\gamma\gamma)$ events (i.e., we assume that in the regime we are considering, the top is not boosted), we can reconstruct the hadronic decay of the top quark by grouping together the correct set of three jets. However, determining which three jets are the correct ones is a nontrivial task; to make this problem tractable, we use a Boosted Decision Tree (BDT) trained in XGBoost ??.

A decision tree is a supervised machine learning algorithm that classifies events (or, in the case of top quarks, permutations) into one or more classes based on a set of input variables. It is fed labelled training data (that is, vectors of event properties and a numerical label corresponding to the event class), it then uses a greedy splitting algorithm to determine the criteria at each of a number of nodes that best separate the training data into classes.

Because individual decision trees are on their own considered weak learners (that is, they often under-fit or over-fit the training data), they are often aggregated into an ensemble in a procedure called "boosting", which weights the output of a number of decision trees together to construct one overall prediction ?? . There are many boosting methods, including Adaptive Boosting (AdaBoost) ?? which iteratively weights training events based on whether or not they are correctly classified, and gradient boosting, which trains each subsequent decision tree on the residuals of the prior. All the BDTs discussed in this dissertation employ gradient boosting, implemented in a variety of packages including XGBoost ??, the Toolkit for MultiVariate Analyses (TMVA) ??, and LightGBM ??.

The goal of the top reconstruction BDT is to correctly identify the jet triplets that correspond to the top decay products, from the set of all possible jet triplets. Two iterations of the BDT were developed following the same training procedure, one for the CP-Analysis that used EMTopo-Jets, and one for the Couplings Analysis that used ParticleFlowJets.

The training events come from the $t\bar{t}H$ Monte Carlo sample requiring 0 leptons, 3 jets and 1 b-tagged jet, identified using the criteria discussed in this chapter. From this sample, we denote "signal" events as the set of jet triplets truth-matched to a top quark, while the set of all non-truth-matched triplets in this sample is taken as background. The truth matching is done as follows:

1. starting from the truth top, identify W and b candidate daughter jets
2. Match the W candidates to 2 truth jets (and b candidates to 1 truth jet) by identifying the jets containing the most descendants of the truth particle weighted by p_T
3. Match each truth jet to the closest reconstructed jet in ΔR ;

4. if the 3 jets obtained this way are distinct, this top is considered as truth-matched and marked as a signal triplet. All other sets of 3 jets are considered background.

The input variables used for the top reconstruction BDT are:

- the four-vector information of the W boson candidate constructed from the W decays
- the four-vector information of the b quark candidate (matched to a truth b-quark)
- ΔR between the W and b candidates
- ΔR between the two jets comprising the W candidate
- the pseudo-continuous b-tag score of all three jets
- the tri-jet mass, i.e. mass of the top candidate

This trained BDT is then applied to both fully hadronic (i.e., events containing two hadronic tops) and semi-leptonic events (i.e., events containing only one hadronic top); the algorithm has been shown to be able to perform accurately in both cases.

Fully hadronic events For events passing hadronic-top pre-selection (0 leptons, 3 jets, 1 b-jet), the triplet with the highest Top Reco BDT score is taken as the primary top candidate. In hadronic events with six or more jets, a second top is reconstructed by selecting the triplet with highest score from the remaining jets. In events with four or five jets, the second top is taken as the sum of the jets remaining after reconstructing the primary top. In events with exactly three jets, only the primary top is reconstructed.

Leptonic events In single-lepton events passing leptonic-top pre-selection (1 lepton, 1 b-jet), the primary top candidate is reconstructed from a leptonic W candidate and a jet. The W candidate is constructed from a lepton and E_T^{miss} , which is assumed to originate from a neutrino.

The W candidate four-momentum is derived using a dedicated leptonic-top reconstruction algorithm. First, the combined lepton + E_T^{miss} system is assigned the mass of a W boson, which constrains the z -component of the neutrino momentum:

$$m_W^2 = \left[E_\ell + \sqrt{p_{\nu,x}^2 + p_{\nu,y}^2 + p_{\nu,z}^2} \right]^2 - [\vec{p}_\ell + (p_{\nu,x}, p_{\nu,y}, p_{\nu,z})]^2. \quad (3.2)$$

This returns two solutions for $p_{\nu,z}$, and therefore two solutions for \vec{p}_W ; the solution with smaller $|p_{\nu,z}|$ is selected. In the event that the $p_{\nu,z}$ solutions obtained are imaginary numbers (a scenario that occurs roughly 50% of the time), a constraint is applied requiring

$$m_T = m_W \quad (3.3)$$

where m_T is the total transverse mass of the W candidate. This forces $p_{\nu,z}$ to assume a real value. In addition to this constraint, the values of $p_{\nu,x}$ and $p_{\nu,y}$ are chosen so that the Euclidean distance from $(p_{\text{MET},x}, p_{\text{MET},y})$ is minimized.

Once a leptonic W candidate has been constructed, the BDT score is evaluated for every (W, jet) combination in a similar manner as the fully-hadronic scenario. To handle this correctly, the pseudo-continuous b-tag scores for the W candidates are set to the minimum value and the angle between the W daughters is set to 0. The top is then reconstructed from the leptonic W and the jet giving the highest BDT score.

In single-lepton events with four or more jets, a second top is reconstructed by selecting the jet triplet with highest BDT score, excluding the jet associated to the primary top. In events with two or three jets, the second top is taken as the sum of the jets left over after reconstructing the primary top. In events with exactly one jet, only the primary top is reconstructed.

No top reconstruction is performed in dileptonic events since using E_T^{miss} to solve for the kinematics of two neutrinos cannot easily be done.

Retraining with PFlow jets Two distinct trainings of the top reconstruction BDT were performed. In the CP Analysis, the BDT was trained using EMTopo jets, while for the Coupling analysis, PFlowJets are used. Figure ?? reports the primary top four-momenta components reconstructed using the same set of PFlow jets for both trainings. From these results, we conclude that the re-training provides a better description of the truth level distribution for PFlow jets, but that in both cases, the BDT trainings accurately model the truth top kinematics distributions.

An alternative likelihood-based top reconstruction method is explored in Appendix ??.

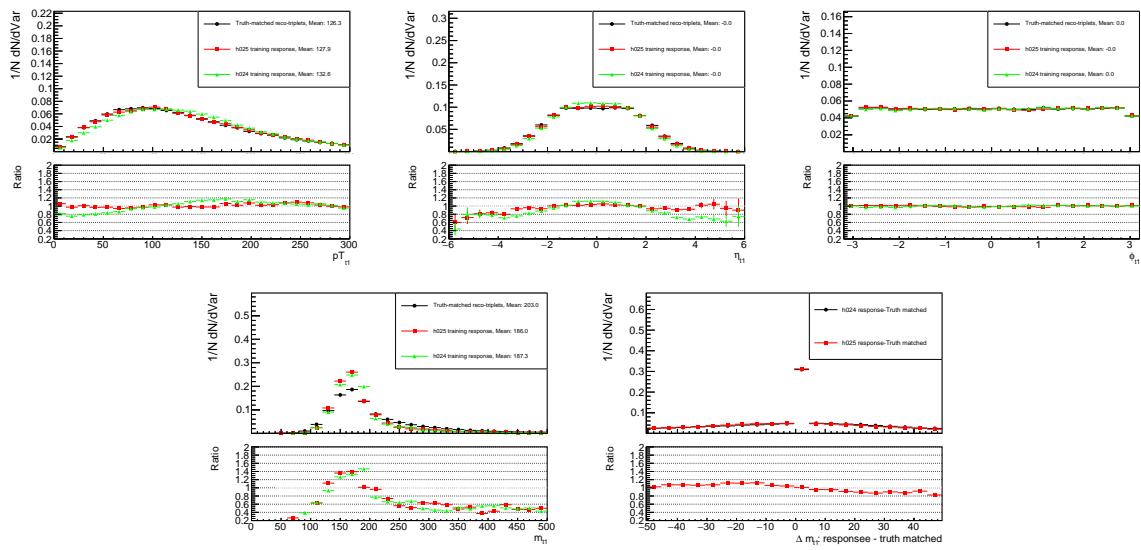


Figure 3.3: Performance of the top reconstruction BDT for the primary top in $t\bar{t}H$ events. The green "h025" line represents the EMTopo training applied to PFlow reconstructed jets, while the red "h024" line represents the output of the dedicated PFlow training with PFlow jets. The black line represents the truth-matched reco level distribution.

CHAPTER 4

Data and Monte Carlo Samples

4.1 Data, Monte Carlo, and HGam Pre-selection

4.1.1 Data

In both Run-2 HGam analyses discussed in this dissertation, we use the full LHC Run-2 dataset, consisting of $139.0 \pm 2.4 \text{ fb}^{-1}$ of proton-proton collisions with a center of mass energy $\sqrt{s} = 13 \text{ TeV}$ recorded by the ATLAS detector between 2015 and 2018 ???. Figure ?? shows the instantaneous luminosity gathered by the detector as a function of time; figure ?? shows the number of pileup events per bunch-crossing as a function of time.

The trigger used to select events is the diphoton trigger HLT_g35_loose_g25_loose (for 2015-2016 data) and HLT_g35_medium_g25_medium (for 2017-2018 data). Both triggers require two photon candidates, one with a transverse energy E_T of at least 35 GeV and the other with transverse energy of at least 25 GeV. The 2015-2016 trigger requires two photons that pass the "loose" ID requirement, while the 2017-2018 trigger requires two photons that pass the "medium" ID requirement (the cut was tightened due to increased luminosity and pileup). The trigger was calibrated

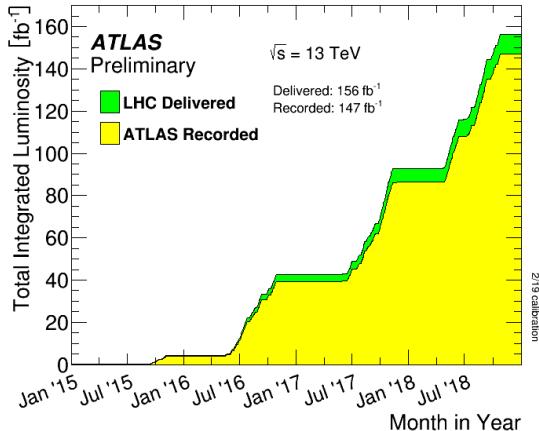


Figure 4.1: Integrated luminosity for the Run-2 ATLAS data-taking period.

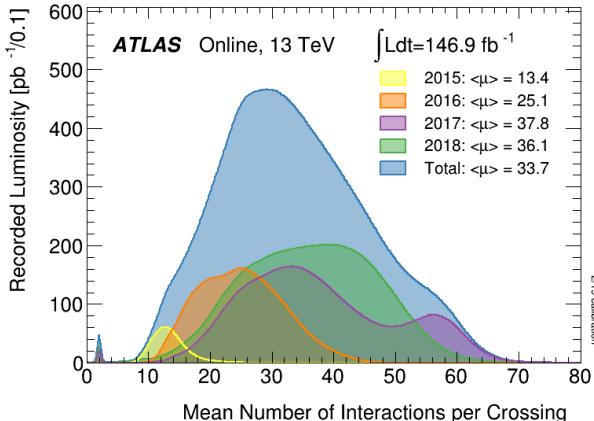


Figure 4.2: Pileup for the Run-2 ATLAS data-taking period.

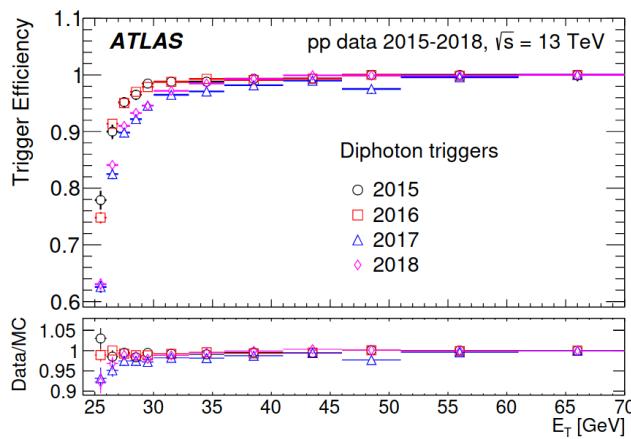


Figure 4.3: Efficiency of the trigger for the different years of the run-2 data taking period as a function of subleading photon E_T .

using radiative Z decays, and is observed to be greater than 95% efficient for each photon as long as it is 5 GeV above the trigger threshold ??.

4.1.2 Higgs Preselection and Data CRs

In both analyses discussed in this dissertation, we use the following preselection to define the $H \rightarrow \gamma\gamma$ signal region. These requirements are applied for both data and Monte Carlo simulation.

- *$\gamma\gamma$ Preselection:* The event is required to contain two photons passing the loose isolation and ID requirements.
- *PV:* At least one primary vertex is required to be identified in the event.
- *Trigger – Matching:* The leading two photons observed are required to match those identified by the trigger.

- *Relative p_T* : The leading and subleading photons are required to have $p_T/m_{\gamma\gamma}$ larger than 0.35 and 0.25, respectively.
- *TI*: For events in the signal-region, we require both photons to pass the "Tight" Photon ID and Isolation requirements (for events in the 'NTI' data control region, this selection criterion is ignored).
- *InvariantMass*: The invariant mass of the diphoton system must satisfy $105\text{GeV} < m_{\gamma\gamma} < 160\text{GeV}$

In order to accurately model the continuum diphoton background, it is often useful to invert the *TI* requirement- that is, we construct a data sample consisting of those events passing all other preselections, but for which one or more photons does not pass either or both of the "tight" isolation and ID requirements. Data control samples in this not-tight-isolated "*NTI*" region allow us to model the kinematic properties of objects in the event other than the diphoton system, such as top quark jet variables, with some amount of accuracy.

Additionally, it is useful to consider the *TI* or *NTI* data sidebands- this consists of the set of events passing either *TI* or *NTI* selection in the region $105\text{GeV} < m_{\gamma\gamma} < 160\text{GeV}$ but outside the signal region ($123\text{GeV} < m_{\gamma\gamma} < 127\text{GeV}$ for the CP analysis, $120\text{GeV} < m_{\gamma\gamma} < 130\text{GeV}$ for the Couplings analysis). This allows us to model the shape of the background while safely away from the Higgs signal peak near 125 GeV. This is detailed further in the subsequent analysis-specific chapters.

4.1.3 Nominal and Alternative Monte Carlo Samples

In order to generate Monte Carlo simulated events, several software programs must work in concert. In a real physics event, different proton constituents (usually gluons or quarks, often called "partons"), interact with one another in a process called a hard scatter to produce new particles.

To simulate such a collision, the quantum field-theoretic "matrix element" that governs the dynamics of the hard-scatter process must be calculated, the momentum distributions of the individual partons must be modelled using Parton Distribution Functions (or PDFs), and the decay showers of any final-state hadrons produced in the event must be dynamically modelled. Additionally, the dynamics of proton constituents not contributing to the hard scatter, called the "underlying event", must also be simulated. Because showering and underlying event processes are difficult to accurately model, they are "tuned" on the data using a variety of particular parameter sets ??.

The Monte Carlo samples used in this dissertation are constructed using either the *MadGraph5_aMC@NLO* ?? ?? or Powheg 7.03 ?? ?? ?? generators for matrix element calculation, *Pythia8v2* ?? ?? ?? or *Herwig7v2.6.2* ?? ?? for showering, *EvtGen* 1.1 for hadronization of final-state jets ??, and *Sherpa*

1.2.0 ?? ?? ?? ?? ?? for the matrix element calculation, showering, and hadronization of certain high-statistics samples.

After generation, Monte Carlo simulations are then passed through the Geant4 software package ?? to simulate the detector response. After this, the events are passed through the same object-reconstruction software used for data events ?. All Monte Carlo samples used in this dissertation are generated using the Geant4 full simulation, with the exception of the QCD continuum diphoton samples used for background modelling, which were treated using the GEANT4 fast simulation settings due to the number of events generated. For all MC samples, pile-up interactions were simulated by overlaying each Monte Carlo event with a different number of minimum-bias events simulated using Pythia 8.186 ?? with the ATLAS "A3" tune ??.

In all Monte Carlo simulations, the Higgs mass was set to 125 GeV and the decay width was set to 4.07MeV ?. All samples described include the small contribution from Dalitz decays (that is, decays in which one photon converts into two real final-state leptons), which is accounted for in the normalization of the samples. For all Monte Carlo samples generated using Madgraph5_aMC@NLO, the renormalization (μ_R) and factorization (μ_F) scales are defined as the scalar sum of the transverse masses of all final-state particles divided by two (i.e. $H_T/2$), and the top and W boson decays are handled by MadSpin ?? in order to ensure the correct treatment of the spin correlations of the decay products.

For all simulated samples, several corrections are applied to ensure the simulated samples correspond to data, including a beam-spot width correction, photon shower-shape and calorimetric isolation corrections, energy scale and resolution corrections, photon identification and isolation corrections, jet selection efficiency corrections, and electron and muon identification, reconstruction and isolation corrections.

Prod. Mode	Generator	PDF (Matrix Element)	PDF+Tune (Parton Shower)
ggF	POWHEG NNLOPS + PYTHIA	PDF4LHC15	AZNLOCTEQ6
VBF	POWHEG+PYTHIA	PDF4LHC15	AZNLOCTEQ6
W^+H	POWHEG+PYTHIA	PDF4LHC15	AZNLOCTEQ6
W^-H	POWHEG+PYTHIA	PDF4LHC15	AZNLOCTEQ6
$qq \rightarrow ZH$	POWHEG+PYTHIA	PDF4LHC15	AZNLOCTEQ6
$gg \rightarrow ZH$	POWHEG+PYTHIA	PDF4LHC15	AZNLOCTEQ6
b \bar{b} H	POWHEG+PYTHIA	PDF4LHC15	A14NNPDF23
t \bar{t} H	POWHEG+PYTHIA	PDF4LHC15	A14NNPDF23
t \bar{t} H	MADGRAPH5_aMC@NLO + PYTHIA	NNPDF3.0	A14NNPDF23
tHbj	MADGRAPH5_aMC@NLO + PYTHIA	NNPDF3.0	A14NNPDF23
tHW	MADGRAPH5_aMC@NLO + PYTHIA	NNPDF3.0	A14NNPDF23

Table 4.1: Summary of nominal signal samples

Nominal gluon-gluon fusion events are simulated using the Powheg NNLOPS tool ?? using the PDF4LHC15 PDF set ??, while the nominal VBF , WH , ZH , and $ggZH$ are all generated using

the Powheg generator with the PDF4LHC15 PDF set. Showering is performed with Pythia using the AZNLO tune ?? and the CTEQ6 PDF set ?? for these samples. Nominal Standard-Model tH and bbH samples are generated using the PowhegBOX tool ?? using the PDF4LHC15 PDF set; for these samples, showering is performed with Pythia using the A14 tune ?? and the NNPDF23 PDF set ?. Nominal tWH and $tHjb$ samples are generated using the Madgraph5_aMC@NLO generator with the NNPDF30 PDF Set ??, and are showered using Pythia with the A14 tune and the NNPDF23 PDF set. tWH samples are generated using the five-flavor PDF scheme, while $tHjb$ samples are generated using the four-flavor PDF scheme. At first order, NLO corrections to the tWH process contain final-state bottom quark jets that thus lead to interference with the more-common ttH process ??; to resolve this, a diagram subtraction method is used ?. Additionally, a ttH sample is generated using the Madgraph5_aMC@NLO generator with the NNPDF30 PDF Set ?? for use with the dedicated $ttHCP$ samples, in order to confirm the validity of the effective field theory model used to generate them.

Alternative Standard-Model samples are developed using Herwig for showering rather than Pythia, in order to evaluate the parton-showering uncertainty. The matrix element generators and PDFs are all the same as for the nominal samples, with the exception of the VBF sample, for which generator weights were unable to be included using the Herwig showering tool (and thus the NNPDF30 PDF is quoted for both the generator and the showering tool). ggF , VBF , VH , tWH , $tHjb$ and the supplementary Madgraph5_aMC@NLO ttH sample are all showered using the H7P1 Tune ??, and the Powheg ttH sample is showered using the ?? tune.

Additionally, a supplementary ggF sample is generated using Madgraph5_aMC@NLO containing an additional two partons at the matrix element stage. To avoid double-counting events with additional final-state partons, the "FxFx" merging scheme is employed. This sample is showered with Pythia using the NNPDF30 PDF and the A14 tune.

Prod. Mode	Generator	PDF (Matrix Element)	PDF+Tune (Parton Shower)
ggF	POWHEG NNLOPS + HERWIG	PDF4LHC15	NNPDF3.0 +H7p1
VBF	POWHEG + HERWIG	NNPDF3.0	NNPDF3.0 +H7p1
W^+H	POWHEG + HERWIG	PDF4LHC15	NNPDF3.0 +H7p1
W^-H	POWHEG + HERWIG	PDF4LHC15	NNPDF3.0 +H7p1
$qq \rightarrow ZH$	POWHEG + HERWIG	PDF4LHC15	NNPDF3.0 +H7p1
$gg \rightarrow ZH$	POWHEG + HERWIG	PDF4LHC15	NNPDF3.0 +H7p1
$t\bar{t}H$	POWHEG + HERWIG	PDF4LHC15	NNPDF3.0 +H7UE
$t\bar{t}H$	MADGRAPH5_aMC@NLO + HERWIG	NNPDF3.0	NNPDF3.0 +H7p1
$tHbj$	MADGRAPH5_aMC@NLO + HERWIG	NNPDF3.0	NNPDF3.0 +H7p1
tHW	MADGRAPH5_aMC@NLO + HERWIG	NNPDF3.0	NNPDF3.0 +H7p1
ggF	MADGRAPH5_aMC@NLO + PYTHIA	NNPDF3.0	NNPDF3.0 +A14

Table 4.2: Summary of alternative signal samples

The cross-sections for all Higgs processes are normalized to the state-of-the-art precision mea-

surements discussed in ?? using K-factor scaling factors. The accuracy of the cross-sections is N3LO QCD + NLO Electroweak for ggF , NNLO QCD + NLO Electroweak for VBF and VH , and NLO QCD + NLO Electroweak for tH , while $tHjb$ and tWH are produced at NLO QCD with no electroweak correction. Additionally, samples are normalized to account for the $H \rightarrow \gamma\gamma$ branching ratio of 2.270×10^{-3} calculated with HDECAY ?? and PROPHECY4F ??.

Prod. Mode	XSxBR [pb]
ggF	0.1101404
VBF	0.00857833
W^+H	0.00190226
W^-H	0.00120605
$qq \rightarrow ZH$	0.001724519
$gg \rightarrow ZH$	0.000278529
tH	0.001149755
$tHbj$	0.00016857
tHW	3.44359e-05
$b\bar{b}H$	0.00110390

Table 4.3: Cross sections times branching ratio values used to normalize each production mode. The values correspond to the state-of-the-art predictions and are taken from the CERN Yellow Report ??.

The background for $H \rightarrow \gamma\gamma$ events is generally treated as a smoothly-falling, continuous distribution in diphoton mass $m_{\gamma\gamma}$. It is comprised of non-Higgs events containing real final-state or initial-state radiated photons, as well as hadronic jets that behave like photons. To model this, a functional-form based data-driven method is used; however, Monte Carlo templates are nonetheless used to facilitate and validate the background modelling.

Both QCD continuum $\gamma\gamma$ +jets production and the various $V\gamma\gamma \rightarrow ll\gamma\gamma$ samples are simulated at leading order in QCD using the Sherpa event generator with the CT10 PDF set ?. Showering is performed using the default Sherpa showering tool.

$t\bar{t}\gamma\gamma$, one of the leading backgrounds in top-enriched regions, is simulated using Madgraph5_aMC@NLO with the NNPDF30 PDF set. It is showered using Pythia and tuned using the A14 tune.

Prod. Mode	Generator	PDF (Matrix Element)	Showering Tool	PDF+Tune (Parton Shower)
$\gamma\gamma+0,1(\text{NLO}),2,3(\text{LO}), m_{\gamma\gamma} \in 50-90 \text{ GeV}$	SHERPA	CT10	SHERPA	CT10
$\gamma\gamma+0,1(\text{NLO}),2,3(\text{LO}), m_{\gamma\gamma} \in 90-175 \text{ GeV}$	SHERPA	CT10	SHERPA	CT10
$ee\gamma\gamma; m_{\gamma\gamma} > 80 \text{ GeV}$	SHERPA	CT10	SHERPA	CT10
$\mu\mu\gamma\gamma; m_{\gamma\gamma} > 80 \text{ GeV}$	SHERPA	CT10	SHERPA	CT10
$\tau\tau\gamma\gamma; m_{\gamma\gamma} > 80 \text{ GeV}$	SHERPA	CT10	SHERPA	CT10
$\nu\nu\gamma\gamma; m_{\gamma\gamma} > 80 \text{ GeV}$	SHERPA	CT10	SHERPA	CT10
$ev\gamma\gamma; m_{\gamma\gamma} > 80 \text{ GeV}$	SHERPA	CT10	SHERPA	CT10
$\mu\nu\gamma\gamma; m_{\gamma\gamma} > 80 \text{ GeV}$	SHERPA	CT10	SHERPA	CT10
$\tau\nu\gamma\gamma; m_{\gamma\gamma} > 80 \text{ GeV}$	SHERPA	CT10	SHERPA	CT10
$t\bar{t}\gamma\gamma \text{ (noallhad)}$	MADGRAPH5_aMC@NLO	NNPDF3.0	PYTHIA	NNPDF3.0 +A14
$t\bar{t}\gamma\gamma \text{ (allhad)}$	MADGRAPH5_aMC@NLO	NNPDF3.0	PYTHIA	NNPDF3.0 +A14

Table 4.4: Summary of nominal background samples

4.1.4 ttHCP Monte Carlo Samples

In the $ttHCP$ analysis, an Effective Field Theory (EFT) setting a cutoff scale of 1 TeV, below which no new BSM particles coupling to the Higgs exist, is used to generate Monte Carlo samples. The EFT used is the Higgs Characterization (HC) model ??, implemented in the MadGraph5_aMC@NLO generator with the NNPDF30 PDF set. In all BSM samples generated, we use Pythia for showering with the NNPDF23 PDF set and the A14 PDF set.

As previously mentioned, the top-Higgs interaction term of the Lagrangian in the presence of CP-violation can be parameterized as

$$\mathcal{L} = \kappa_t g_t \bar{t}(cos(\alpha) + sin(\alpha)i\gamma^5)th \quad (4.1)$$

where $g_t = \frac{-m_t}{v} = \frac{-173.26\text{GeV}}{246\text{GeV}} = -0.703$, κ_t is the dimensionless coupling-strength term ($\kappa_t = 1$ in the Standard Model), and α is an angle that parameterizes the CP-mixing strength ($\alpha = 0$ in the Standard Model, $\alpha = \frac{\pi}{2}$ in the fully CP-odd case). The interpretation of the $H \rightarrow \gamma\gamma$ and ggF dependence on α is handled in several different ways, as discussed in ??.

As in the nominal case, $t\bar{t}H$ and tWH samples are generated using the five-flavor scheme, while the four-flavor scheme is used for the $tHjb$ process. The Standard Model cross-sections and branching ratios for all process are normalized to those given in the CERN Higgs Yellow Report 4 ??, in which fixed scales and the five-flavor scheme are used. Those cross-sections are calculated at NLO QCD accuracy (without electroweak correction) for the $tHjb$ and tWH processes, while $t\bar{t}H$ is calculated at both NLO QCD and NLO Electroweak accuracies. $ggF + 2jets$ samples are also generated using the MadGraph5_aMC@NLO generator as the ggF cross-section varies with α and κ_t . To ensure that the W-boson Higgs Coupling is kept constant, tWH samples are generated by treating $\cos\alpha \times \kappa_{SM}=1$.

K-factors are then computed to scale the Higgs Characterization Monte Carlo cross-sections to the Yellow Report cross-sections. The obtained K-factors are shown to be similar for different CP mixing angles; thus, the K-factors derived for the SM case can be safely used for the various samples with different α values.

κ_t	α	$\cos\alpha$	κ_{SM}
1	0(SM, CP-even)	1	1
1	15	0.965926	1.035276
1	30	0.866025	1.154701
1	45 (CP max mixing)	0.707107	1.414214
1	60	0.5	2
1	75	0.258819	3.863703
1	90 (CP-odd)	0.000001	10^6
-1	0	1	1
0.5	0	1	1
2	0	1	1
2	45	0.707107	1.414214

Table 4.5: Parameters used in the HC model in order to allow for CP-couplings only in the top quark sector while fixing the HWW coupling to its SM value by imposing $\cos\alpha \kappa_{SM}=1$. In the first set of samples, κ_t is fixed to 1 and α is varied. In the second set, κ_t is set to values different from 1. For $\alpha = 90$; due to numerical precision, $\cos\alpha$ strictly equal to 0 can't be generated and thus a value approaching it (10^{-6}) and a defined value for κ_{SM} (10^6) are used.

Cross-section (pb)					
κ_t	α	ttH	tHjb	tWH	ggF
1	0(SM, CP-even)	0.458	0.0606	0.0167	14.1
1	15	0.443	0.0636	0.0184	–
1	30	0.396	0.0743	0.0231	–
1	45 (CP max mixing)	0.329	0.0958	0.0308	24.0
1	60	0.265	0.1374	0.0422	–
1	75	0.217	0.1970	0.0563	–
1	90 (CP-odd)	0.199	0.2707	0.0726	32.4
-1	0	–	0.6971	0.1490	–
0.5	0	–	0.0939	0.0148	–
2	0	–	0.2406	0.0924	–
2	45	–	0.2055	0.1095	–

Table 4.6: NLO cross-sections for the $t\bar{t}H$, $tHjb$, tWH , and ggF processes for different CP-scenarios (see parameters in Table ??). The top part refers to MC samples in which the κ_t value is fixed to 1 and α is varied, while the bottom part refers to samples with κ_t different from 1.

Normalized Cross-section(fb)						
κ_t	α	ttH	tHjb	tWH	ggF	
1	0(SM, CP-even)	1.150	0.169	0.034	33.1	
1	15	1.113	0.177	0.038	–	
1	30	0.995	0.207	0.048	–	
1	45 (CP max mixing)	0.827	0.266	0.064	53.7	
1	60	0.666	0.382	0.087	–	
1	75	0.545	0.548	0.116	–	
1	90 (CP-odd)	0.500	0.753	0.150	74.4	
-1	0	–	1.980	0.307	–	
0.5	0	–	0.264	0.030	–	
2	0	–	0.666	0.190	–	
2	45	–	0.570	0.226	–	

Table 4.7: NLO cross-sections for the $t\bar{t}H$, $tHjb$, tWH , and ggF processes for different CP-scenarios normalized with the K-factor and $BR(H \rightarrow \gamma\gamma)$. The top part refers to MC samples in which the κ_t value is fixed to 1 and α is varied, while the bottom part refers to samples with κ_t different from 1.

CHAPTER 5

Signal Parameterization, Background Parametrization, and Statistical Methods

In the analyses discussed in this dissertation, smooth functional forms are used to model both the signal and background. This allows an unbinned likelihood fit to be performed, the preferable statistical analysis method in the lower-statistics $t\bar{t}H$ regime (an alternative to the binned likelihood fit, also common in ATLAS analyses, which offers more statistics per bin, but a general loss of precision).

Additionally, parameterizing signal and background in this way allows for the relatively straightforward calculation of several key systematics, including "spurious signal" background mismodelling, photon energy scale, and photon energy resolution.

5.1 Signal Modelling

In collider physics analyses, one of the most common forms of signal is a "resonance", a bump in a smooth energy spectrum indicating the presence of a particle with mass given by the center of the resonance bump and lifetime given by the width Γ of the resonance bump ???. The "true underlying form" of resonances generally follow the Breit-Wigner distribution described in ??; however, due to detector and beam effects, this form does not accurately describe observed data.

Instead, for both analyses discussed in this dissertation, a "Double-Sided Crystal Ball" (DSCB) function ???? is used. The function has six parameters, two that describe the shape of its Gaussian core μ_{CB} , and σ_{CB} , and two that describe the shape of each of its exponential tails: α_{low} and n_{low} ; α_{high} and n_{high} . The function is defined as:

$$f_{DSCB}(m_{\gamma\gamma}) = \begin{cases} e^{\frac{-\alpha_{low}^2}{2}} \left(\frac{R_{low} - \alpha_{low} - t}{R_{low}} \right)^{n_{low}} & if t < -\alpha_{low} \\ e^{\frac{-t^2}{2}} & if -\alpha_{low} \leq t \leq \alpha_{high} \\ e^{\frac{-\alpha_{high}^2}{2}} \left(\frac{R_{high} - \alpha_{high} + t}{R_{high}} \right)^{n_{high}} & if t > \alpha_{high} \end{cases}$$

where $t = \frac{(m_{\gamma\gamma} - \mu_{CB})}{\sigma_{CB}}$ and $R = \text{frac}\alpha$.

To parameterize the signal in each analysis category for both analyses discussed in this dissertation, all Monte Carlo for the various Higgs production modes (VBF , VH , ggF , tH , tWH , $tHjb$ and bbH) generated using a Higgs mass of 125 GeV are combined. The resulting distribution is then fit with a DSCB function, and a rigid transformation of 0.09 GeV is performed such that the mean of the fitted DSCB corresponds to the experimentally measured Higgs mass of $125.09\text{GeV}(\text{stat}) \pm 0.21\text{GeV}(\text{syst})$??.

Because the Double-Sided Crystal Ball function depends strongly on the photon resolution and energy scale, these systematics can be parameterized in the final fit as variations in the DSCB parameters ???. Examples of DSCB shapes in two categories of the CP Analysis are shown in figure XXX.

5.2 Background Modelling and Spurious Signal

Like the signal, the background is also parameterized as a smoothly-falling functional form in each category. This is done in a data-driven manner: first, a functional form is chosen using simulation-derived (or NTI-derived) templates to minimize the "spurious signal" systematic uncertainty. The background normalization and parameters of this functional form are then allowed to float in the final fit to the data, i.e., they are not fixed as a result of the spurious signal test.

Background templates for the spurious signal test are constructed from Monte Carlo or NTI data to resemble the expected TI data as closely as possible in each category. The construction of the templates is detailed further in section .

The spurious signal test is a signal-plus-background functional fit to a background-only distribution. This provides a relatively simple way to measure background mismodelling- the background template contains no true Higgs signal, so the functional form that best models the background is the one for which the extracted "spurious" signal is closest to zero. This is illustrated in figure ???. Spurious signal can be positive or negative- if positive, the functional form chosen "undershoots" the true background, while if negative, the functional form chosen "overshoots" the true background.

The spurious signal fit is performed for signal masses from 121 GeV to 129 GeV at intervals of 1 GeV; the number of spurious signal events N_{sp} is defined as the maximum of the absolute value of the number of signal events extracted from this range. The functional form chosen is one of the following functional forms:

- Exponential Function: $f(m_{\gamma\gamma}) = e^{c \cdot m_{\gamma\gamma}}$
- Exponential Function of 2nd Order Polynomial: $f(m_{\gamma\gamma}) = e^{c_1 \cdot m_{\gamma\gamma}^2 + c_2 \cdot m_{\gamma\gamma}}$

- Bernstein polynomial of order N : $B_N(m_{\gamma\gamma}) = \sum_{i=0}^N c_i \cdot b_{i,N}$ with $b_{i,N} = \binom{N}{i} m_{\gamma\gamma}^i (1 - m_{\gamma\gamma})^{N-i}$
- First-Order Power Law Function: $f(m_{\gamma\gamma}) = m_{\gamma\gamma}^c$

Functional forms chosen are then required to satisfy one of the two following criteria:

- $N_s < 0.1 \times N_{s,exp}$, where $N_{s,exp}$ is the expected true signal in a given category
- $N_s < 0.2 \times \sigma_{s,exp}$, where $\sigma_{s,exp}$ is the statistical uncertainty on the expected true signal in a given category

If more than one function passes these criteria, the function with the lower number of degrees of freedom is selected. If there are multiple functions that pass the criteria and have the same number of degrees of freedom, the function with the lowest resulting spurious signal is selected.

In low statistics categories, it is not uncommon that no functional form will satisfy the above criteria. In this case, the "relaxed" spurious signal fit is performed, which replaces N_s with a new variable ζ_s that is designed to accommodate up to 2σ fluctuations in the spurious signal:

$$\zeta_s = \begin{cases} N_s + 2\delta_{MC} & \text{if } N_s + 2\delta_{MC} < 0 \\ N_s - 2\delta_{MC} & \text{if } N_s - 2\delta_{MC} > 0 \\ 0 & \text{otherwise} \end{cases}$$

Though ζ_s is used to define the spurious signal criteria, N_s is still used for the final spurious signal uncertainty.

An additional requirement that the χ^2 probability of the signal-plus-background fit in the spurious signal test is greater than 1% is applied in the Couplings analysis; while this is not a requirement in the CP analysis, all spurious signal fits nonetheless satisfy this criterion as well.

In the Couplings analysis, several of the very low-stat categories nevertheless produce unphysical fits using these criteria. Thus, in categories of this analysis containing fewer than 100 events in the data sidebands, a Wald test is performed ??: the functional forms are restricted to Exp, ExpPoly2, and ExpPoly3, and their respective maximum-likelihoods L_1 , L_2 , and L_3 are computed from a fit to the TI data sidebands. The quantities $q_{ij} = -2\ln(\frac{L_i}{L_j})$ are then calculated and converted into p-values, assuming that q_{ij} follows a χ^2 distribution with one degree of freedom. If the p-value calculated is ≤ 0.05 , the function with the larger number of degrees of freedom is chosen.

An illustration of the Wald test in one of the Couplings categories is shown in figure XXX.

These failed fits show the limitations of current spurious signal criteria: in low-statistics categories, the Monte Carlo templates often contain large statistical fluctuations that can be fit as

spurious signal. However, because spurious signal is intended to measure the mismodelling due to the choice of functional form, the presence of such fluctuations can artificially inflate the spurious signal. If the functional forms chosen because of these fluctuations have too many degrees of freedom, this can cause disastrous effects when fitting to the true data sidebands, as statistical fluctuations in data will dominate. This is the motivation for introducing Gaussian Process Regression (GPR) smoothing, discussed in appendix ???. At the time of this writing, this novel smoothing procedure is currently being implemented in the Coupling analysis; a discussion of the procedure and the improvements to the spurious signal that result are given in appendix ???, while extensive validation of the smoothing procedure using toy tests is given in appendix ??.

5.2.0.1 Background Templates

The background templates constructed for both analyses discussed in this dissertation are designed to approximate the continuum diphoton background.

In all categories of the CP analysis, either $t\bar{t}\gamma\gamma$ Monte Carlo or NTI data is used to model the continuum background, scaled to the TI data sidebands. The statistical uncertainty of both template candidates is checked in each region by taking the integral and sum of errors; the template chosen is the one with the smallest statistical uncertainty. $t\bar{t}\gamma\gamma$ is ultimately used to construct the template in all CP analysis categories but one.

Similarly, in the leptonic VH categories of the Coupling analysis, a combination of $\gamma\gamma + V\gamma\gamma$ Monte Carlo, scaled to match the TI sidebands, is used. However, in the other ggH and VBF categories, a more sophisticated data-driven method is used due to the nontrivial presence of fake photons arising from misidentified jets.

First, the purity fraction, i.e. the fraction of true vs. fake diphoton events in data, is calculated in each category. This is done using a double two-dimensional ABCD method ??.

For a given choice of photon (either leading or subleading) four regions are constructed in each category:

- Region A, in which the photon passes the Tight ID criterion and the Tight isolation criterion
- Region B, which the photon passes the Tight ID criterion and fails the Tight isolation criterion
- Region C, in which the photon fails the Tight ID criterion (but passes the LoosePrime3 ID criterion) and fail the Tight isolation criterion
- Region D, in which the photon fails the Tight ID criterion (but passes the LoosePrime3 ID criterion) and passes the Tight isolation criterion

Assuming that the ID and isolation cuts are independent with ϵ_{ID} and ϵ_{iso} , and that the photons in all regions but region A are definitively fakes, it is possible to define the number of fake photons in region A as:

$$N_A = \epsilon_{iso} N_B = \epsilon_{ID} N_D = \epsilon_{iso} \epsilon_{ID} N_C \quad (5.1)$$

Extending this to the diphoton pair, eight total regions can be defined in each category, allowing us to quantify the ($\gamma\gamma/\gamma j/jj$) fraction in each category. However, the assumption that the isolation and ID cuts are independent is not strictly valid, so it is not enough to perform a simple counting-experiment to determine the fraction of jets: various correlation terms must be considered, so a fit must be performed. The fit has sixteen equations with nineteen variables, and is performed following the process in ??.

After the purity fractions are calculated in each category, the shapes of the γj and jj distributions are derived using NTI data. The $\gamma\gamma$ Monte Carlo is then reweighted to match these shapes, and the $\gamma\gamma$, γj , and jj contributions are added according to their purities. As in other categories, the templates are scaled to the TI sidebands.

5.3 Likelihood Fitting and Asimov Data

The fits performed in all analyses discussed in this dissertation use a test statistic called a Profile Likelihood Ratio. The likelihood ratio is a Bayesian variable that is simply defined as $\mathcal{L}(\vec{x}, \vec{\theta}) = P(\vec{x}|vec{\theta})$ for a set of observed data points \vec{x} and a set of additional nuisance-parameter variables $\vec{\theta}$ such as background normalization, shape, and systematic uncertainties ??.

For independent, identically distributed Poisson variables (i.e., the number of signal and background events in a given histogram bin i), we can parameterize the likelihood as:

$$\mathcal{L}(\mu, N_b, \vec{\zeta}) = \prod_i^N \frac{(\mu \times s_i + N_b \times b_i(\vec{\zeta}))^{n_i} \times e^{-(\mu \times s_i + N_b \times b_i(\vec{\zeta}))}}{n_i!} \quad (5.2)$$

for signal strength μ , expected signal and background probability distributions s_i and b_i in each bin, nuisance parameters $\vec{\theta} = (N_b, \vec{\zeta})$ governing the total background normalization and shape, and n_i observed number of events in bin i. Likelihoods can also be constructed in an unbinned manner by taking the continuous form of s_i and b_i as a function of $m_{\gamma\gamma}$. For a multi-category analysis, the total likelihood is taken as the product of the likelihood in each category.

Using this, we can construct the Profile Likelihood Ratio as:

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\vec{\theta}}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\vec{\theta}})} \quad (5.3)$$

where $\hat{\theta}(\mu)$ is the set of values of $\vec{\theta}$ that conditionally maximize \mathcal{L} for each given μ , and $\hat{\theta}, \hat{\mu}$ are the values of $\vec{\theta}$ and μ that maximize \mathcal{L} globally.

By maximizing λ , or, more commonly, by minimizing $-2\ln(\lambda)$ (due to its comparative ease of computation) it is possible to extract a best-fit value for μ given a set of data. By definition, μ is the ratio of the observed signal yield to the one expected in the SM; however, it is also possible to parameterize μ in terms of other variables, such as the various κ couplings. This parameterization is performed in both analyses discussed in this dissertation, and is detailed further in ?? and ??.

Systematic uncertainties enter into the fit as nuisance parameters. In order to treat additional systematic uncertainties on the signal shape and yield in a manner similar to the treatment of the systematic uncertainties on the background, the Double-Sided Crystal Ball parameters μ_{CB} and σ_{CB} that enter into the parameterization of s_i are modified by energy scale and resolution shifts of magnitude σ and sign θ :

$$\mu_{CB} = \mu_{CB} \prod_k (1 + \sigma_{ES,k} \theta_{ES,k}) \quad (5.4)$$

$$\sigma_{CB} = \sigma_{CB} \prod_j (1 + \sigma_{ER,j} \theta_{ER,j}) \quad (5.5)$$

Similar response-term modifications can be performed to the luminosity and trigger uncertainties affecting overall signal yield. Additional constraint terms (given a Gaussian, Log-Normal, or Asymmetric ??) that are equivalent to the Bayesian prior distribution for each nuisance parameter are also multiplied to the total likelihood, in order to correctly model their allowed spread ??.

When performing a likelihood fit, it is often useful to construct a representative "Asimov" dataset generated by fixing both μ and particular nuisance parameters.

Pre-fit Asimov data is constructed by performing an unbinned, background-only profile likelihood ratio fit to the TI sidebands. The nuisance parameters are then set to the values extracted from this fit, and a Standard Model signal with significance $\mu = 1$ (that is, exactly corresponding to the predicted signal yield) is then superimposed atop the background shape. Similarly, after unblinding the signal region, post-fit Asimov data is constructed by fitting to the entire dataset, extracting the values for the nuisance parameters, and again superimposing a Standard Model signal with significance $\mu = 1$ atop the background. The reason for constructing both pre-fit and post-fit Asimov is to evaluate the "pull" of various nuisance parameters based on the data in the signal region: if the fit after including signal-region data causes a nuisance parameter to deviate substantially from the norm, it is possible that one or more systematics are not well-modelled.

The list of systematic uncertainties and the expected significances derived from pre- and post-fit Asimov data for both analyses discussed in this dissertation are given in Chapters ?? and ??.

CHAPTER 6

Study of the CP properties of the top-quark Yukawa interaction in $t\bar{t}H$ and tH events wth $H \rightarrow \gamma\gamma$

6.1 Categorization

In order to properly measure the dependence of the top Yukawa coupling on the CP-mixing angle α , a region of $t\bar{t}H$ -and- tH -enriched phase space is divided into a number of different categories based both upon the similarity of events it contains to signal (Higgs processes like $t\bar{t}H$ and tH) rather than background (non-Higgs continuum diphoton processes), as well as the similarity of events it contains to CP-odd rather than CP-even Higgs processes. By creating many such categories and fitting the event yield in each, detailed constraints can be set on the value of α .

First, two sets of regions are defined. The "hadronic" region targets events containing fully-hadronic top decays, requiring two loose-ID photons, one b-tagged jet at the 77% working point with $p_T > 25$ GeV, as well as two additional jets with $p_T > 25$ GeV and exactly zero electrons or muons. Similarly, the "leptonic" region targets events containing semi-leptonic top decays, requiring two loose-ID photons, one b-tagged jet at the 77% working point with $p_T > 25$ GeV, as well as at least one isolated electron or muon.

To perform the categorization in these regions, two multivariate Boosted Decision Trees (BDTs) are used, one to separate Higgs events from background and one to separate CP-odd Higgs events from CP-even Higgs events. Both BDTs are trained on low-level kinematic features using the XG-Boost package ?? . A series of cuts on these BDT scores are then defined, delineating a total of 20 orthogonal regions, each with differing sensitivity both to the $t\bar{t}H + tH$ signal and to the CP-odd $t\bar{t}H + tH$ hypothesis. An illustration of the categorization strategy is shown in figure ?? .

6.1.1 SBBDT

The signal-versus-background BDT (SBBDT) developed for use in the CP Analysis is identical to that developed first in the $79.8 fb^{-1}$ measurements of $t\bar{t}H$ in the diphoton channel ?? in and later

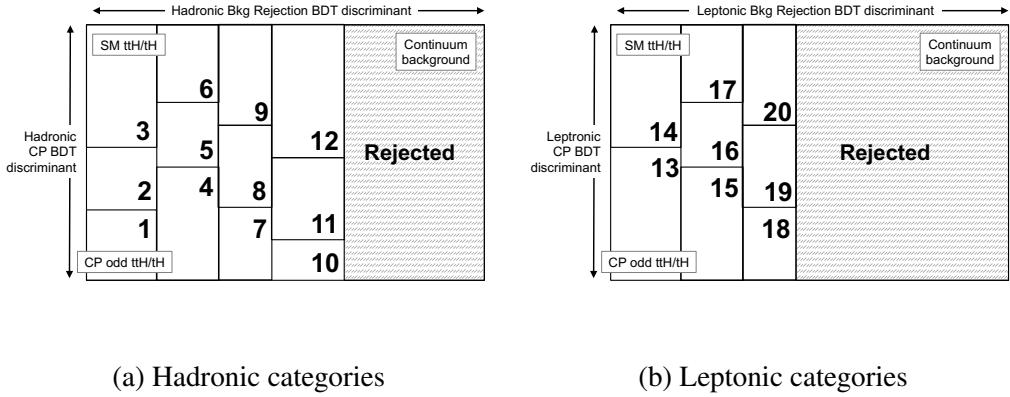


Figure 6.1: Diagram of the 2-dimensional categorization scheme in the hadronic (a) and leptonic (b) channels. The x -axis indicates the background-rejection BDT (SBBDT) score distribution, and the y -axis represents the CP-BDT score distribution. The shaded region indicates rejected events.

retrained for 139fb^{-1} measurements of $t\bar{t}H$ in the diphoton channel ???. It is trained separately for both the hadronic and leptonic regions.

Both the hadronic and leptonic BDTs are trained using a Standard-Model Powheg $t\bar{t}H$ Monte Carlo sample to model the signal and NTI data control events to model the continuum diphoton background.

6.1.1.1 Hadronic Region

In the hadronic region, 60% of the $t\bar{t}H$ Monte Carlo signal events are used for training, 20% are reserved for categorization and BDT hyperparameter optimization, and the final 20% are reserved for validation. 60% of the NTI events are used for training, 20% are reserved for hyperparameter optimization, and the remaining 20% are reserved for testing and significance evaluation.

The input variables chosen are:

- $p_T/m_{\gamma\gamma}$, η and ϕ of the two photons. Photon p_T is scaled by $m_{\gamma\gamma}$ to reduce unwanted sculpting of the diphoton mass spectrum.
- p_T , η , ϕ and E of the six jets with highest p_T
- Boolean b-tag flag (77% working point) for each of the six jets with highest p_T
- E_T^{miss} and direction of E_T^{miss}

Distributions of the BDT input variables using 79.8fb^{-1} of data are shown in Figure ??.

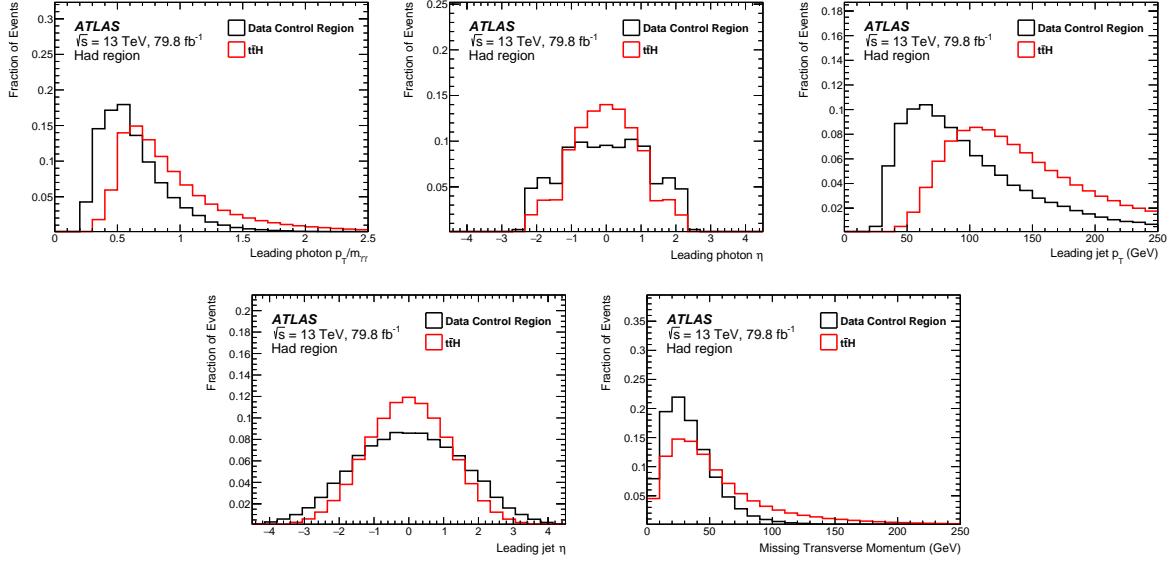


Figure 6.2: Distributions of training variables for the hadronic background-rejection BDT, trained at 79.8fb^{-1} . Taken from ??.

6.1.1.2 Leptonic Region

As in the hadronic region, in the leptonic region, 60% of the $t\bar{t}H$ Monte Carlo signal events are used for training, 20% are reserved for categorization and BDT hyperparameter optimization, and the final 20% are reserved for validation. 75% of the NTI events containing zero b-jets but at least one un-tagged jet were used for training and the remaining 25% are reserved for hyperparameter optimization, while 50% of the NTI events containing one or more b-jets are used for categorization and the remaining 50% are reserved for testing and significance evaluation.

However, due to lower statistics in the leptonic top decay channel due to the smaller top-quark branching ratio to leptons, two cuts are relaxed for the development of the leptonic BDT:

- The relative p_T cuts are loosened from $\frac{p_T}{m_{\gamma\gamma}} > 0.35$ for the leading photon and $\frac{p_T}{m_{\gamma\gamma}} > 0.25$ for the subleading photon to a flat $p_T > 35 \text{ GeV}$ for the leading photon and $p_T > 25 \text{ GeV}$ for the subleading photon.
- The diphoton invariant mass window is extended from $105\text{GeV} < m_{\gamma\gamma} < 160\text{GeV}$ to $80\text{GeV} < m_{\gamma\gamma} < 250\text{GeV}$.

The cuts are again tightened to define the signal region after BDT training is complete- that is, the loosening of these cuts is only utilized to increase BDT training statistics, and does not carry through to other stages of the analysis.

The input variables chosen are:

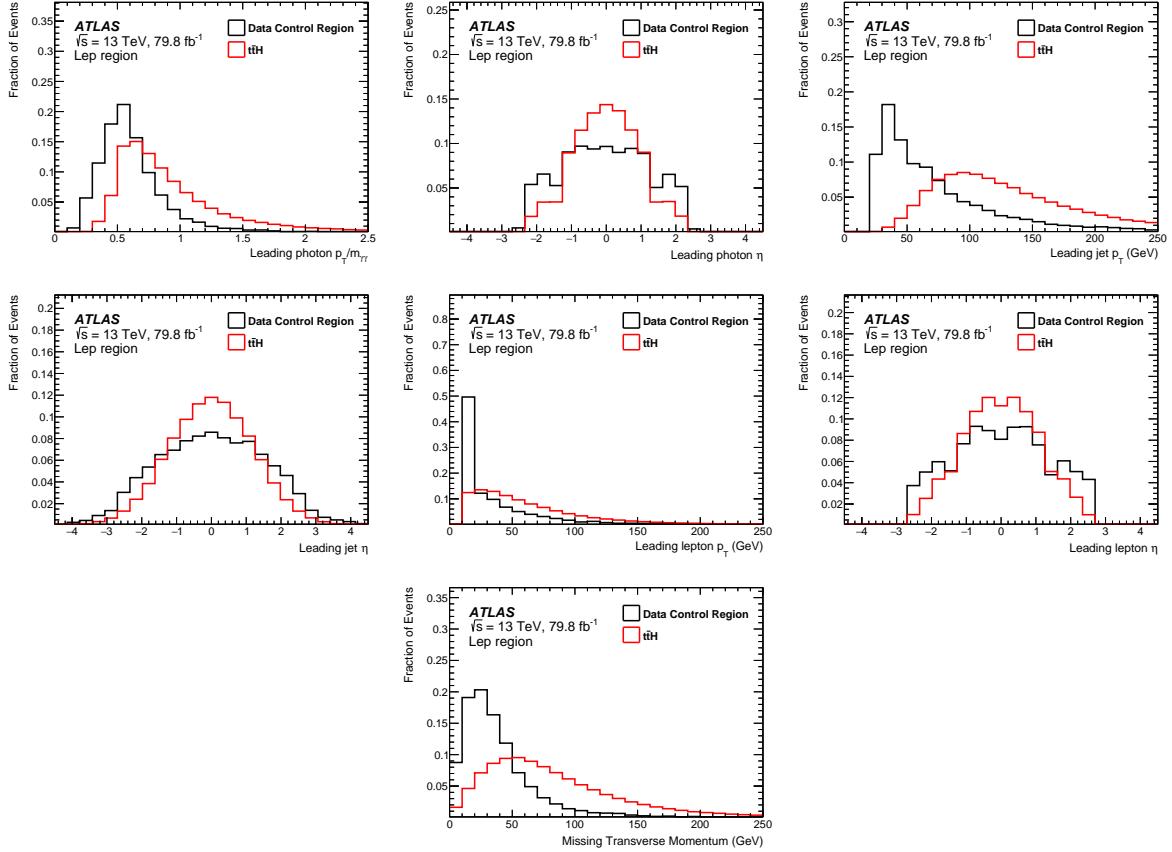


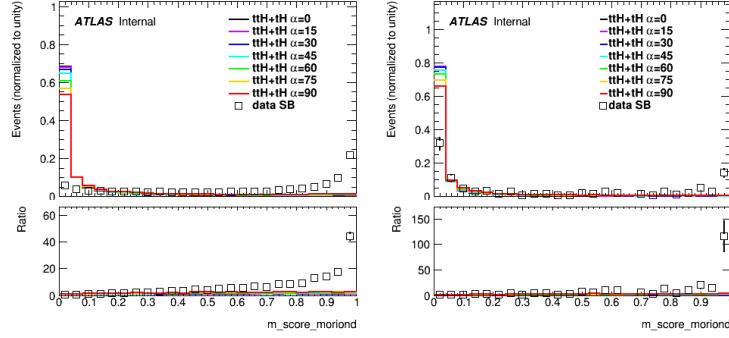
Figure 6.3: Distributions of training variables for the leptonic background-rejection BDT, trained at 79.8fb^{-1} . Taken from ??.

- $p_T/m_{\gamma\gamma}$, η and ϕ of the two photons. Photon p_T is scaled by $m_{\gamma\gamma}$ to reduce unwanted sculpting of the diphoton mass spectrum.
- p_T , η and ϕ of up to two leptons.
- p_T , η , ϕ and E of the four jets with highest p_T
- Boolean b-tag flag (77% working point) for each of the four jets with highest p_T
- E_T^{miss} and direction of E_T^{miss}

Distributions of the BDT input variables using 79.8fb^{-1} of data are shown in Figure ???. The BDT output distributions are shown in figures ???. The SBBDT performance is independent of α .

6.1.2 CP-Sensitive Observables

To construct a BDT to discriminate between CP-even and CP-odd $t\bar{t}H + tH$, a number of variables must be plotted at truth-level in order to determine their dependence on α . The HC model $t\bar{t}H$ and



(a) Hadronic $ttH + tHjb + tWH$
(b) Leptonic $ttH+tHjb+tWH$

Figure 6.4: SB BDT score for the sum of ttH , $tHjb$ and tWH , with relative weights according to the predicted cross sections. Shown in (a) for the hadronic channel and (b) for the leptonic channel, for various CP mixing scenarios. The open squares show data in the NTI sideband region, giving an approximation of the shape of the continuum background.

tH samples with alternative values of α generated using MadGraph5_aMCatNLO are used, added according to their calculated cross-sections given in table XXX.

From these plots, it is observed that the strongest variable is the Higgs boson p_T : CP-odd ttH and tH have a much more boosted Higgs than CP-even ttH and tH , and are more central in η . Similarly, the angular separation $\Delta\eta$ between the top and anti-top is much larger in CP-odd ttH , while the top and anti-top are more back-to-back in azimuthal angle $\Delta\phi$ in CP-even ttH than in CP-odd ttH . In $tHjb$ events, it is apparent that the top p_T and η also have discriminatory power. The mass of the Higgs + leading top system also offers discriminatory power- for ttH and tWH events it increases with α , while for $tHjb$ events it decreases with α .

These variables are shown in Figures

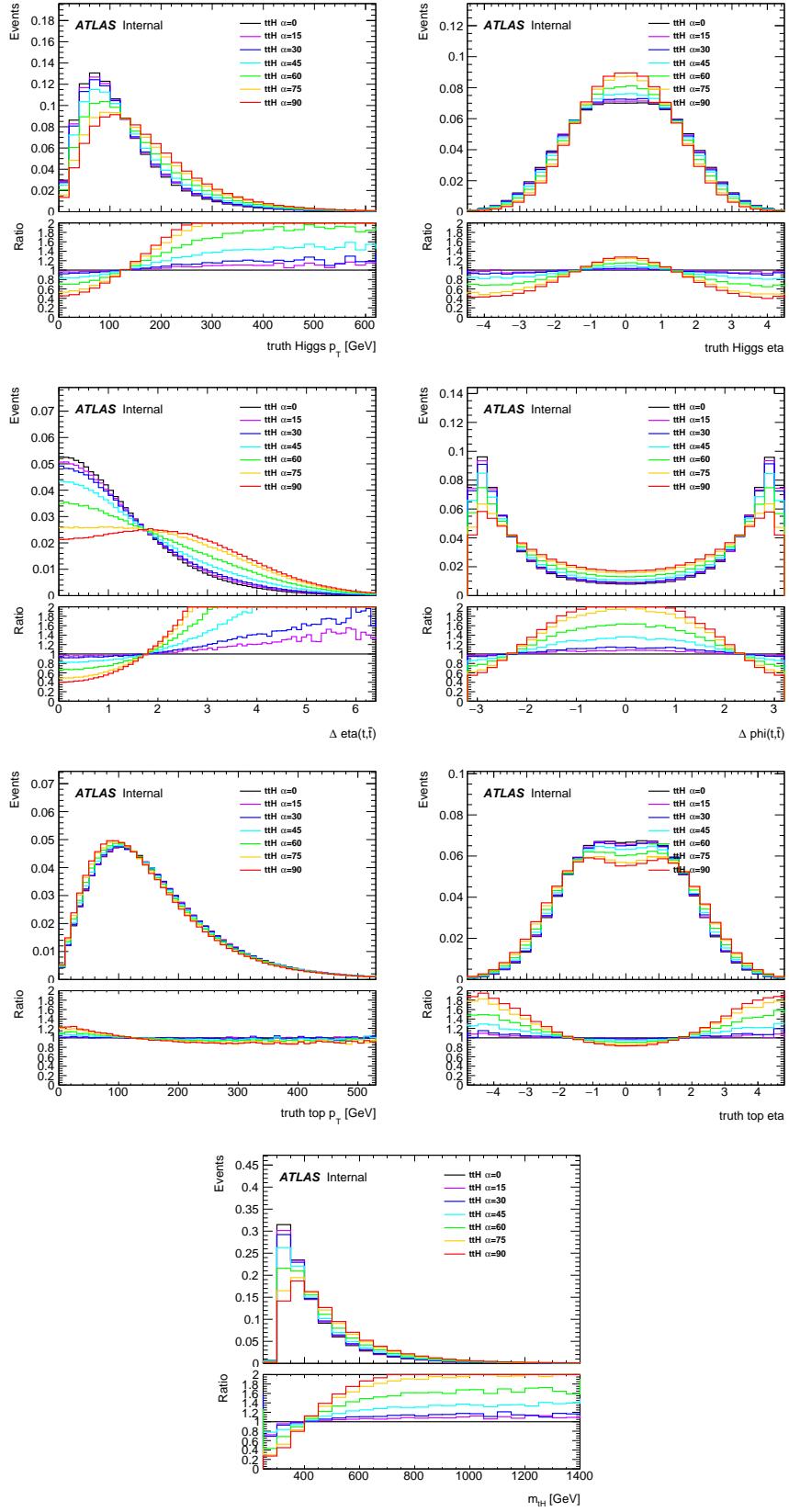


Figure 6.5: Truth-level distributions in $t\bar{t}H$ Monte Carlo of the Higgs boson p_T and η (top), angular separation between top and anti-top quarks (second row), top quark p_T and η (third row) and invariant mass of the top-Higgs system (bottom) for different values of the CP mixing angle α .

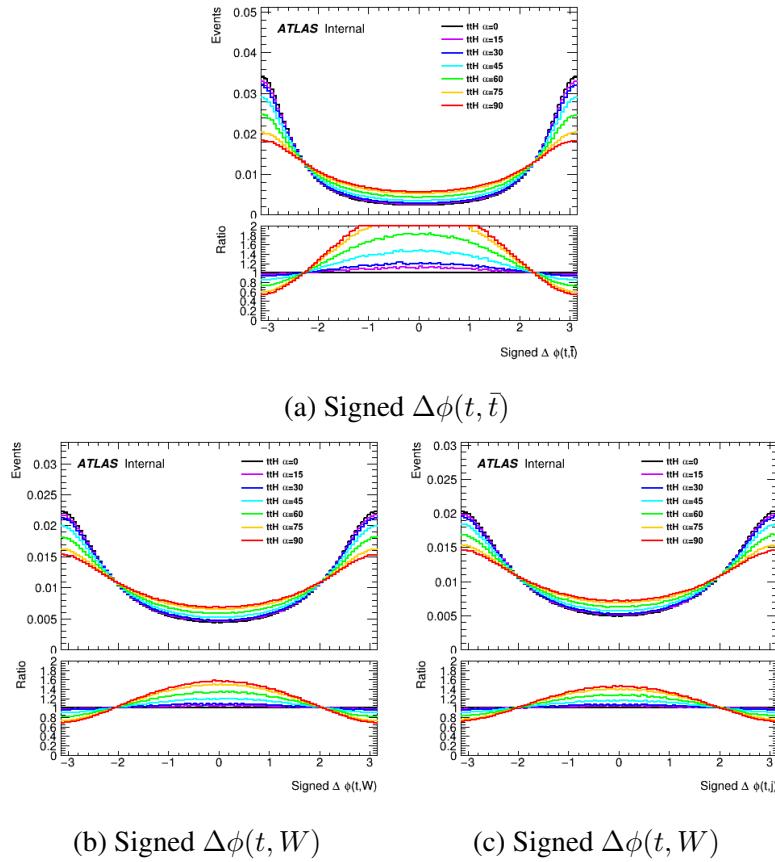


Figure 6.6: The parton-level signed $\Delta\phi$ between (a) the two top quarks in $t\bar{t}H$, (b) a top quark and the daughter W of the other top quark, and (c) a top quark and the highest p_T light jet from the hadronic decay of the other top's daughter W .

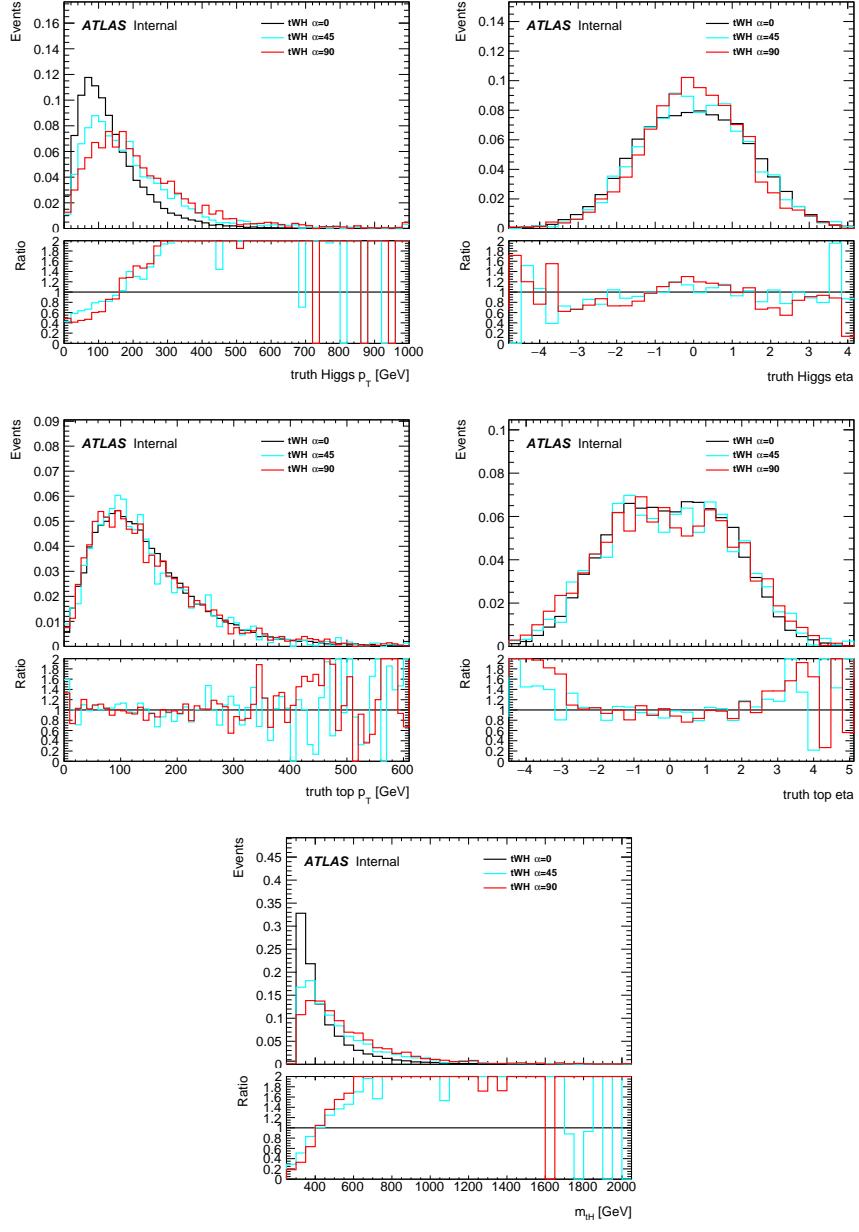


Figure 6.7: Truth-level distributions in tWH Monte Carlo of the Higgs boson p_T and η (top), top quark p_T and η (middle) and invariant mass of the top-Higgs system (bottom) for different values of the CP mixing angle α .

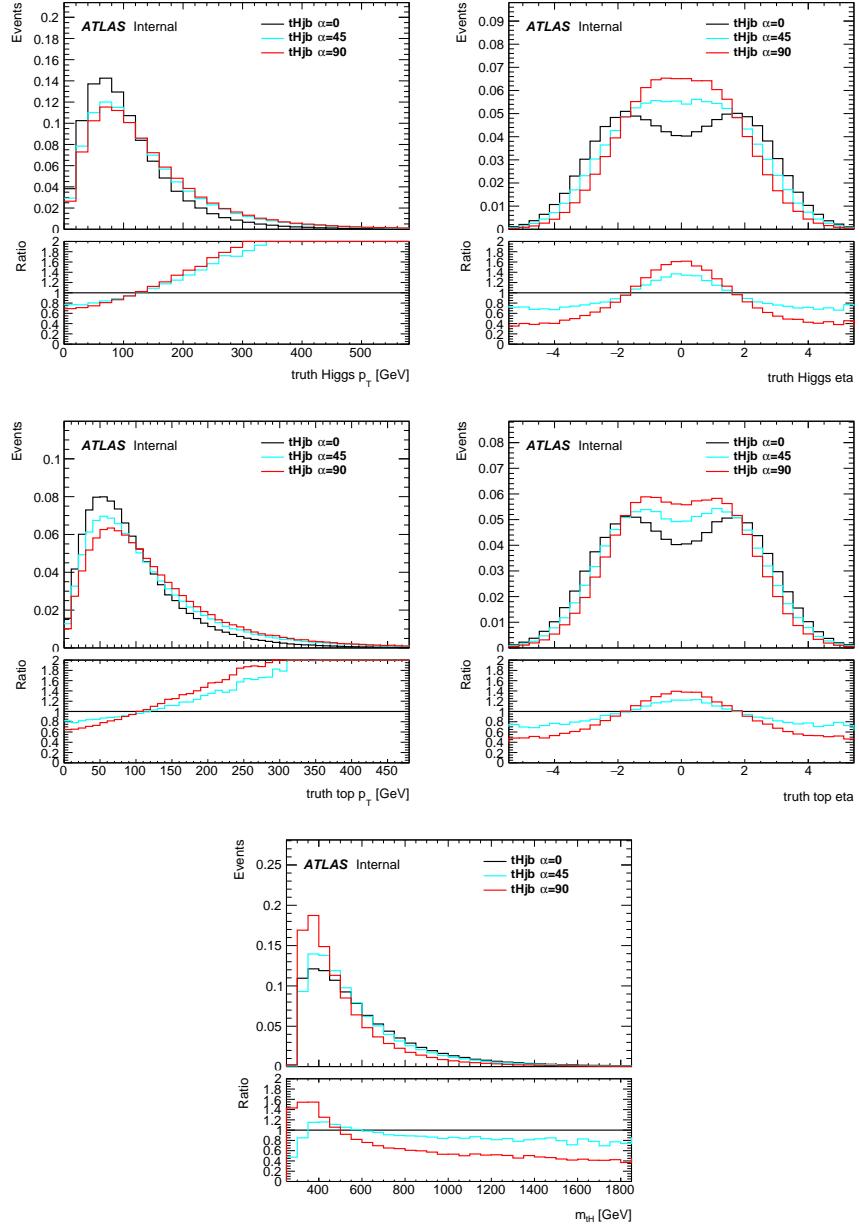


Figure 6.8: Truth-level distributions in $tHjb$ Monte Carlo of the Higgs boson p_T and η (top), angular separation between top and anti-top quarks (second row), top quark p_T and η (third row) and invariant mass of the top-Higgs system (bottom) for different values of the CP mixing angle α .

6.1.3 CPBDT

Similar to the SBBDT case, two dedicated CPBDTs are trained using XGBoost, one in the hadronic preselection region and one in the leptonic preselection region. An alternative implementation of the BDT is developed using the Toolkit for Multivariate Analysis (TMVA) ?? and is discussed in Appendix ???. Studies performed with the TMVA BDT helped guide the implementation of this BDT, including making the determination that only one dedicated CP-even vs CP-odd BDT was needed (rather than training for multiple α points) and the discovery of several useful high-level variables.

The signal samples are the CP-odd $tH + tWH + tHjb$, added according to their predicted cross-sections, while the background samples are the CP-even $tH + tWH + tHjb$. 50% of the signal and background samples are used for training, 25 % are used for categorization, and 25% are used for testing and significance evaluation.

For both the hadronic and leptonic CPBDTs, input variables chosen are:

The training variables used in the hadronic CP BDT training are:

- p_T and η of the Higgs candidate
- p_T, η, ϕ (with respect to the Higgs candidate), and BDT score of the first and second reconstructed tops. Due to its potential to be composed of fewer than three jets, the second top is referred to as the 'hybrid' top. In the case where no hybrid top is reconstructed, a dummy value is passed to XGBoost.
- Angles $\Delta\eta$ and $\Delta\phi$ between the top candidates. In the case where no hybrid top is reconstructed, a dummy value is passed to XGBoost.
- Two-object invariant masses m_{t1H} , and m_{t1hy} . In the case where no second hybrid top is reconstructed, a dummy value is passed to XGBoost.
- $H_T = \sum_{\text{jet } j} p_T^j$
- The minimum ΔR between a photon and a jet in the event (out of all photon-jet combinations)
- The second-smallest ΔR between a photon and a jet in the event (out of all photon-jet combinations)
- Jet multiplicity and b -jet multiplicity (77% working point)
- Missing E_T significance = $E_T^{\text{miss}} / \sqrt{H_T}$

The XGBoost BDT parameters are optimized by running the training 100 times with hyper-parameters selected according to a Bayesian minimization procedure [?]. The integral of the Receiver Operating Curve (ROC AUC) is used as the optimization metric, evaluated on the validation set for each training. The ROC-AUC measures true-positives versus true-negatives, parameterized as a function of classifier threshold cuts: a completely random classifier has a ROC-AUC of 0.5, while a perfect classifier has a ROC-AUC of 1.0 ???. The ROC-AUC for the hadronic CPBDT is 0.7839, while the ROC-AUC for the leptonic BDT is 76.69.

Plots of the CPBDT input variables for the leptonic BDT are shown in figures ?? - ??, while plots of the hadronic CPBDT input variables are shown in figures ?? - ??.

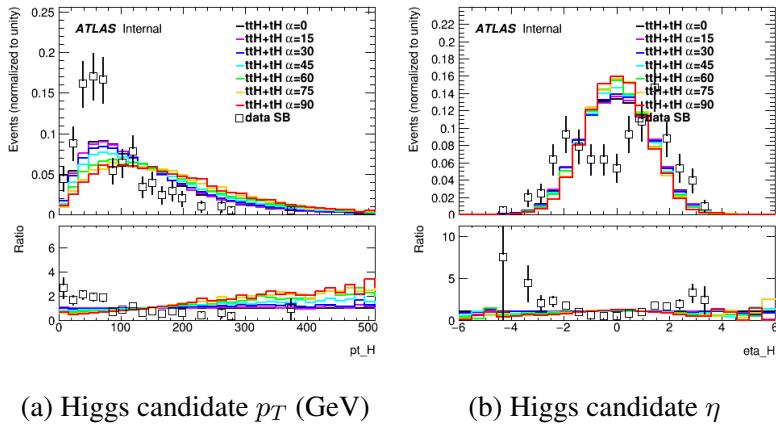


Figure 6.9: Leptonic BDT training variables related to the photons. The open squares show data in the NTI sideband region, giving an approximation of the shape of the continuum background.

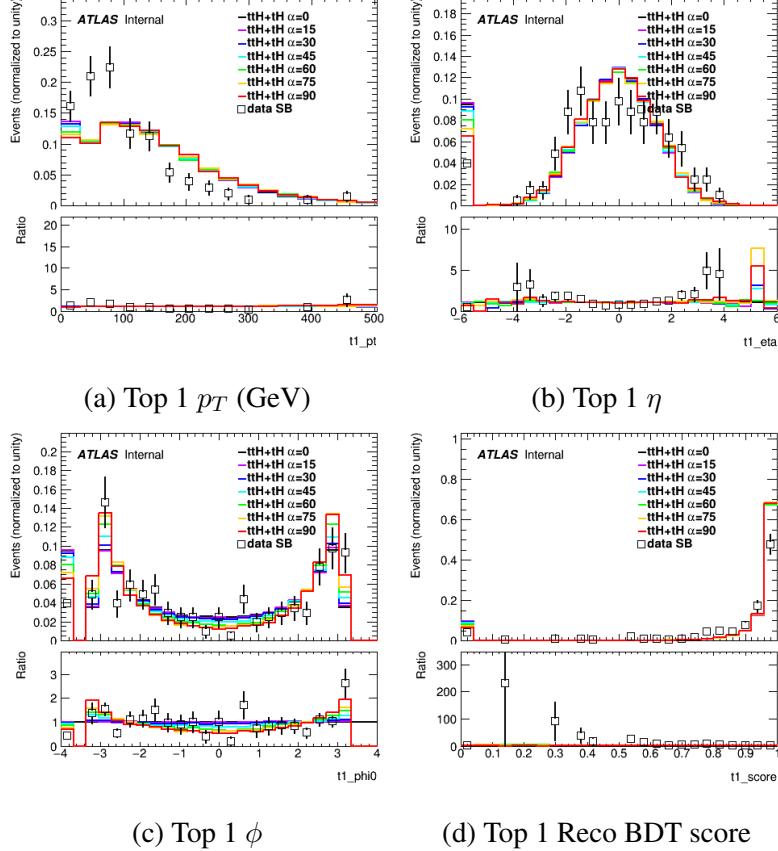


Figure 6.10: Leptonic BDT training variables related to the first reconstructed top quark. The ϕ is calculated with respect to the Higgs candidate. The open squares show data in the NTI sideband region, giving an approximation of the shape of the continuum background.

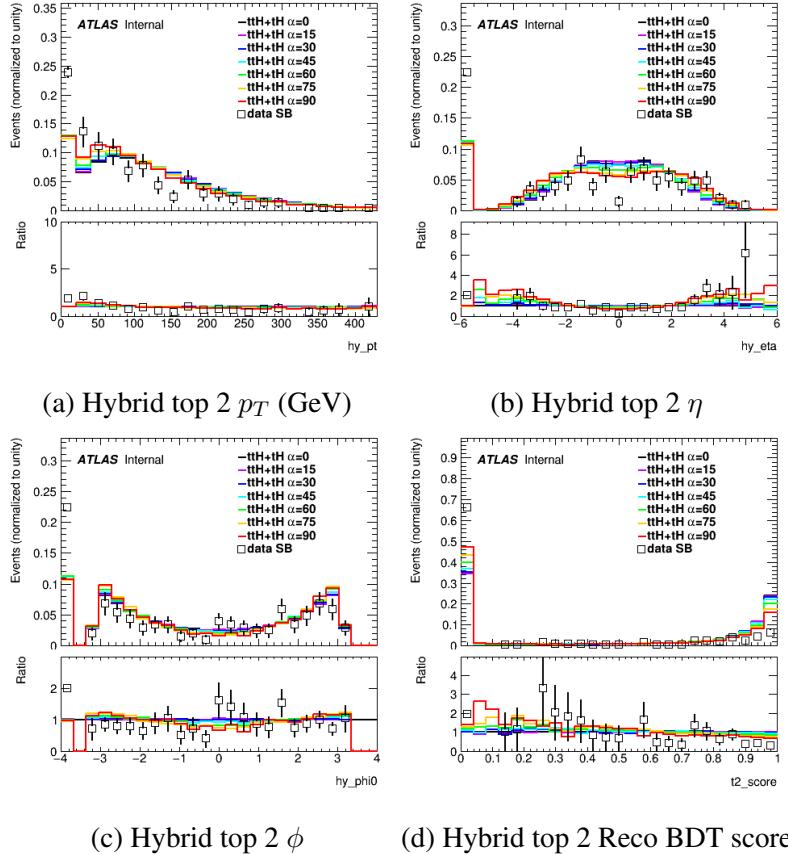


Figure 6.11: Leptonic BDT training variables related to the second reconstructed top quark. The ϕ is calculated with respect to the Higgs candidate. The underflow bins in p_T , η , and ϕ represent events where no second ("hybrid") top is reconstructed. The underflow bin in the Reco BDT score represents events with fewer than 6 jets. The open squares show data in the NTI sideband region, giving an approximation of the shape of the continuum background.

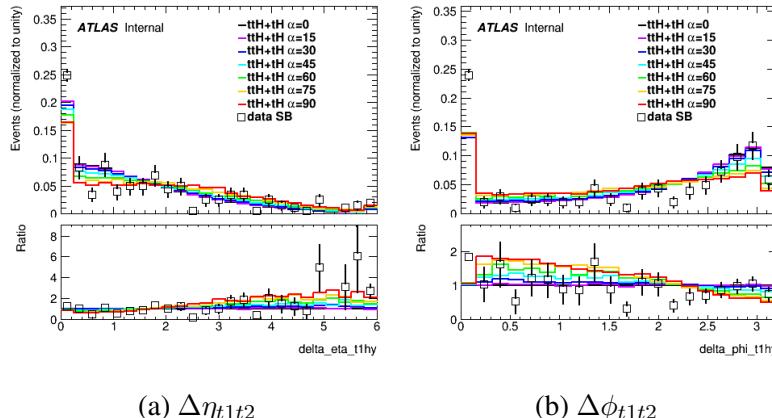


Figure 6.12: Leptonic BDT training variables related to angles between the top quarks in the event. The underflow bins represent events where no second ("hybrid") top is reconstructed. The open squares show data in the NTI sideband region, giving an approximation of the shape of the continuum background.

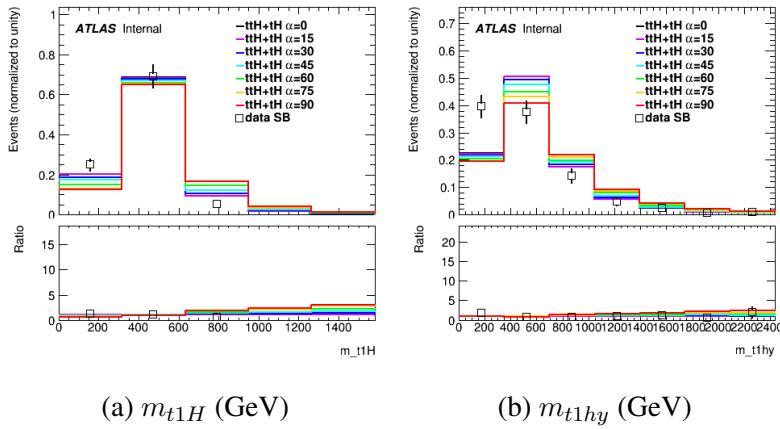


Figure 6.13: Multi-object invariant masses used as training variables in the leptonic BDT. The underflow bins in m_{t2H} and m_{t1t2} represent events where no second ("hybrid") top is reconstructed. The open squares show data in the NTI sideband region, giving an approximation of the shape of the continuum background.

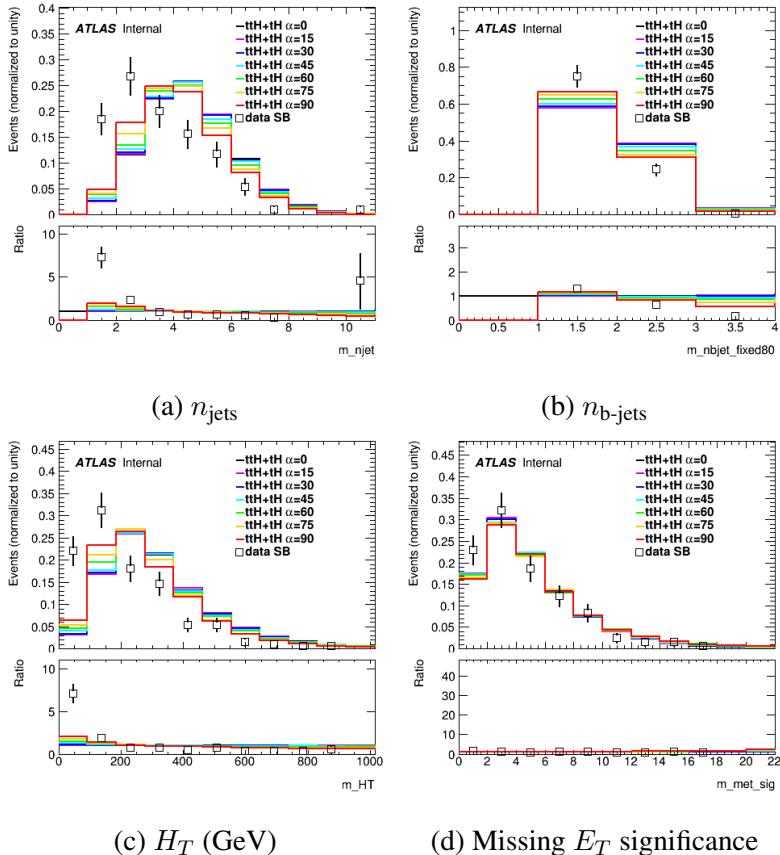


Figure 6.14: Leptonic BDT training variables jet multiplicity, b-jet multiplicity, H_T , and missing E_T significance. The open squares show data in the NTI sideband region, giving an approximation of the shape of the continuum background.

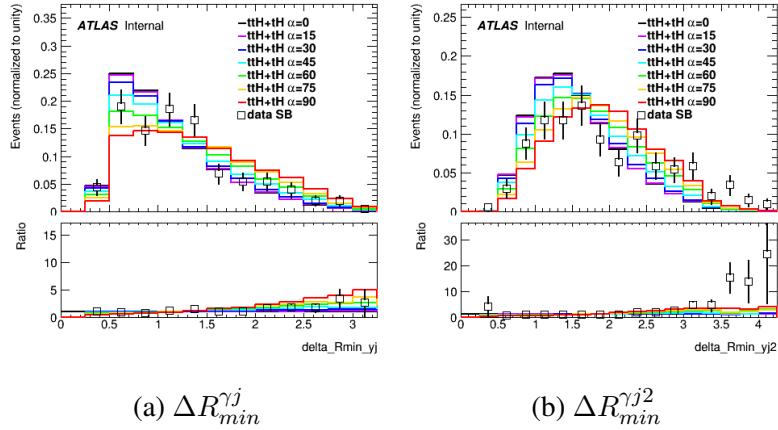


Figure 6.15: Leptonic BDT training variables $\Delta R_{min}^{\gamma j}$ and $\Delta R_{min}^{\gamma j2}$. The open squares show data in the NTI sideband region, giving an approximation of the shape of the continuum background.

The XGBoost BDT parameters are optimized by running the training 100 times with hyper-parameters selected according to a Bayesian minimization procedure [?]. The area under the ROC curve (ROC AUC) is evaluated on the validation set for each training, and the training that maximizes the ROC AUC is selected as optimal. The values of the optimal leptonic training are shown in Table ???. Details of each parameter can be found in the XGBoost documentation [?]. Note that the large value of the regularization parameter alpha (compared to default $\text{alpha}=1$) is responsible for the asymmetry between signal and background (red and black) shown in Figures ?? and ??.

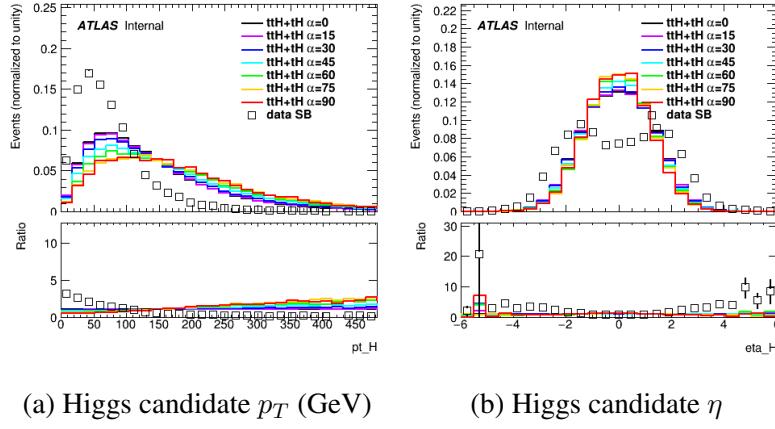


Figure 6.16: Hadronic BDT training variables related to the photons. The open squares show data in the NTI sideband region, giving an approximation of the shape of the continuum background.

Plots of the CPBDT output score for both the hadronic and leptonic CPBDT are shown in figures ?? and ??, while two-dimensional plots of CPBDT score versus SBBDT score for data in both the hadronic and leptonic regions are shown in figure XXX.

6.1.4 Poisson Number-Counting Significance

In order to determine the optimal categorization based on BDT score, two Poisson number-counting significance metrics are used.

In physics analyses such as those discussed in this dissertation, the probability of observing k events in a given region of phase space given an expected number of events λ is modeled by a Poisson distribution:

$$P(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \quad (6.1)$$

To determine the compatibility of the observed number of events with a given signal hypothesis H vs the Standard Model, it is useful to define a test statistic called the Poisson Number-Counting Significance:

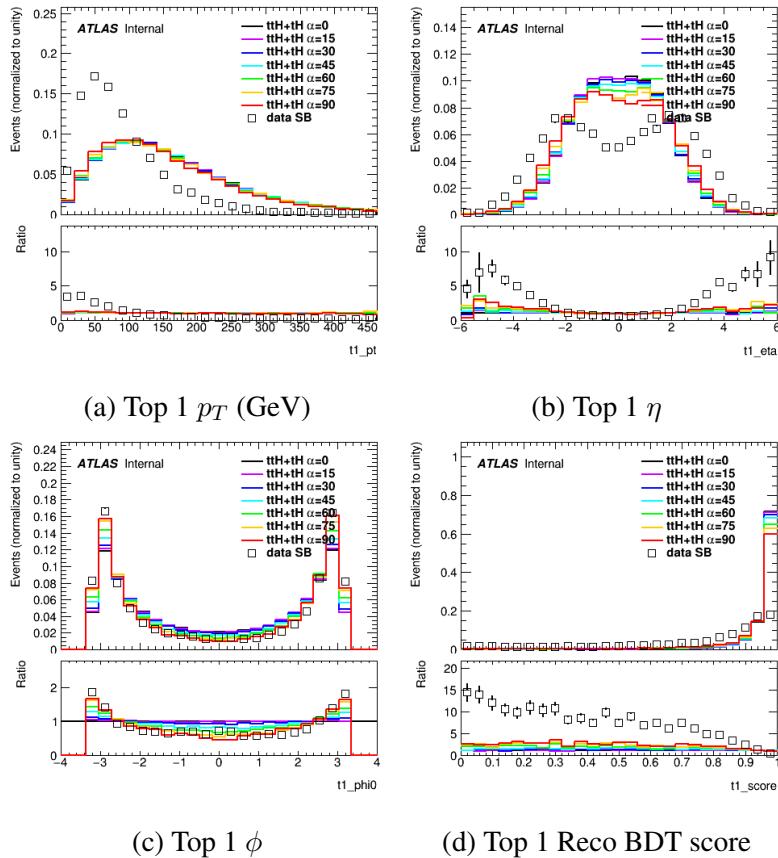


Figure 6.17: Hadronic BDT training variables related to the first reconstructed top quark. The ϕ is calculated with respect to the Higgs candidate. The open squares show data in the NTI sideband region, giving an approximation of the shape of the continuum background.

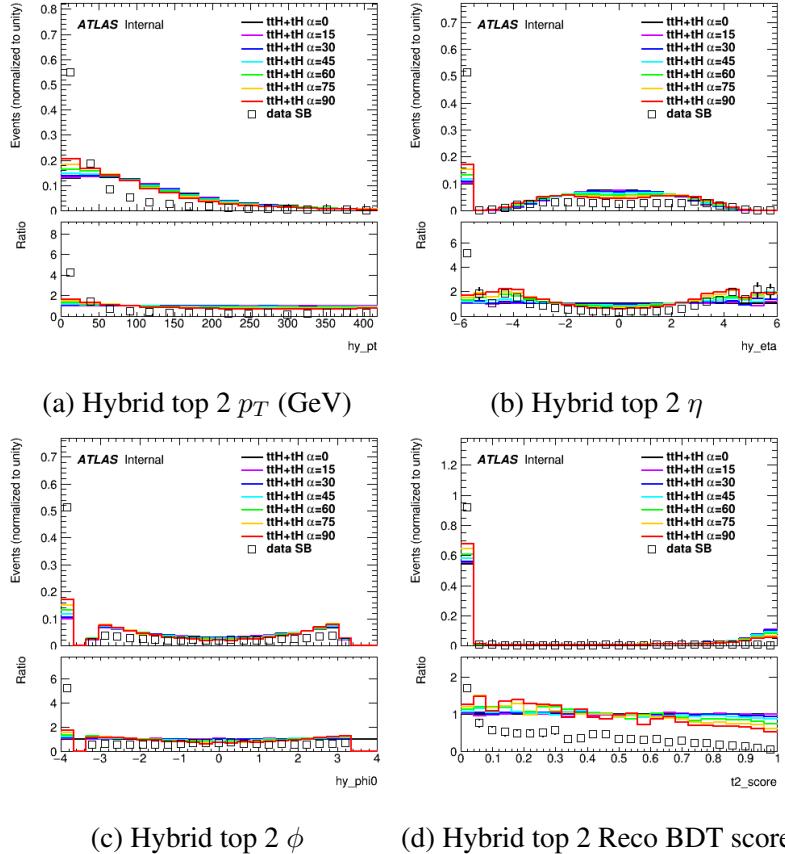


Figure 6.18: Hadronic BDT training variables related to the second reconstructed top quark. The ϕ is calculated with respect to the Higgs candidate. The underflow bins in p_T , η , and ϕ represent events where no second ("hybrid") top is reconstructed. The underflow bin in the Reco BDT score represents events with fewer than 6 jets. The open squares show data in the NTI sideband region, giving an approximation of the shape of the continuum background.

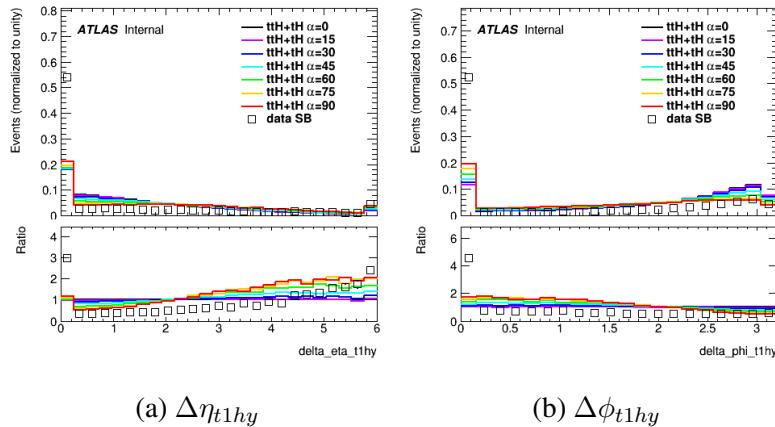


Figure 6.19: Hadronic BDT training variables related to angles between the top quarks in the event. The underflow bins represent events where no second ("hybrid") top is reconstructed. The open squares show data in the NTI sideband region, giving an approximation of the shape of the continuum background.

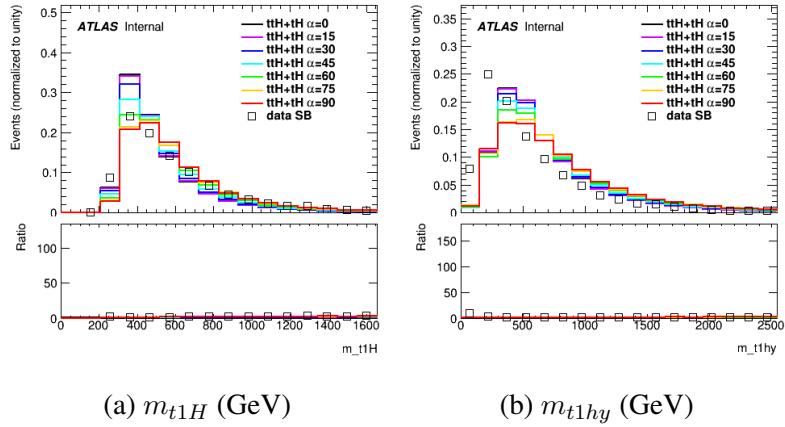


Figure 6.20: Multi-object invariant masses used as training variables in the hadronic BDT. The underflow bins in m_{t2H} and m_{t1t2} represent events where no second ("hybrid") top is reconstructed. The open squares show data in the NTI sideband region, giving an approximation of the shape of the continuum background.

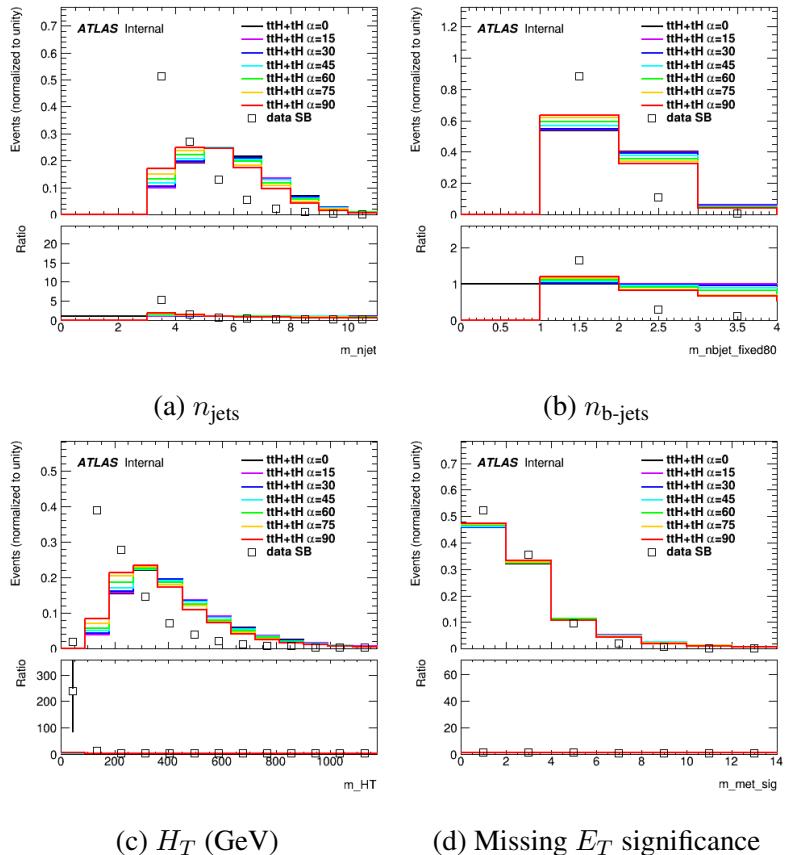


Figure 6.21: Hadronic BDT training variables jet multiplicity, b-jet multiplicity, H_T , and missing E_T significance. The open squares show data in the NTI sideband region, giving an approximation of the shape of the continuum background.

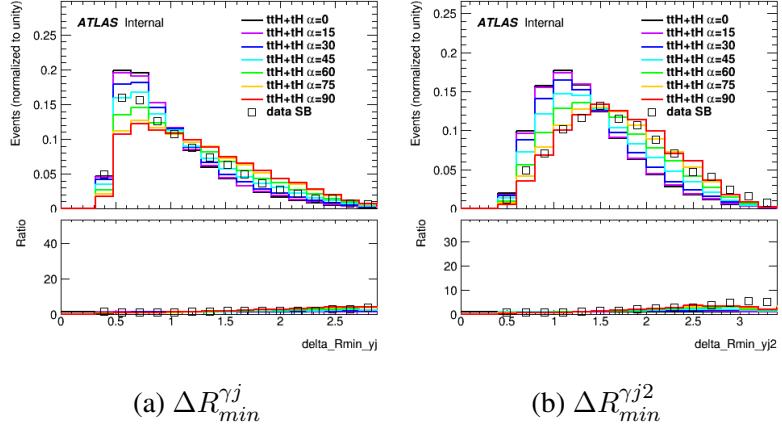


Figure 6.22: Hadronic BDT training variables $\Delta R_{min}^{\gamma j}$ and $\Delta R_{min}^{\gamma j^2}$. The open squares show data in the NTI sideband region, giving an approximation of the shape of the continuum background.

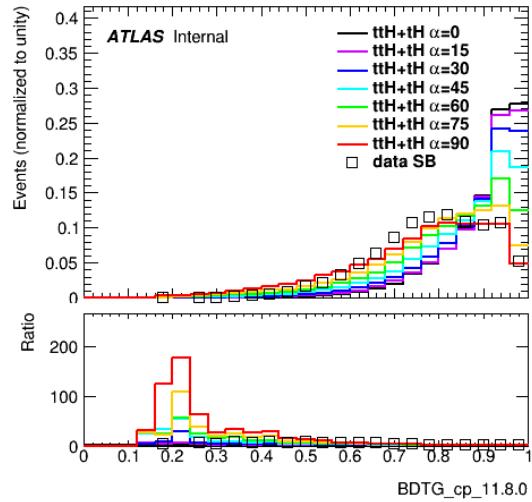


Figure 6.23: Hadronic channel CP BDT score for $t\bar{t}H + tHjb + tWH$, with relative weights according to the predicted cross section for various CP mixing scenarios. The open squares show data in the NTI sideband region.

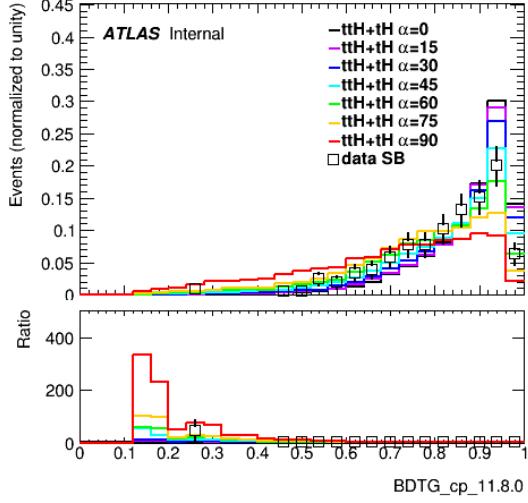


Figure 6.24: Leptonic channel CP BDT score for $t\bar{t}H+tHjb+tWH$, with relative weights according to the predicted cross section for various CP mixing scenarios. The open squares show data in the NTI sideband region.

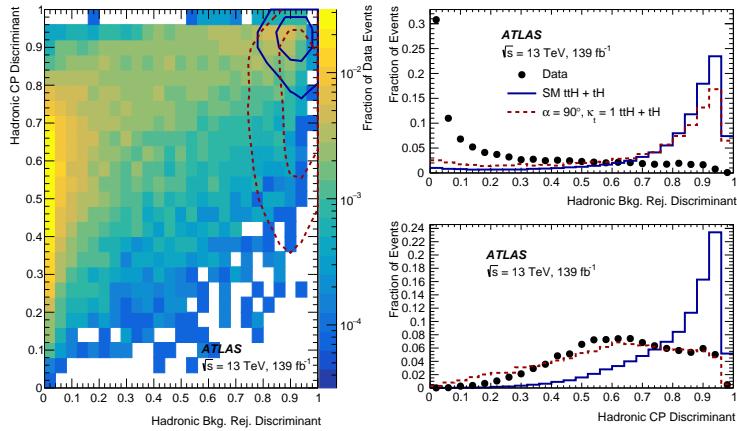


Figure 6.25: Distribution of events from TI sidebands, CP even signal, and CP odd signal in the 2D background rejection BDT vs. CP BDT plane in the hadronic category are shown in full color, black, and red contours, respectively, along with 1D projections onto each BDT score. Inner (outer) contours contain 25% (50%) of signal events.

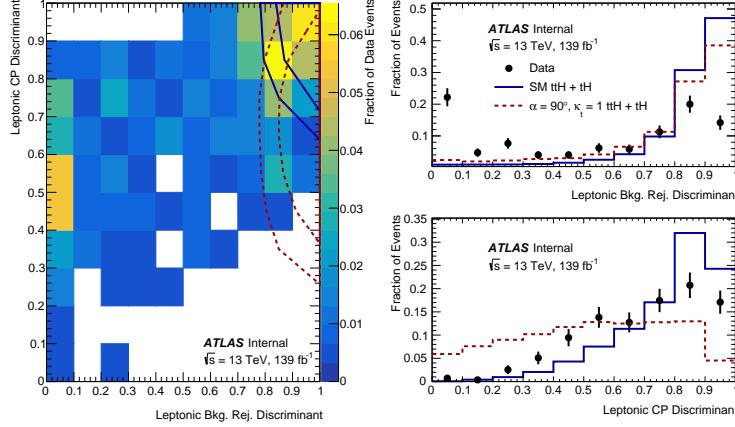


Figure 6.26: Distribution of events from TI sidebands, CP even signal, and CP odd signal in the 2D background rejection BDT vs. CP BDT plane in the leptonic category are shown in full color, black, and red contours, respectively, along with 1D projections onto each BDT score. Inner (outer) contours contain 25% (50%) of signal events.

$$\begin{aligned}
 Z^2 &= -2\ln\left(\frac{P(S+B, S_H+B)}{P(S+B, S+B)}\right) \\
 &= 2(S+B)\ln\left(\frac{(S+B)}{(S_H+B)}\right) - 2(S+B) + 2(S_H+B)
 \end{aligned} \tag{6.2}$$

Where S and B are the Standard-Model predicted signal and background and S_{h1} is the predicted signal under hypothesis H.

Thus, the $t\bar{t}H + tH$ Standard-Model number-counting significance (that is, the significance of observing the amount of $t\bar{t}H + tWH + tHjb$ in analysis categories that predicted by the Standard Model if in fact the signal process does not exist and the background-only "null hypothesis" is true) can be defined as:

$$Z_{t\bar{t}H+tH} = \sqrt{2((S+B)\ln(1+S/B) - S)} \tag{6.3}$$

and the CP-Odd number-counting significance as:

$$Z_{CP}(90) = \sqrt{2(S_e+B)\ln\left(\frac{S_e+B}{S_o+B}\right) - 2(S_e+B) + 2(S_o+B)} \tag{6.4}$$

where S_e is the amount of signal expected if $\alpha = 0^\text{deg}$ and S_o is the amount of signal expected if $\alpha = 90^\text{deg}$.

To determine the total significance for either of these metrics across a number of categories, the number-counting significances for each category are added in quadrature.

To estimate the number of continuum background events B in the diphoton mass signal region in each category (that is, $123\text{GeV} < m_{\gamma\gamma} < 127\text{GeV}$), a scaling method is applied to the NTI data

control region in the mass sidebands (that is, $105\text{GeV} < m_{\gamma\gamma} < 123\text{GeV}$ or $127\text{GeV} < m_{\gamma\gamma} < 160\text{GeV}$):

$$\begin{aligned} N(TI, signalwindow) &= N(NTI, masssidebands) \times \frac{N(TI, masssidebands)}{N(NTI, masssidebands)} \times \frac{N(NTI, signalwindow)}{N(NTI, masssidebands)} \\ &= N(NTI, masssidebands) \times f_1 \times f_2 \end{aligned} \quad (6.5)$$

Where the scale-factors f_1 and f_2 are calculated separately using the hadronic and leptonic preselection regions. The combined scale factor, $f_1 \times f_2$, is found to be 0.013 for the hadronic region and 0.016 for the leptonic region.

In order to ensure that there is sufficient data to perform a likelihood fit, at least 0.8 continuum events (as modelled by scaled NTI sidebands) are required in the $123\text{GeV} < m_{\gamma\gamma} < 127\text{GeV}$ signal window in each category.

6.1.5 2D Categorization

To determine the categorization, a brute-force scan is performed across many different sets of category boundaries, calculating both Z_{ttH+tH} and $Z_{CP}(90)$ on the validation set. Because the same categorization scheme does not optimize both metrics, a compromise scenario is constructed, for which Z_{ttH+tH} is maximized while $Z_{CP}(90)$ is required to be no more than 0.15σ less than its maximal determined value. The brute-force scan, and this effect, is shown in figure ???. The categories selected are given in table ???. The expected yield and purity in each category is given in figure .

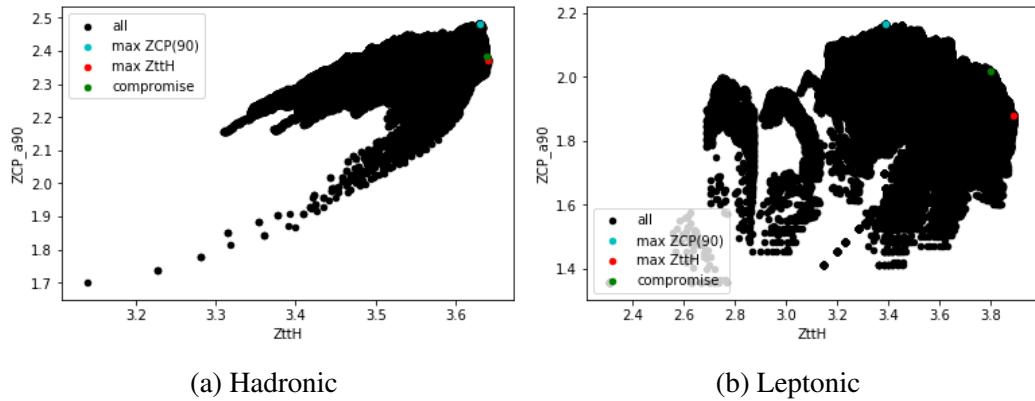


Figure 6.27: Z_{CP} vs. Z_{ttH} for all sets of boundaries considered.

The background-only statistical uncertainty in a given category is given by:

$$s = \frac{S}{Z_{ttH+tH}} \quad (6.6)$$

Category	Bkg. Rej. BDT score	CP BDT score
1	[0 – 0.005]	[0 – 0.90]
2	[0 – 0.005]	[0.90 – 0.97]
3	[0 – 0.005]	[0.97 – 1]
4	[0.005 – 0.009]	[0 – 0.88]
5	[0.005 – 0.009]	[0.88 – 0.96]
6	[0.005 – 0.009]	[0.96 – 1]
7	[0.009 – 0.019]	[0 – 0.84]
8	[0.009 – 0.019]	[0.84 – 0.96]
9	[0.009 – 0.019]	[0.96 – 1]
10	[0.019 – 0.091]	[0 – 0.61]
11	[0.019 – 0.091]	[0.61 – 0.86]
12	[0.019 – 0.091]	[0.86 – 1]
13	[0 – 0.012]	[0 – 0.91]
14	[0 – 0.012]	[0.91 – 1]
15	[0.012 – 0.085]	[0 – 0.82]
16	[0.012 – 0.085]	[0.82 – 0.93]
17	[0.012 – 0.085]	[0.93 – 1]
18	[0.085 – 0.748]	[0 – 0.72]
19	[0.085 – 0.748]	[0.72 – 0.86]
20	[0.085 – 0.748]	[0.86 – 1]

Table 6.1: Category boundaries that optimize the number-counting rejection of the CP odd scenario in the 12 hadronic and 8 leptonic categories.

It is also useful to define:

$$n_s(\alpha) = n_{ttH}^{SM} + n_{tWH}^{SM} + n_{tHjb}^{SM} - n_{ttH}(\alpha) - n_{tWH}(\alpha) - n_{tHjb}(\alpha) \quad (6.7)$$

as an indication of the "signal" (i.e., difference between Standard-Model and alternative- α yield) in each category.

Using this, the effects of major systematics can be computed, ensuring they do not need to be accounted for at the categorization stage. The 100% ggF yield uncertainty and the Parton Shower Uncertainty (calculated by subtracting the yield in each category using the Herwig Monte Carlo samples from the yield using the Pythia samples) are shown in Table ??- in all categories, it is clear that the statistical uncertainty dominates. This can be confirmed by plotting the number-counting rejection for the CP-odd category, $Z_{CP}(90^{\text{deg}})$, both with and without systematics in figure ??.

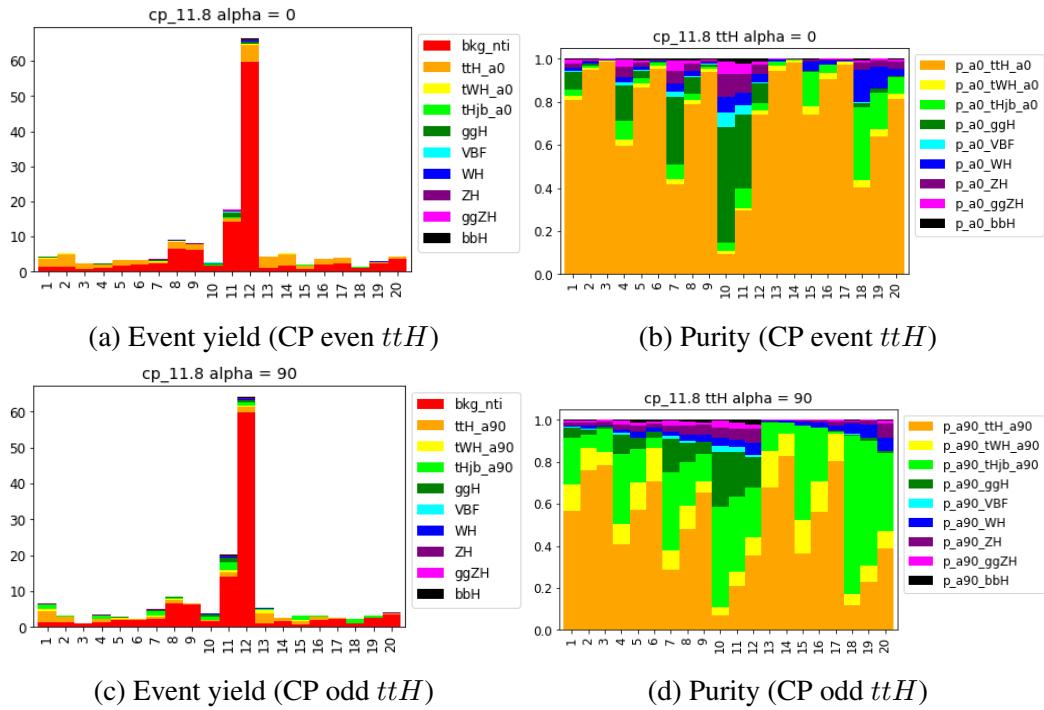


Figure 6.28: (Left) Event yields in the CP categories at 139 fb^{-1} . Shown separately for $\alpha = 0^{\text{deg}}$ and $\alpha = 90^{\text{deg}}$. (Right) purity of the Higgs yield in each category for $\alpha = 0^{\text{deg}}$ and $\alpha = 90^{\text{deg}}$. Yields are calculated in the signal window $m_{\gamma\gamma} = 125 \pm 2 \text{ GeV}$.

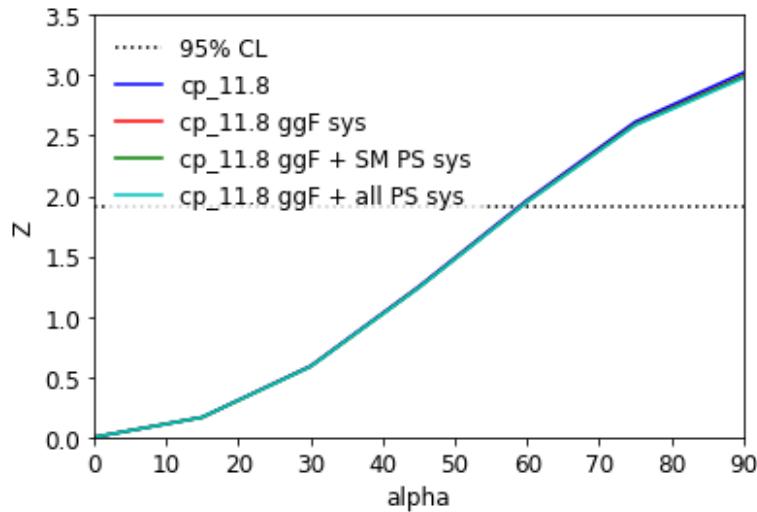


Figure 6.29: The impact on the number counting limit as systematic uncertainties are progressively added. There is a small change of 0.3° on the number-counting limit.

Category	$n_s(90^\circ)$	δ	n_{ggF}	PS (ttH)	PS (tWH)	PS (tHjb)
1	-2.226	1.462	0.247	-0.157	0.008	0.01
2	2.038	1.581	0.044	-0.049	-0.007	-0.005
3	1.111	1.136	0.002	0.098	0	0
4	-0.956	1.275	0.208	-0.039	-0.043	-0.038
5	0.57	1.534	0.054	-0.008	-0.02	-0.006
6	1.086	1.554	0.009	0.08	-0.005	-0.007
7	-1.393	1.597	0.448	-0.053	-0.02	-0.014
8	0.505	2.713	0.186	0.015	-0.01	-0.002
9	1.287	2.623	0.021	0.094	0.007	-0.001
10	-1.09	1.363	0.557	-0.003	0.019	-0.012
11	-2.365	3.858	1.25	-0.053	-0.038	0.026
12	2.33	7.86	0.6	0.007	-0.025	0.006
13	-0.835	1.398	0.001	-0.149	-0.024	-0.006
14	2.542	1.64	0.001	0.156	-0.008	0.001
15	-1.218	1.168	0.004	-0.089	-0.024	-0.011
16	0.473	1.643	0	-0.015	-0.012	-0.014
17	1.32	1.697	0.001	-0.036	0.001	0
18	-0.899	1.082	0.008	-0.016	0.032	-0.027
19	-0.284	1.627	0.008	-0.024	0.022	-0.01
20	0.309	1.973	0.003	-0.014	-0.013	0.013

Table 6.2: Comparison of statistical uncertainty with key systematics and CP-Odd vs. SM separation in each category. PS indicates parton showering uncertainty, calculated by subtracting the yields from the Herwig and the Pythia Monte Carlo samples.

Number-counting ttH significance (if even)	5.00σ
Number-counting ttH significance (if odd)	2.75σ
Number-counting $tHjb + tWH$ significance (if even)	0.321σ
Number-counting $tHjb + tWH$ significance (if odd)	2.23σ
Number-counting CP-odd rejection	3.02σ
Number-counting $\alpha = 45^\circ$ rejection	1.25σ

Table 6.3: Figures of merit for the twenty-category CP BDT categorization, calculated from event yields in the signal $m_{\gamma\gamma}$ region 125 ± 2 GeV.

APPENDIX A

Alternative Top Reconstruction with the KLFitter

A.1 The KLFitter

In order to validate the effectiveness of the top reconstruction BDT, we test its performance against that of the Kinematic Likelihood Fitter (KLFitter) ??.

The KLFitter is a likelihood-based framework that, like the BDT, attempts to solve the top reconstruction combinatorial jet-matching problem. To correctly identify the jet triplet originating from the top, the KLFitter maximizes a likelihood function that depends on the kinematics of each of the final-state jets for each possible jet permutation. Based on this likelihood-maximization, an event probability is calculated that allows us to identify the best potential top candidate from the jet permutations.

For semileptonic tops (the decay mode in which we perform this study), the form of the likelihood function is:

$$\begin{aligned} \mathcal{L} = & \mathcal{B}(m_{q_1 q_2 q_3} | m_t, \Gamma_t) \times \mathcal{B}(m_{q_1 q_2} | m_W, \Gamma_W) \times \\ & \mathcal{B}(m_{q_4 l\nu} | m_t, \Gamma_t) \times \mathcal{B}(m_{l\nu} | m_W, \Gamma_W) \times \\ & \prod_{i=1}^{\infty} C_{jet}(E_{jet,i}^{meas} | E_{jet,i}) \times C_l(E_l^{meas} | E_l) \times \\ & C_{miss}(E_x^{miss} | p_x^\nu) \times C_{miss}(E_y^{miss} | p_y^\nu) \end{aligned} \quad (\text{A.1})$$

where the \mathcal{B} s denote Breit-Wigner functions (similar to Gaussians, dependent on both mass and decay width) and the C terms indicate transfer functions, designed to model the difference between measured kinematics and their true values due to detector effects. At the time of this study, transfer functions for the ATLAS detector were defined only for jets with $|\eta| < 4.5$, so we restrict our study to this range.

The log-likelihood algorithm is minimized using the Minuit algorithm ?? as implemented in the Bayesian Analysis Toolkit (BAT) ??.

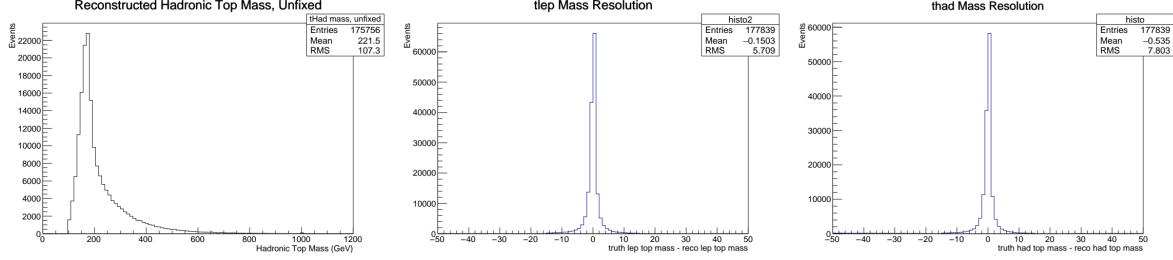


Figure A.1: Reconstructed top-mass and top-mass resolution of the KLFitter (using the "unfixed" top-mass setting to illustrate performance).

A.2 Comparison

In order to compare the KLFitter and BDT methods, we construct a validation set of truth-tagged semileptonic top events. We use the "mc16a" (2015-like pileup profile) PowhegPy8 $t\bar{H}$ sample, and select only events which contain at least four jets and exactly one lepton, all with $|\eta| < 4.5$.

We find that KLFitter performs optimally when we fix the top mass to its measured value of 172.5 GeV ?? and use the 'kWorkingPoint' b-tag handling method, which adds an additional multiplier to the event probability to account for the b-tagging efficiency and light-jet rejection.

	Correct Leptonic Tops (%)	Correct Hadronic Tops (%)	Both Correct (%)
KLFitter	55.59	21.92	16.20
Top Reco BDT	60.19	21.35	17.92

Table A.1: Comparison of KLFitter and top-reconstruction BDT.

A.3 Additional CP BDT Studies with TMVA

A.3.1 Four-Vector BDT, $t\bar{H}$ only

Prior to the generation of alternative-CP tH samples, an alternative $t\bar{H}$ -only CP-BDT was developed using only the four-vectors of objects in the event in order to check its performance against the nominal $t\bar{H}$ -only CPBDT developed using the top reconstruction variables. Utilizing four-vectors provides another potential solution to the large underflow of the top reconstruction BDT, and can be seen as an alternative to the "hybrid top" method. The BDT is trained using ROOT's TMVA package [?] using the same aMCnlo+Pythia8 training samples as the nominal CP-BDT, and utilizes the same train/test/significance subsampling scheme. It is reproduced in XGBoost for the significance comparison - a replication of the Nominal CP-BDT in TMVA performs near-identically, confirming that, for this analysis, the choice of MVA package appears to have no effect

on BDT performance.

A.3.1.1 Hadronic Channel

Similar to the nominal BDT, the hadronic four-vector CP BDT is trained to separate $t\bar{t}H$ CP even and CP odd aMCnlo+Pythia8 Monte Carlo passing the hadronic pre-selection (0 leptons, ≥ 1 b-jet at 77% working point).

The training variables used in the four-vector CP BDT training are:

- p_T , η , ϕ , and pseudocontinuous b-tag score of the leading 6 jets in the event
- The p_T of the Higgs candidate (scaled by mass)
- $\cos(\theta^*)$ (the cosine of the angle between the photons in the Collins-Soper frame ??)
- The η and ϕ of the two leading photons in the event
- The magnitude and ϕ of the missing transverse energy in the event
- The summed invariant mass of all jets in the event
- The minimum ΔR between a photon and a jet in the event
- The second-smallest ΔR between a photon and a jet in the event

The linear correlations between these variables in $t\bar{t}H$ CP even and CP odd aMCnlo+Pythia8 Monte Carlo are shown in Figure ?? . Figures ?? - ?? compare the distribution of each training variable in $t\bar{t}H$ CP even and CP odd Monte Carlo.

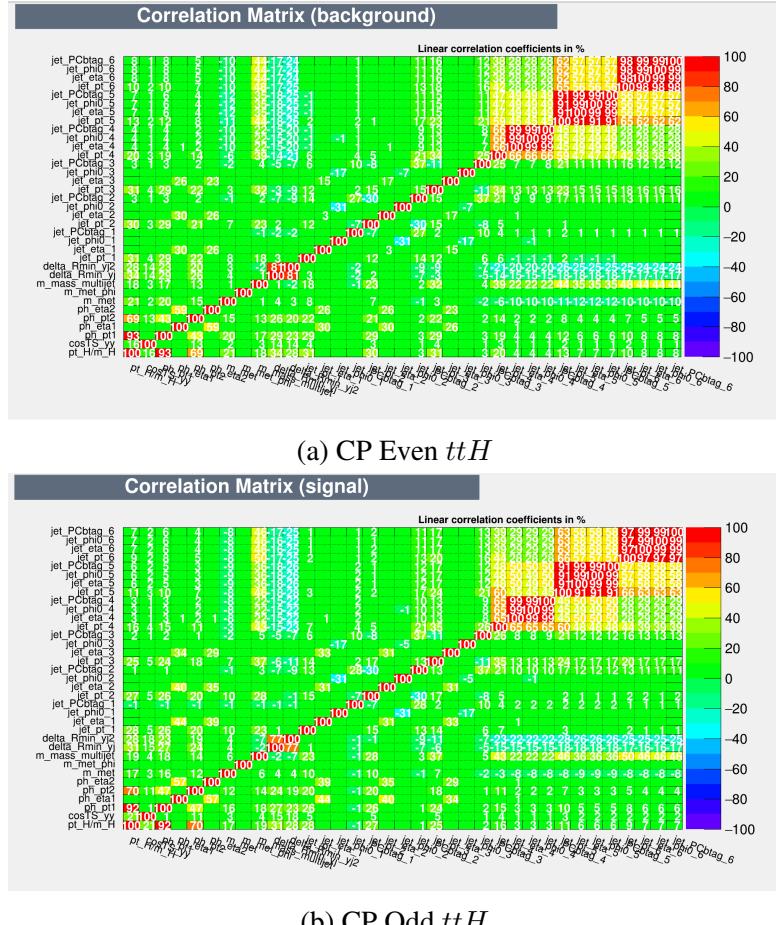


Figure A.2: Training variable correlations for events passing hadronic pre-selection.

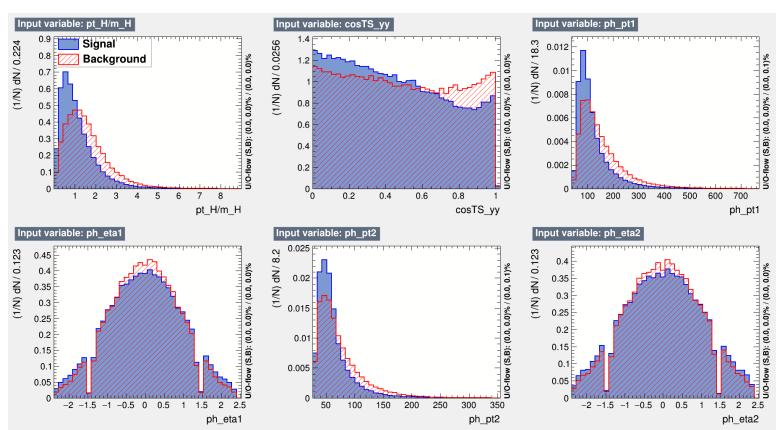


Figure A.3: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: Higgs candidate p_T (scaled by mass), $\cos(\theta^*)$, leading photon p_T , leading photon η , subleading photon p_T , and subleading photon η .

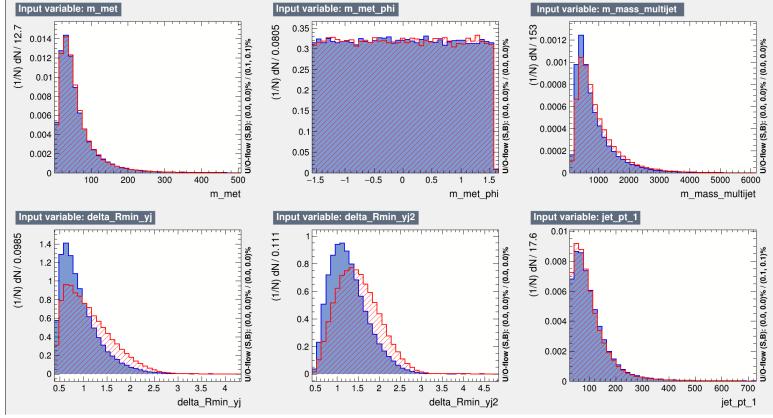


Figure A.4: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: Magnitude of MET, MET ϕ (branch cut chosen to range from $-\pi/2$ to $\pi/2$), invariant mass of all jets in the event, minimum ΔR between a photon and a jet, second-smallest ΔR between a photon and a jet, p_T of highest b-tag scoring jet.

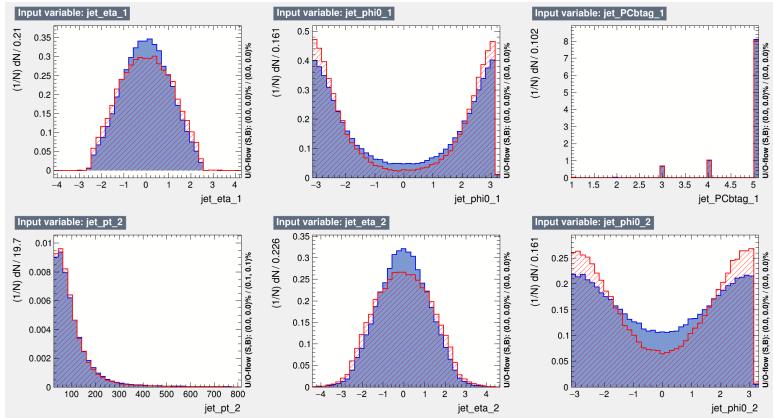


Figure A.5: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: η of highest b-tag scoring jet, ϕ of highest btag-scoring jet (measured with respect to the Higgs candidate), pseudo-continuous b-tag score of highest btag-scoring jet, p_T of second-highest b-tag scoring jet, η of second-highest b-tag scoring jet, ϕ of second-highest btag-scoring jet (measured with respect to the Higgs candidate).

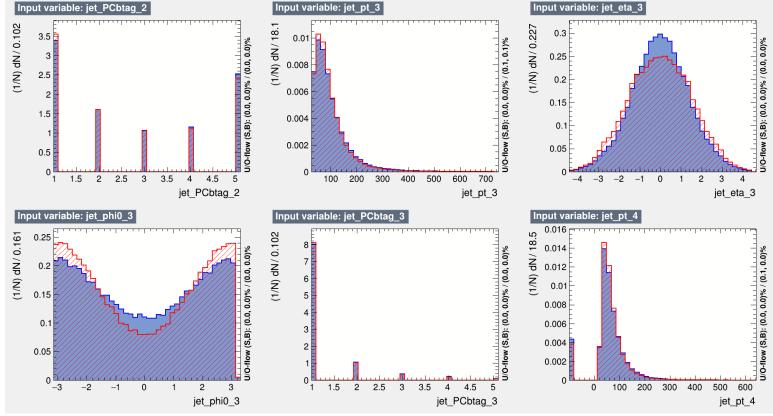


Figure A.6: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: pseudo-continuous b-tag score of second-highest btag-scoring jet, p_T of third-highest b-tag scoring jet, η of third-highest b-tag scoring jet, ϕ of third-highest btag-scoring jet (measured with respect to the Higgs candidate), pseudo-continuous b-tag score of third-highest btag-scoring jet, p_T of fourth-highest b-tag scoring jet.

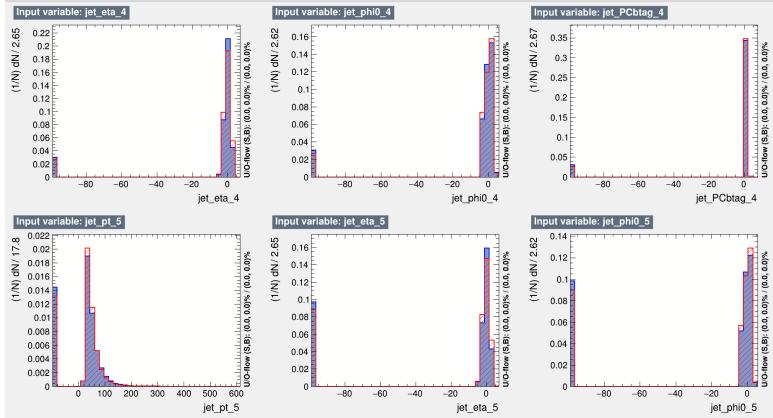


Figure A.7: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: η of fourth-highest b-tag scoring jet, ϕ of fourth-highest btag-scoring jet (measured with respect to the Higgs candidate), pseudo-continuous b-tag score of fourth-highest btag-scoring jet, p_T of fifth-highest b-tag scoring jet, η of fifth-highest b-tag scoring jet, ϕ of fifth-highest btag-scoring jet (measured with respect to the Higgs candidate)

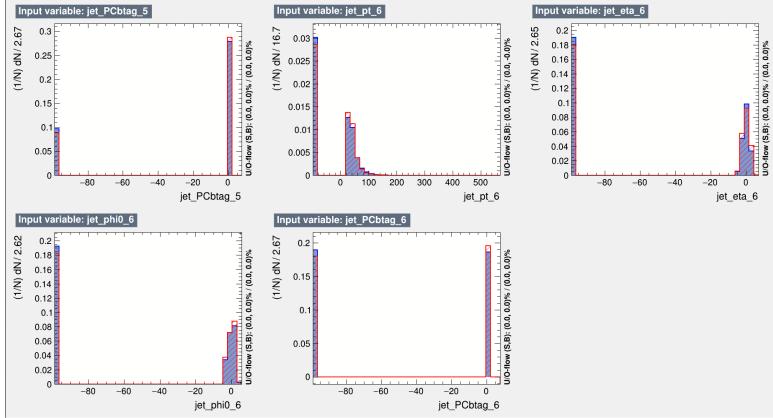


Figure A.8: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: Pseudo-continuous b-tag score of fifth-highest btag-scoring jet, p_T of sixth-highest b-tag scoring jet, η of sixth-highest b-tag scoring jet, ϕ of sixth-highest btag-scoring jet (measured with respect to the Higgs candidate), pseudo-continuous b-tag score of sixth-highest btag-scoring jet.

A.3.1.2 Leptonic channel

Likewise, the leptonic four-vector CP BDT is trained to separate $t\bar{t}H$ CP even and CP odd aMCnlo+Pythia8 Monte Carlo passing the leptonic pre-selection (≥ 0 leptons, ≥ 1 b-jet at 77% working point).

- p_T , η , ϕ , and pseudocontinuous b-tag score of the leading 4 jets in the event
- p_T , η , ϕ , and pseudocontinuous b-tag score of the leading 2 leptons in the event
- The p_T of the Higgs candidate (scaled by mass)
- $\cos(\theta^*)$ (the cosine of the angle between the photons in the Collins-Soper frame ??)
- The η and ϕ of the two leading photons in the event
- The magnitude and ϕ of the missing transverse energy in the event
- The summed invariant mass of all jets in the event
- The minimum ΔR between a photon and a jet in the event
- The second-smallest ΔR between a photon and a jet in the event

The linear correlations between these variables in $t\bar{t}H$ CP even and CP odd aMCnlo+Pythia8 Monte Carlo are shown in Figure ???. Figures ?? - ?? compare the distribution of each training variable in $t\bar{t}H$ CP even and CP odd Monte Carlo.

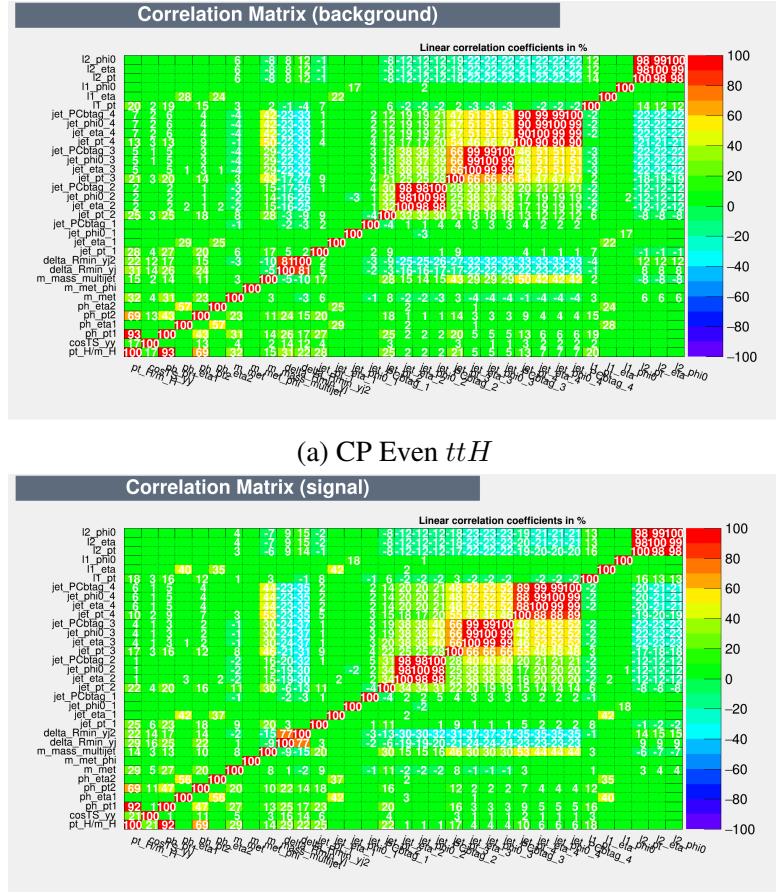


Figure A.9: Training variable correlations for events passing leptonic pre-selection.

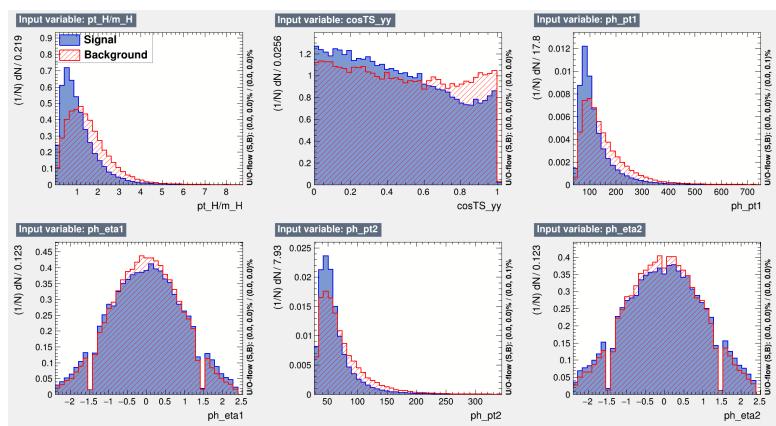


Figure A.10: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: Higgs candidate p_T (scaled by mass), $\cos(\theta^*)$, leading photon p_T , leading photon η , subleading photon p_T , and subleading photon η .

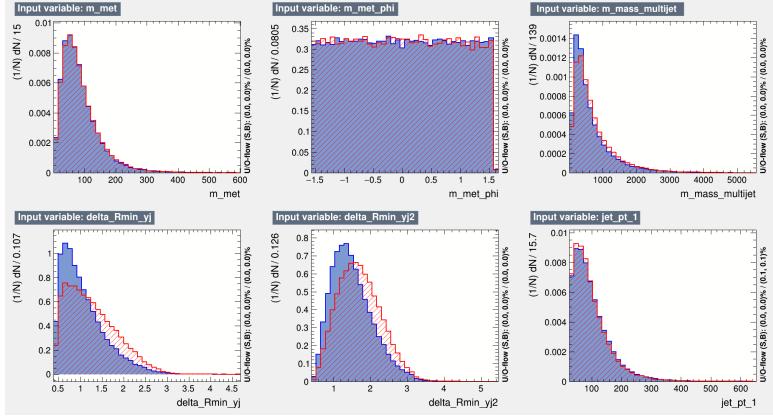


Figure A.11: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: Magnitude of MET, MET ϕ (branch cut chosen to range from $-\pi/2$ to $\pi/2$), invariant mass of all jets in the event, minimum ΔR between a photon and a jet, second-smallest ΔR between a photon and a jet, p_T of highest b-tag scoring jet.

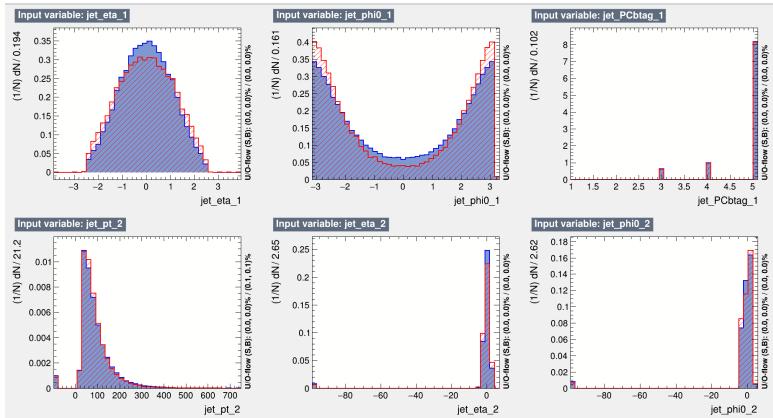


Figure A.12: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: η of highest b-tag scoring jet, ϕ of highest btag-scoring jet (measured with respect to the Higgs candidate), pseudo-continuous b-tag score of highest btag-scoring jet, p_T of second-highest b-tag scoring jet, η of second-highest b-tag scoring jet, ϕ of second-highest btag-scoring jet (measured with respect to the Higgs candidate).

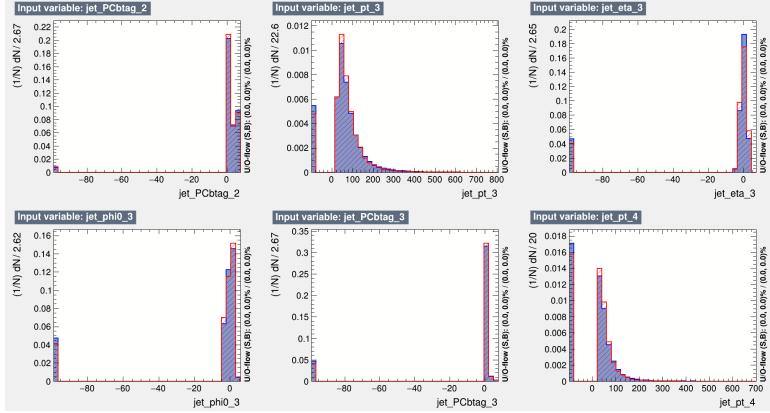


Figure A.13: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: pseudo-continuous b-tag score of second-highest btag-scoring jet, p_T of third-highest b-tag scoring jet, η of third-highest b-tag scoring jet, ϕ of third-highest btag-scoring jet (measured with respect to the Higgs candidate), pseudo-continuous b-tag score of third-highest btag-scoring jet, p_T of fourth-highest b-tag scoring jet.

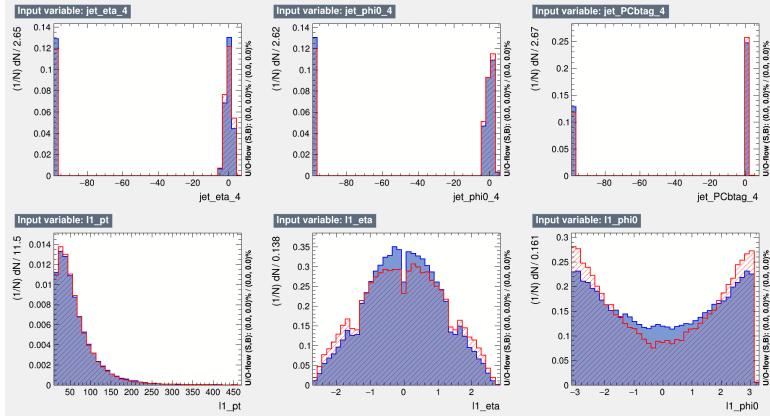


Figure A.14: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: η of fourth-highest btag-scoring jet, ϕ of fourth-highest btag-scoring jet (measured with respect to the Higgs candidate), pseudo-continuous b-tag score of fourth-highest btag-scoring jet, p_T of leading lepton, η of leading lepton, ϕ of leading lepton (measured with respect to the Higgs candidate).

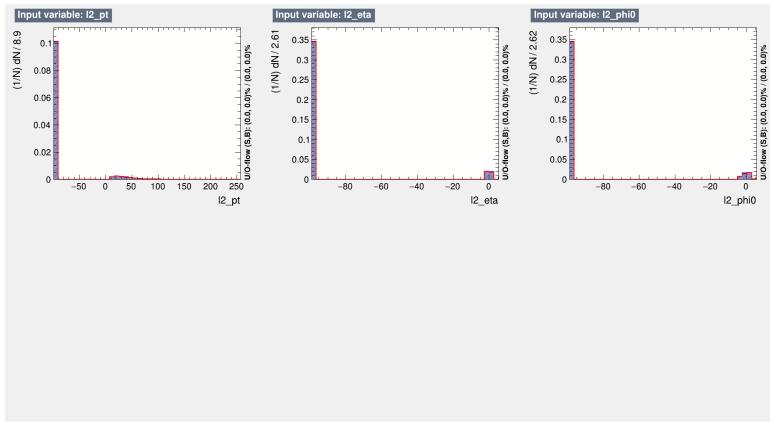


Figure A.15: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right: p_T of sub-leading lepton, η of sub-leading lepton, and ϕ of sub-leading lepton (measured with respect to Higgs candidate).

A.3.1.3 Results

We perform a significance comparison between the four-vector and an early iteration of the nominal CP-BDT, recreating the four-vector CP-BDT in xgboost so as to compare both with similar hyperparameter setups. For this study, we use a modified definition of number-counting significance, treating SM $tHjb$ and tWH as background processes and neglecting the modified top Yukawa coupling's effects on ggH .

We compare to a Nominal BDT with the following architecture:

In the hadronic channel:

- p_T (scaled by mass) and η of the Higgs candidate
- p_T, η, ϕ (wrt. Higgs candidate), and BDT score of the first and second reconstructed hadronic tops (see Section ??). In the case where no second top is reconstructed, a missing value is passed to XGBoost.
- Angles $\Delta\eta$ and $\Delta\phi$ between the top candidates. In the case where no second top is reconstructed, a missing value is passed to XGBoost.
- Two-object invariant masses m_{t1H} , m_{t2H} , and m_{t1t2} . In the case where no second top is reconstructed, a missing value is passed to XGBoost.
- Three-object invariant mass m_{t1t2H} . In the case where no second top is reconstructed, $m_{t1t2H} = m_{t1H}$.
- Jet multiplicity and b-jet multiplicity (77% working point)
- $H_T = \sum_{\text{jet } j} p_T^j$
- $\cos\theta^*$
- Missing E_T significance, $E_T^{\text{miss}}/\sqrt{H_T}$

In the leptonic channel:

- p_T (scaled by mass) and η of the Higgs candidate
- p_T, η, ϕ (wrt. Higgs candidate), and BDT score of the first and second reconstructed hadronic tops (see Section ??). In the case of dilepton events, or if no second top is reconstructed, a missing value is passed to XGBoost.
- Angles $\Delta\eta$ and $\Delta\phi$ between the top candidates. In the case of dilepton events, or if no second top is reconstructed, a missing value is passed to XGBoost.

- Two-object invariant masses m_{t1H} , m_{t2H} , and m_{t1t2} . In the case of dilepton events, or if no second top is reconstructed, a missing value is passed to XGBoost.
- Three-object invariant mass m_{t1t2H} . In the case of dilepton events, a missing value is passed to XGBoost. If no second top is reconstructed, $m_{t1t2H} = m_{t1H}$.
- Jet multiplicity and b-jet multiplicity (77% working point)
- $H_T = \sum_{\text{jet } j} p_T^j$
- $\cos \theta^*$
- Missing E_T significance, $E_T^{\text{miss}}/\sqrt{H_T}$

The results of this comparison are shown in Table ??

	Nominal	Alternative
Hadronic ROC AUC	0.717	0.723
Leptonic ROC AUC	0.708	0.718
$t\bar{t}H$ significance (if even)	4.57σ	4.60σ
$t\bar{t}H$ significance (if odd)	3.26σ	3.33σ
CP odd rejection	3.01σ	2.89σ
CP mix rejection	0.61σ	0.59σ

Table A.2: Figures of merit for the fifteen-category CP BDT categorization. The right-hand column shows that an alternative setup using four-vector training variables in the CP BDT achieves similar sensitivity.

Ultimately, these results show that, even without the presence of CP-alternative tH samples, the top reconstruction-aided "Nominal" BDT is able to produce a higher Number-Counting Significance than the proposed 4-vector only BDT. As is shown in Figures ?? - ??, the difference is primarily due to the Nominal BDT's ability to reduce ggH contamination by relying on top variables.

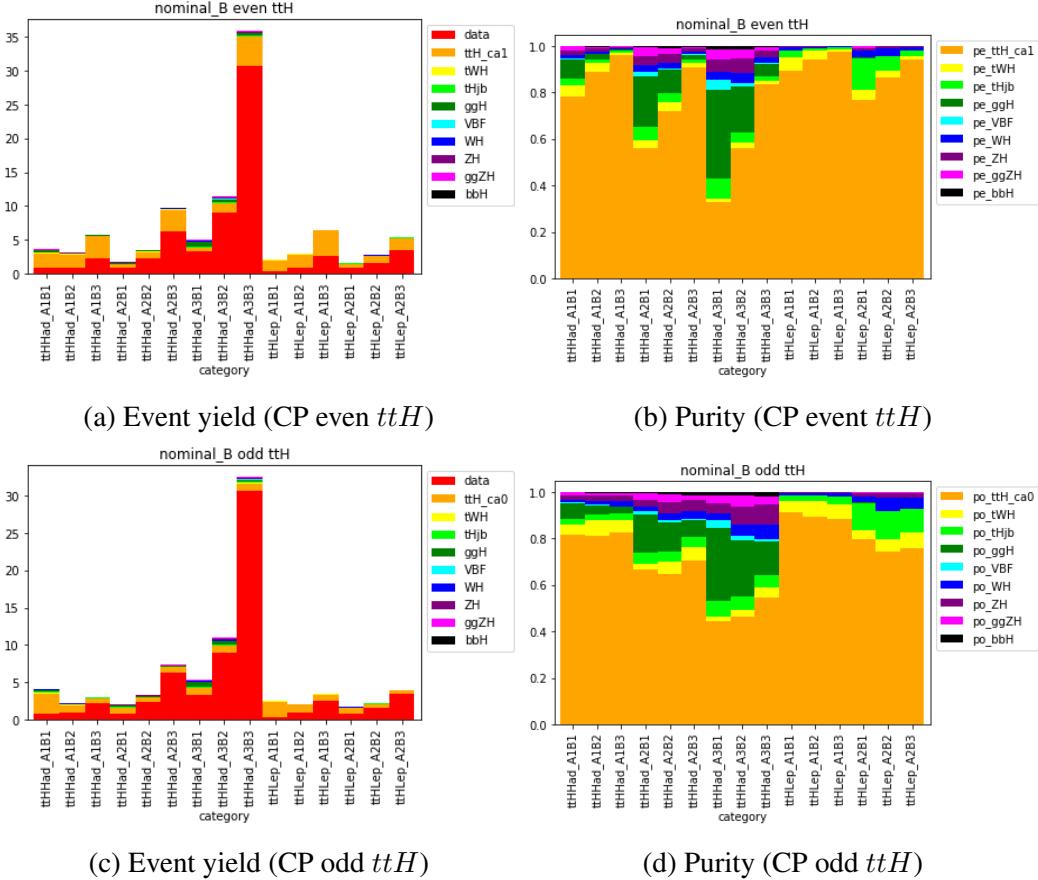


Figure A.16: (Left) Event yields in the CP categories at 139 fb^{-1} , with optimized A-boundaries drawn in the BDT score, using the Nominal CP-BDT. Shown separately for CP even ttH (top) and CP odd ttH (bottom). (Right) purity of the Higgs yield in each category for CP even ttH (top) and CP odd ttH (bottom).

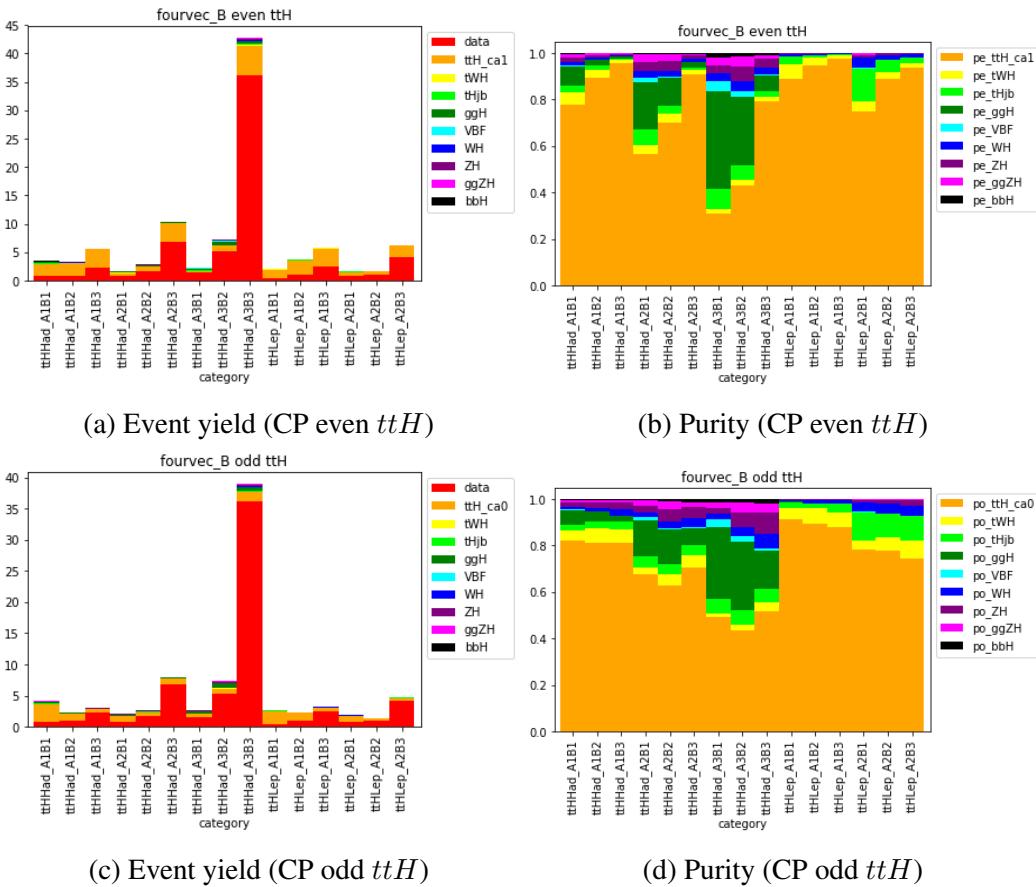


Figure A.17: (Left) Event yields in the CP categories at 139 fb^{-1} , with optimized A-boundaries drawn in the BDT score, using the 4-vector CP-BDT. Shown separately for CP even $t\bar{t}H$ (top) and CP odd $t\bar{t}H$ (bottom). (Right) purity of the Higgs yield in each category for CP even $t\bar{t}H$ (top) and CP odd $t\bar{t}H$ (bottom).

A.3.2 Dilep/Semilep BDT, ttH Only

An additional pair of BDTs are developed with the goal of subdividing the leptonic channel into a dedicated semileptonic and dileptonic channel. In order to fully exploit the effectiveness of this technique, a dedicated SBBDT would need to be trained for dileptonic and semileptonic channels as well; however, this is outside the scope of the analysis at this time. Additionally, low event yields in TI sidebands in the dileptonic channel (8 events passing preselection) call the feasibility of a multi-category fit using this method into question; however, this study may prove useful if this analysis is performed again at the HL-LHC. Dileptonic $t\bar{t}H$ offers special sensitivity to the top Yukawa coupling due to spin-correlations between the leptonic top decay products, so performing such a category division in the future is well-motivated [?].

A.3.2.1 Dileptonic BDT

For the dedicated dileptonic BDT, trained and tested on events which pass the leptonic preselection in addition to requiring that the event contain at least two leptons, we use the following inputs:

- p_T , η , ϕ , and pseudocontinuous b-tag score of the leading jet in the event
- p_T , η , ϕ , and pseudocontinuous b-tag score of the leading 2 leptons in the event
- The p_T of the Higgs candidate, scaled by its mass
- $\cos(\theta^*)$
- The p_T and η of the two leading photons in the event
- The magnitude and ϕ of the missing transverse energy in the event
- The summed invariant mass of all jets in the event
- The minimum ΔR between a photon and a jet in the event
- The ΔR between the two leptons in the event

The linear correlations between these variables in $t\bar{t}H$ CP even and CP odd aMCnlo+Pythia8 Monte Carlo are shown in Figure ?? . Figures ?? - ?? compare the distribution of each training variable in $t\bar{t}H$ CP even and CP odd Monte Carlo. We report a ROC-AUC of 0.707 for this BDT.

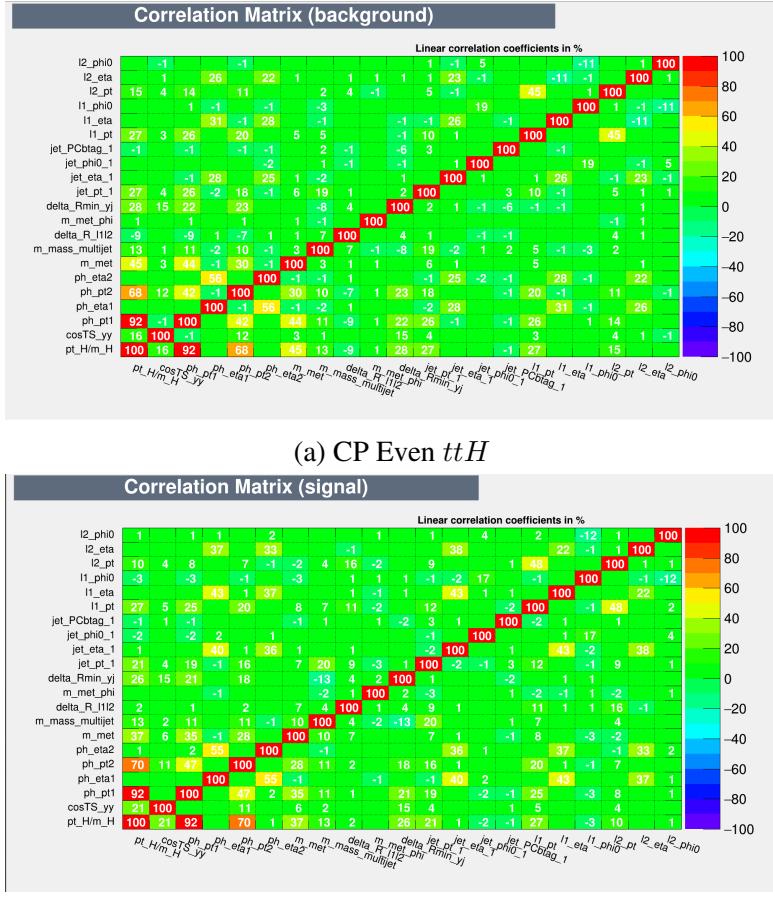


Figure A.18: Training variable correlations for events passing dileptonic pre-selection.

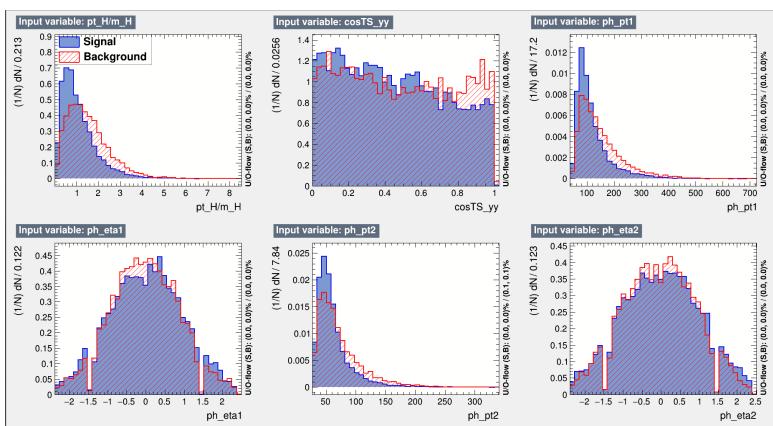


Figure A.19: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd $t\bar{t}H$ is denoted as "signal" (blue); CP-even $t\bar{t}H$ is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: Higgs candidate p_T (scaled by mass), $\cos(\theta^*)$, leading photon p_T , leading photon η , subleading photon p_T , subleading photon η .

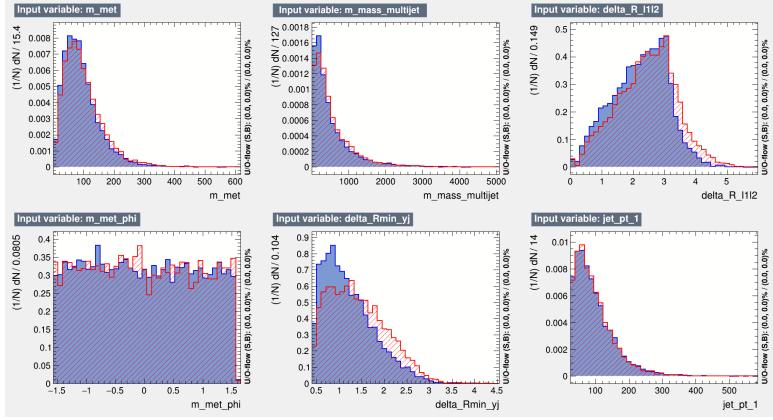


Figure A.20: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: Magnitude of MET, summed invariant mass of all jets in the event, ΔR between the two leptons present in the event, MET ϕ (branch cut chosen to range from $-\pi/2$ to $\pi/2$), minimum ΔR between a photon and a jet, p_T of highest b-tag scoring jet

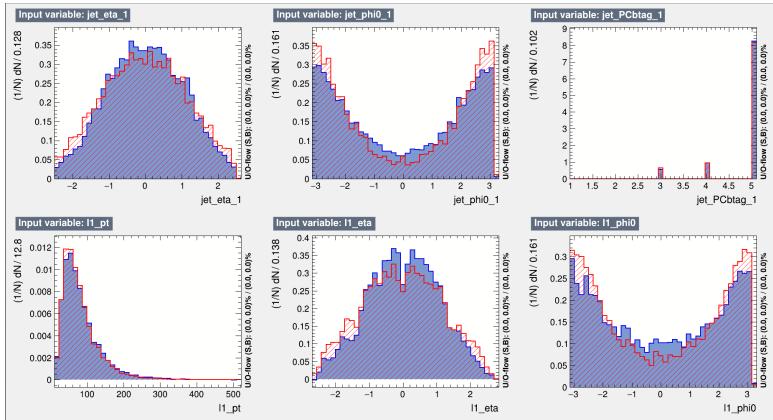


Figure A.21: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: η of highest b-tag scoring jet, ϕ of highest b-tag scoring jet (measured with respect to the Higgs candidate), pseudo-continuous b-tag score of highest b-tag scoring jet, p_T of leading lepton, η of leading lepton, ϕ of leading lepton (measured with respect to the Higgs candidate)

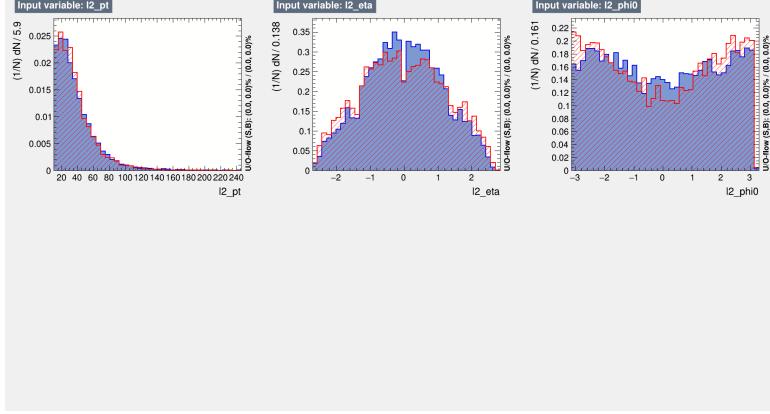


Figure A.22: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: p_T of subleading lepton, η of subleading lepton, ϕ of subleading lepton (measured with respect to the Higgs candidate)

A.3.2.2 Semileptonic BDT

For the dedicated semileptonic BDT, trained and tested on events which pass the leptonic preselection in addition to requiring that the event contain exactly one lepton, we use the following inputs:

- p_T , η , ϕ , and pseudocontinuous b-tag score of the leading 4 jets in the event
- p_T , η , ϕ , and pseudocontinuous b-tag score of the leading lepton in the event
- The p_T of the Higgs candidate (scaled by mass)
- $\cos(\theta^*)$
- The η and ϕ of the two leading photons in the event
- The magnitude and ϕ of the missing transverse energy in the event

The linear correlations between these variables in tH CP even and CP odd aMCnlo+Pythia8 Monte Carlo are shown in Figure ??-. Figures ?? - ?? compare the distribution of each training variable in tH CP even and CP odd Monte Carlo. We report a ROC-AUC of 0.725 for this BDT.

A.3.3 Variable Optimization Studies, Nominal BDT, ttH and tH

We use TMVA in order to perform a number of studies to investigate possible modifications to the nominal CPBDT, using CP-Odd tH + tH as signal and CP-Even tH + tH as background. For each study, we construct a different "Baseline" BDT, starting with the nominal CPBDT, then

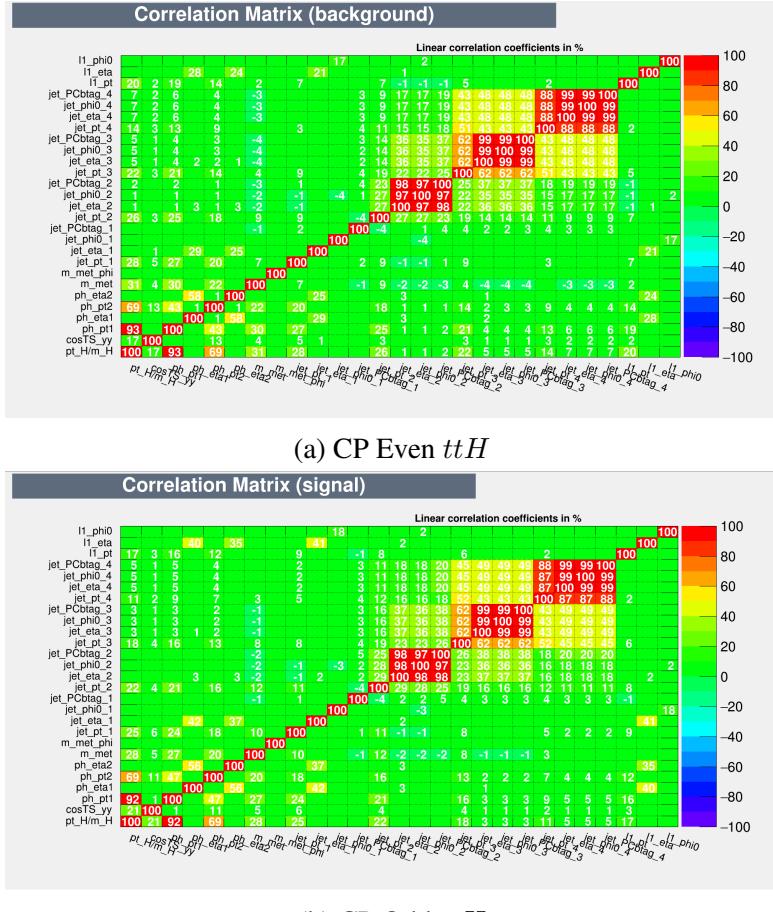


Figure A.23: Training variable correlations for events passing semileptonic pre-selection.

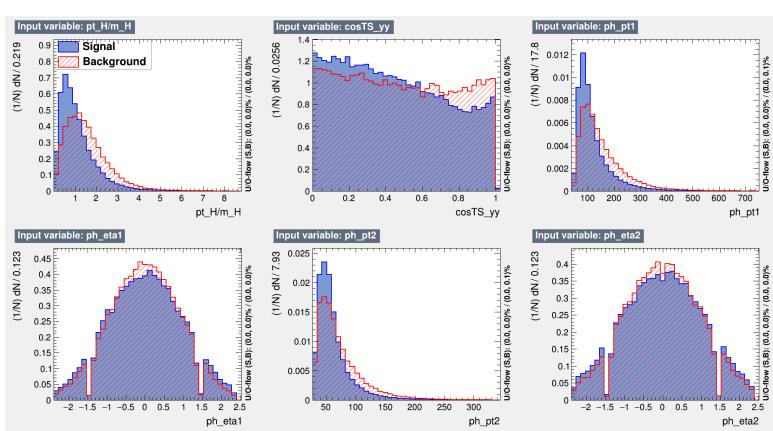


Figure A.24: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: Higgs candidate p_T (scaled by mass), $\cos(\theta^*)$, leading photon p_T , leading photon η , subleading photon p_T , subleading photon η .

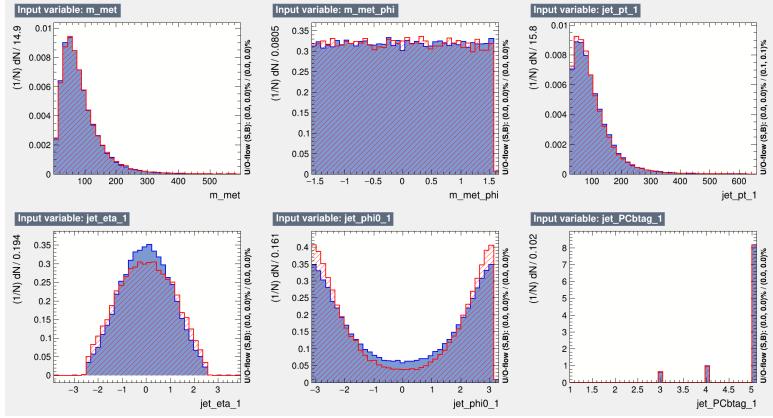


Figure A.25: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: Magnitude of the event MET, MET ϕ (branch cut chosen to range from $-\pi/2$ to $\pi/2$), p_T of highest b-tag scoring jet, η of highest b-tag scoring jet, ϕ of highest btag-scoring jet (measured with respect to the Higgs candidate), pseudo-continuous b-tag score of highest btag-scoring jet

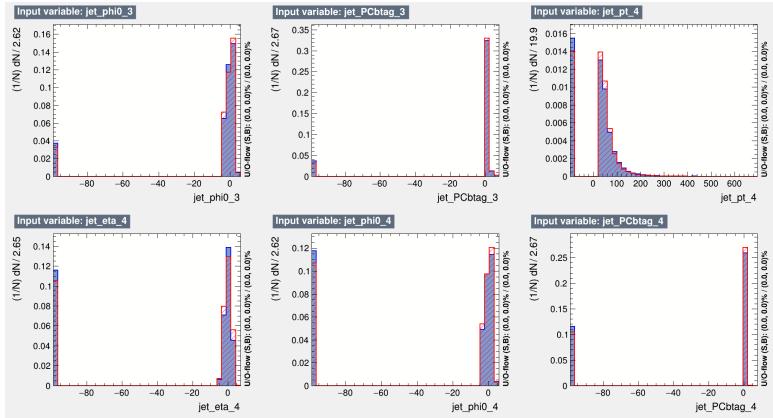


Figure A.26: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: p_T of second-highest b-tag scoring jet, η of second-highest b-tag scoring jet, ϕ of second-highest btag-scoring jet (measured with respect to the Higgs candidate), pseudo-continuous b-tag score of second-highest btag-scoring jet, p_T of third-highest b-tag scoring jet, η of third-highest b-tag scoring jet

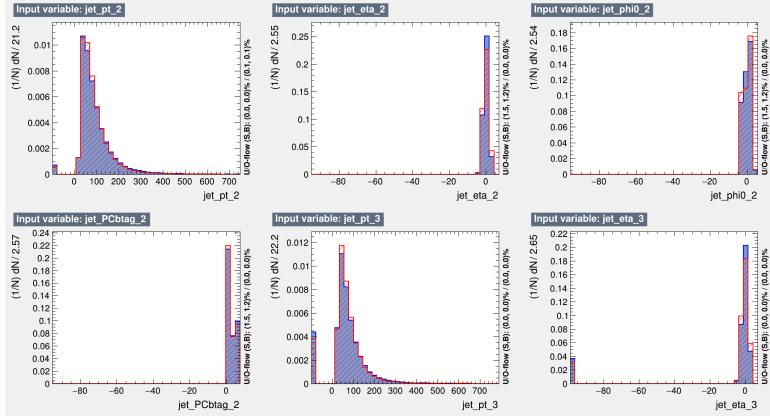


Figure A.27: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: ϕ of third-highest btag-scoring jet (measured with respect to the Higgs candidate), pseudo-continuous b-tag score of third-highest btag-scoring jet, p_T of fourth-highest b-tag scoring jet, η of fourth-highest b-tag scoring jet, ϕ of fourth-highest btag-scoring jet (measured with respect to the Higgs candidate), pseudo-continuous b-tag score of fourth-highest btag-scoring jet

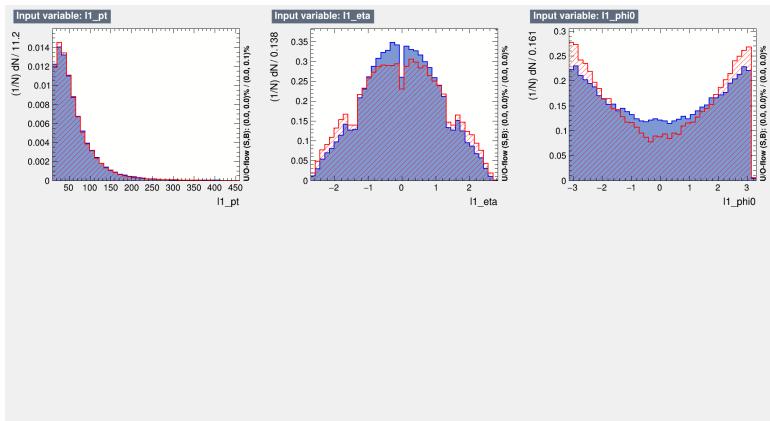


Figure A.28: Normalized training variables for the 4-vector BDT, output by TMVA. CP-odd ttH is denoted as "signal" (blue); CP-even ttH is denoted as "background" (red). Variables shown are, from left to right, top row to bottom row: p_T of leading lepton, η of leading lepton, ϕ of leading lepton (measured with respect to the Higgs candidate)

observe whether or not adding or altering a given variable will substantially impact the CPBDT performance (deeming a "substantial" impact to be one that increases $tH + tH$ Number-Counting Significance by at least 0.1σ , as gains below this threshold are likely to be noise).

We begin with a nominal BDT consisting of the following input variables and modify it in an iterative process, adding or removing variables depending on performance:

- p_T (scaled by mass) and η of the Higgs candidate
- p_T, η, ϕ (wrt. Higgs candidate), and BDT score of the first and second reconstructed hadronic tops (see Section ??). In the case where no second top is reconstructed, we use the p_T, η , and ϕ of the "hybrid" top.
- Angles $\Delta\eta$ and $\Delta\phi$ between the top and the "hybrid" top.
- Two-object invariant masses m_{t1H}, m_{Hhy} , and m_{t1hy}
- The minimum ΔR between a photon and a jet in the event

We come to the following conclusions:

- Adding the p_T, η, ϕ , and pseudocontinuous b-tag score of the leading N jets in the event for $N = 6$, where jets are ordered by p_T does not substantially impact the significance (the largest significance gain from their inclusion was 0.07σ).
- Adding the p_T, η, ϕ , and pseudocontinuous b-tag score of the leading N jets in the event for $N = 6$, where jets are ordered by b-tag score does not substantially impact the significance (the largest significance gain from their inclusion was approximately 0.06σ).
- Changing $\Delta\eta(tops, Higgs)$ and $\Delta\phi(tops, Higgs)$ to $\Delta R(tops, Higgs)$ does not substantially impact the significance (the largest significance gain from this change was approximately 0.07σ).
- Changing the choice of two or more of M_{t1H}, M_{t1hy} , or M_{Hhy} does not substantially impact the significance (all combinations performed almost identically). We note that, because the BDT could in theory learn M_H from combining more than two of these, we recommend only including two, the best-performing pair of which is M_{t1hy} and M_{t1H} .
- Adding the $p_T/M_H, \eta$, and ϕ of the two photons in the event does not substantially impact the significance (the largest significance gain from this change was approximately 0.02σ).
- Adding the p_T, η , and ϕ of the two leading leptons in the event does not substantially impact the significance (the largest significance gain from this change was approximately 0.01σ).

- Adding the missing transverse energy and its azimuthal angle does not substantially impact the significance (the largest significance gain from their inclusion was 0.07σ).
- Adding the minimum or second-smallest ΔR between a photon and a jet in the event to the BDT does substantially impact the significance (the largest significance gain from their inclusion was 0.105σ in the hadronic channel)

We also recommend that, due to its potential dependence on m_H , $\cos(\theta^*)$ should not be included in the final CP-BDT.

A.3.4 Dedicated CP-45 BDT Training

We use TMVA to investigate the performance of a CP-BDT trained against the $\alpha = 45^\circ$ maximal-mixing CP signal sample, rather than the $\alpha = 90^\circ$ CP-odd signal sample used for the Nominal BDT.

As in the previous section, we begin with a different "Baseline" set of BDT variables, starting with the nominal CPBDT, then observe whether or not adding or altering a given variable will substantially impact the BDT performance (deeming a "substantial" impact to be one that increases the $tH + tH$ Number-Counting Significance by at least 0.1σ , as gains below this threshold are likely to be noise). Now, however, we use the Number-Counting Significance at $\alpha = 45^\circ$, rather than at $\alpha = 90^\circ$, as defined in Equation ?? and Equation ??.

As in the previous section, we begin with a nominal $tH + tH$ CPBDT consisting of the following input variables and modify it in an iterative process, adding or removing variables depending on performance:

- p_T (scaled by mass) and η of the Higgs candidate
- p_T, η, ϕ (wrt. Higgs candidate), and BDT score of the first and second reconstructed hadronic tops (see Section ??). In the case where no second top is reconstructed, we use the p_T, η , and ϕ of the "hybrid" top.
- Angles $\Delta\eta$ and $\Delta\phi$ between the top and the "hybrid" top.
- Two-object invariant masses m_{t1H}, m_{Hhy} , and m_{t1hy}
- The minimum ΔR between a photon and a jet in the event

We repeat the same set of tests as in the prior section and come to the same conclusions, namely, that we see no significant variations in the CP-45 significance by deviating from the nominal CPBDT architecture.

Using the BDT trained against the $\alpha = 45^\circ$ maximal-mixing CP signal sample, we observe an inclusive CP-Odd significance of 2.38 and a CP-45 Significance of 1.00 in the hadronic channel, as well as a CP-Odd significance of 1.79 and a CP-45 Significance of 0.78 in the leptonic channel. This can be compared to the TMVA instantiation of the Nominal BDT trained against the CP-Odd signal sample, for which we observe an inclusive CP-Odd significance of 2.50 and a CP-45 Significance of 1.03 in the hadronic channel, as well as a CP-Odd significance of 1.85 and a CP-45 Significance of 0.77 in the leptonic channel. For all four of these significance metrics, we find that the BDT trained against the CP-Odd signal sample performs comparably to the BDT trained against the $\alpha = 45^\circ$ maximal-mixing CP signal sample. Thus, using only one α -point as the BDT signal model rather than many is determined to be a robust choice for the analysis.

In ATLAS analyses structured similarly to those in this dissertation, Monte Carlo statistics can often play a substantial part in driving systematic uncertainty. If the templates used to perform the spurious signal study contain a large number of statistical fluctuations, the spurious signal will often be an overestimate due to the tendency to fit statistical fluctuations in the templates as legitimate spurious signal. Though producing more Monte Carlo simulated samples would resolve this issue, it is often computationally intensive and inefficient.

In order to resolve this, a technique known as Gaussian Process Regression is implemented in the context of the Couplings analysis. This is implemented using the Gaussian Smoothing for BackGrounds (GaSBaG) tool first developed in ??, which interfaces with the Scikit-Learn [?] machine learning package.

A Gaussian Process (GP) is defined as a set of random processes of which all finite subsets have a multivariate normal distribution [?]. Provided statistics in each bin are high enough, a histogram encoding some underlying smooth distribution is an example of such a finite subset: each bin contains a normally-distributed number of events about some true mean, and the bins are correlated according to their covariance. Thus, a Gaussian process can be defined over such a histogram. The mean of the fitted Gaussian Process corresponds to the smooth underlying shape, while the elements of the GP covariance matrix correspond to the error in each bin and the correlation between bins.

The covariance matrix can be simplified using a kernel, which determines the form of the correlation between points (i.e., the length scales in X at which points are expected to influence one another in Y). Two useful kernels are the Radial Basis Function (RBF) kernel and the Gibbs kernel [?, ?].

The RBF kernel has one hyperparameter, the constant length scale l , and it is defined as

$$K_{\text{RBF}}(x, x') = \exp\left(\frac{-(x - x')^2}{2l^2}\right) \quad (\text{A.2})$$

The RBF kernel is useful for mostly-flat functions. However, for smoothly-falling functions,

it is likely that nearby points will be more correlated in some regions than in others, so a constant length scale is a suboptimal model. The Gibbs kernel allows the length scale $l(x)$ to vary linearly as a function of x , and thus has two hyperparameters: the initial length scale and the length scale slope. The Gibbs kernel function is:

$$K_{\text{Gibbs}}(x, x') = \frac{\sqrt{2l(x)l(x')}}{l(x)^2 + l(x')^2} \cdot \exp\left(\frac{-(x - x')^2}{l(x)^2 + l(x')^2}\right) \quad (\text{A.3})$$

The background templates used in the spurious signal test for the analysis categories are all smoothly falling distributions with statistical fluctuations. By utilizing a Gaussian Process Regression fit and then taking the mean of the Gaussian Process as the "true" template shape, it is possible to smooth out some amount of statistical noise, thus decreasing the spurious signal systematic. The GP smoothing technique makes no assumption on the form of the underlying distribution other than that it is smooth and falling, so the choice of functional form from the spurious signal test will not be biased.

The hyperparameters (initial length scale and length scale slope) are allowed to vary over a range specified by the user; the optimal hyperparameters within this range are determined by the Gaussian Process fitting procedure.

A GP is fit to the background template in each category. This can be modelled in a Bayesian manner: the GP is a distribution with a given prior mean that is then conditioned on the template histogram; the smoothed template is the mean of the posterior distribution ???. Due to the expected shape of the templates, the prior is defined as an exponential function with parameters obtained by a fit to the template. However, in cases where the input template has very few statistics, large errors on the data points are compatible with a steep exponential prior though the "true" underlying shape is approximately flat, which can confound the fitter and cause it to simply output the prior distribution. Therefore, a "switch" has been added to re-perform the GP fit using a flat prior in cases where the resulting GP shape and the prior exponential shape disagree with a $\chi^2/\text{DoF} < 0.1$.

The mean of the GP posterior is then saved as the new template (with bin-by-bin errors corresponding to the bin-by-bin errors of the original template). The spurious signal test is then performed with these new distributions.

Validation tests were performed with the GP smoothing technique in order to quantify whether or not GPR induces a bias (that is, artificially reduces the spurious signal beyond just smoothing out statistical fluctuations). These tests primarily use "toy" templates- randomly-generated background templates constructed using a known analytic function as a probability density. One of the goals of these validation studies is to determine a safe statistical regime in which to apply GPR, as in the very low-statistics regime, the distribution of events in each bin follows a Poisson distribution rather than a Gaussian distribution.

We find that GPR remains effectively unbiased for smoothly-falling templates containing more than an average of 20 Monte Carlo events/ bin. The procedures to validate the GPR smoothing are reported in Appendix ??.

A.4 Gaussian Processes smoothed background templates

The background templates of all of the Coupling analysis categories, both before and after the Gaussian Processes (GP) smoothing, are presented in Figures ??–??. The data sidebands, corresponding to 139 fb^{-1} , are shown for comparison, although the GP smoothing technique does not take them into account.

As detailed in ??, the GPR method chosen for use in the Couplings analysis involves extending our templates by 5 GeV on either side to reduce edge effects, as well as the choice of a linear error kernel in order to properly handle template errors.

A.4.1 Spurious Signal GPR-smoothed templates

The SS test was completely re-run with the templates reported in Figures ??–??. We report two sets of results - first, we record the spurious signal extracted from the smoothed templates using the functional form chosen from performing the relaxed spurious signal test on the unsmoothed templates. The results are reported in Table ?? and Table ?? . A comparison with the nominal un-smoothed SS test reported in Section ?? is presented in Table ?? and Table ??.

Second, we record the spurious signal from the smoothed templates using the functional form chosen from performing a non-relaxed spurious signal test on the smoothed templates only (that is, removing the potential two-sigma fluctuation). The results are reported in Table ?? and Table ?? . A comparison with the nominal un-smoothed SS test reported in Section ?? showing the choice of functional form and extracted SS is presented in Table ?? and Table ??.

In categories where GPR is deemed unreliable due to low statistics, we put an N/A rather than numerical values.

A.5 Validation of Gaussian Process Smoothing

A series of tests was performed to determine whether or not the GPR technique induces a bias on the spurious signal, expanding on some of the methodology of ??.

Two sets of bias tests are performed: the first is a measurement of the bias using toy templates pulled from smooth functions (i.e., the bias measure relevant to the Couplings analysis), and the

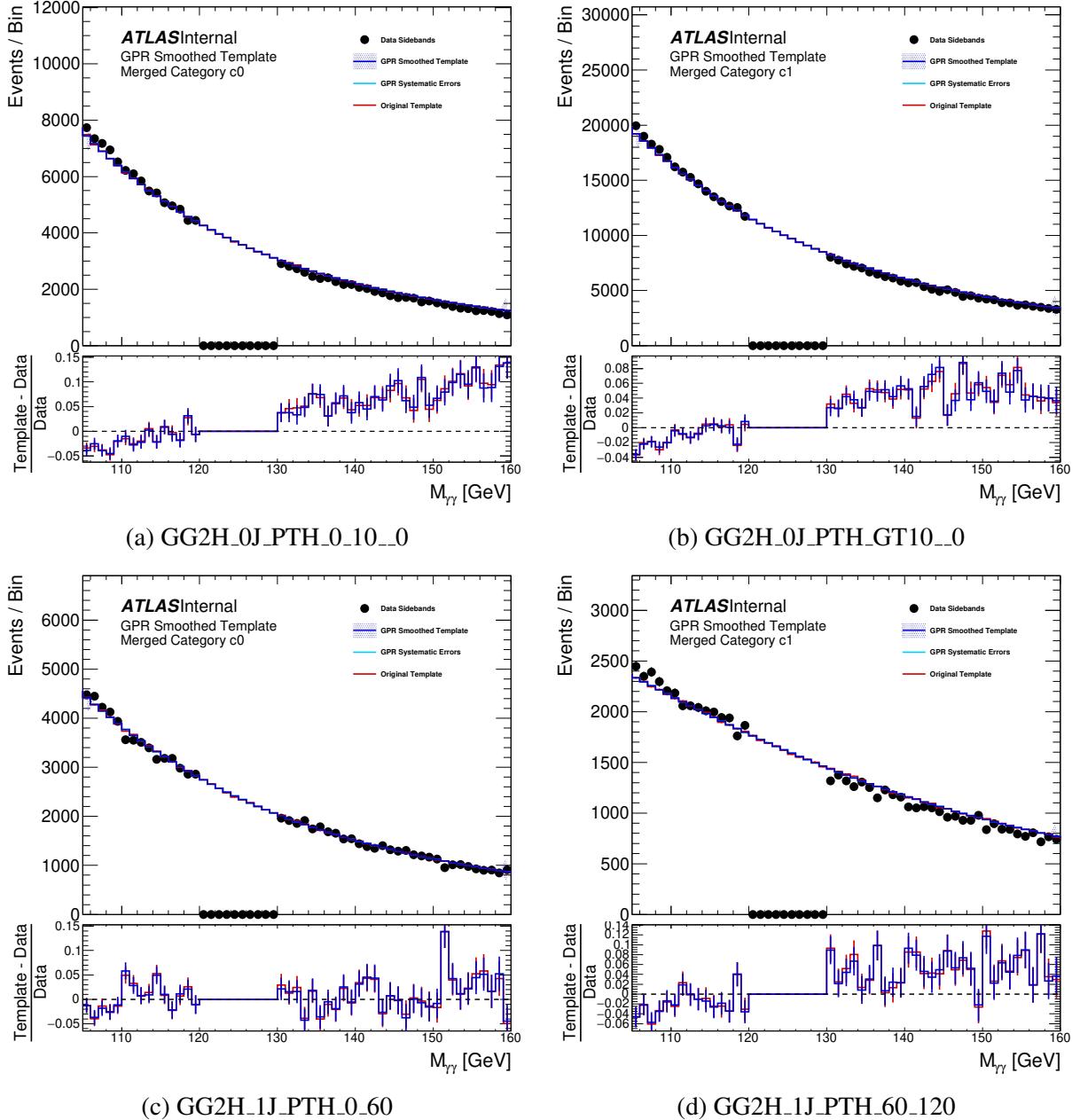


Figure A.29: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

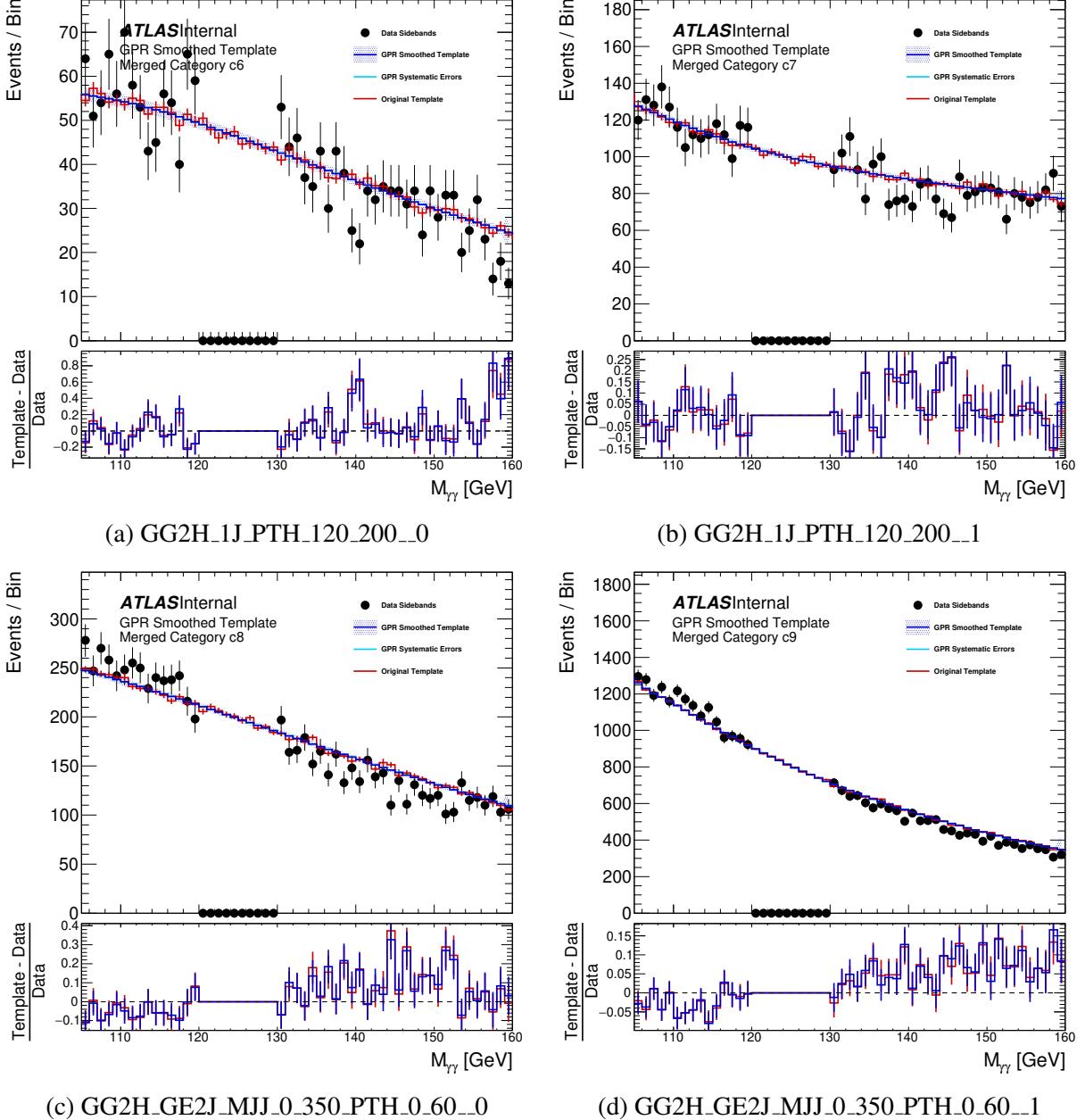


Figure A.30: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

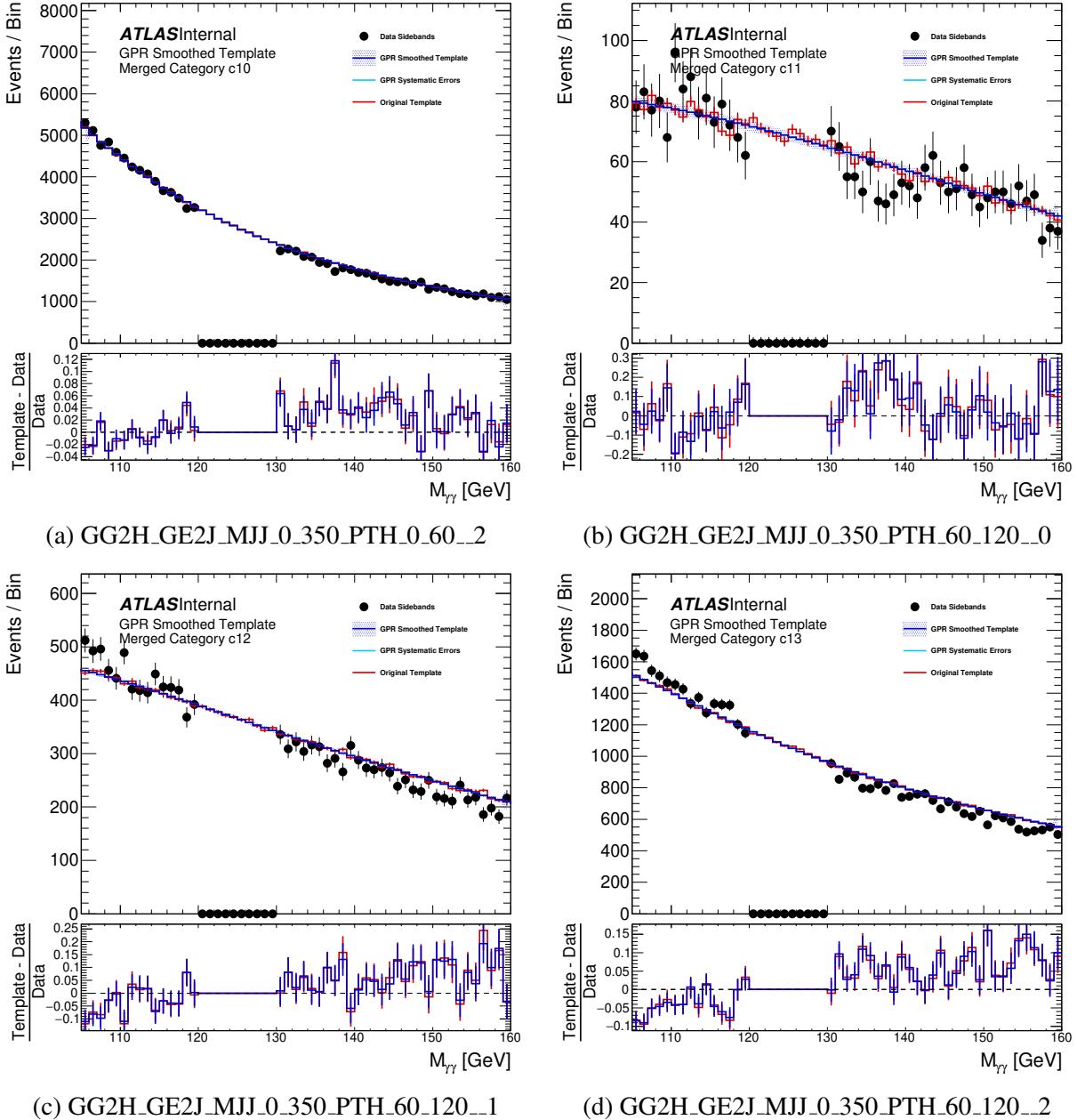


Figure A.31: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

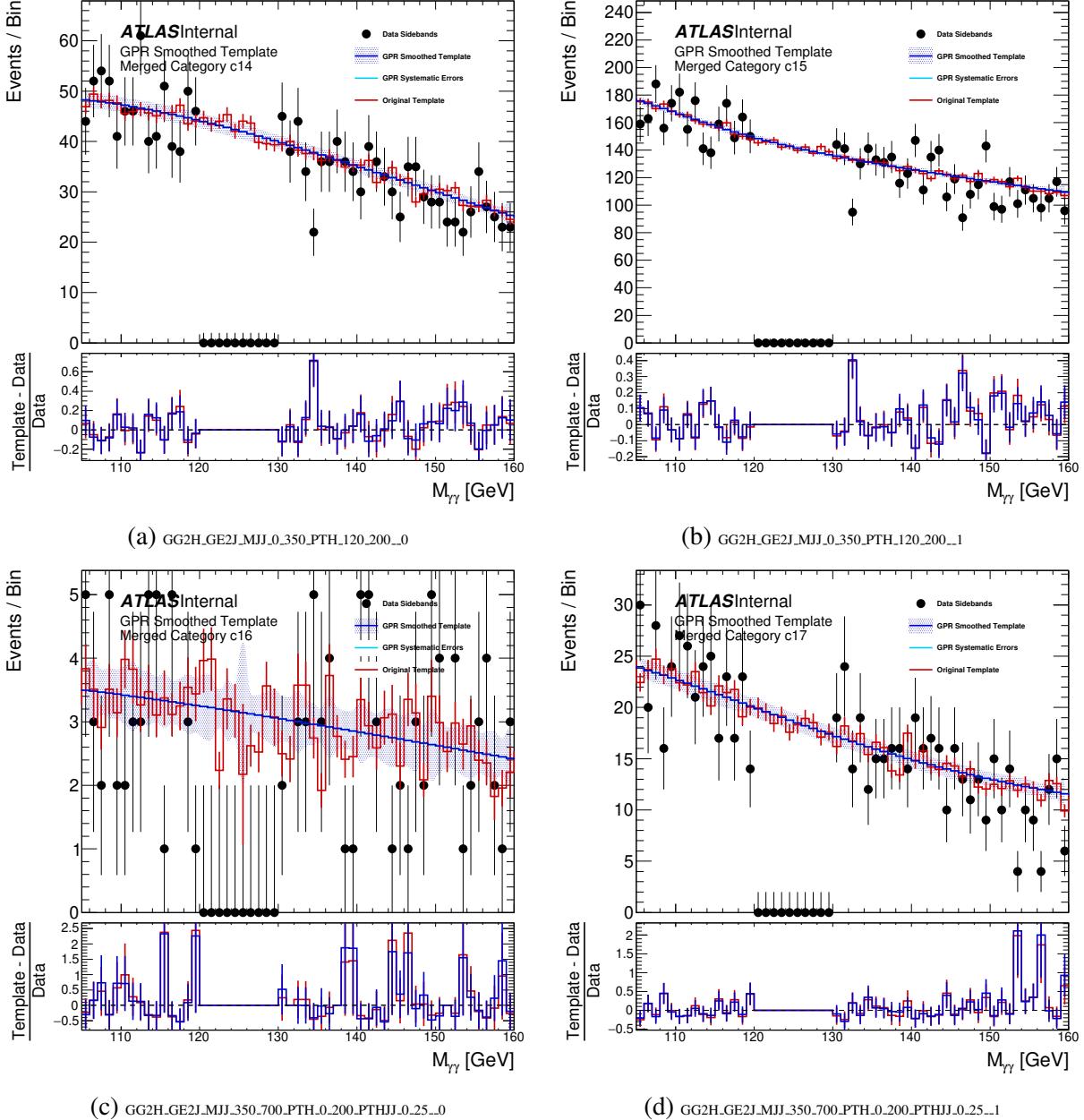


Figure A.32: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

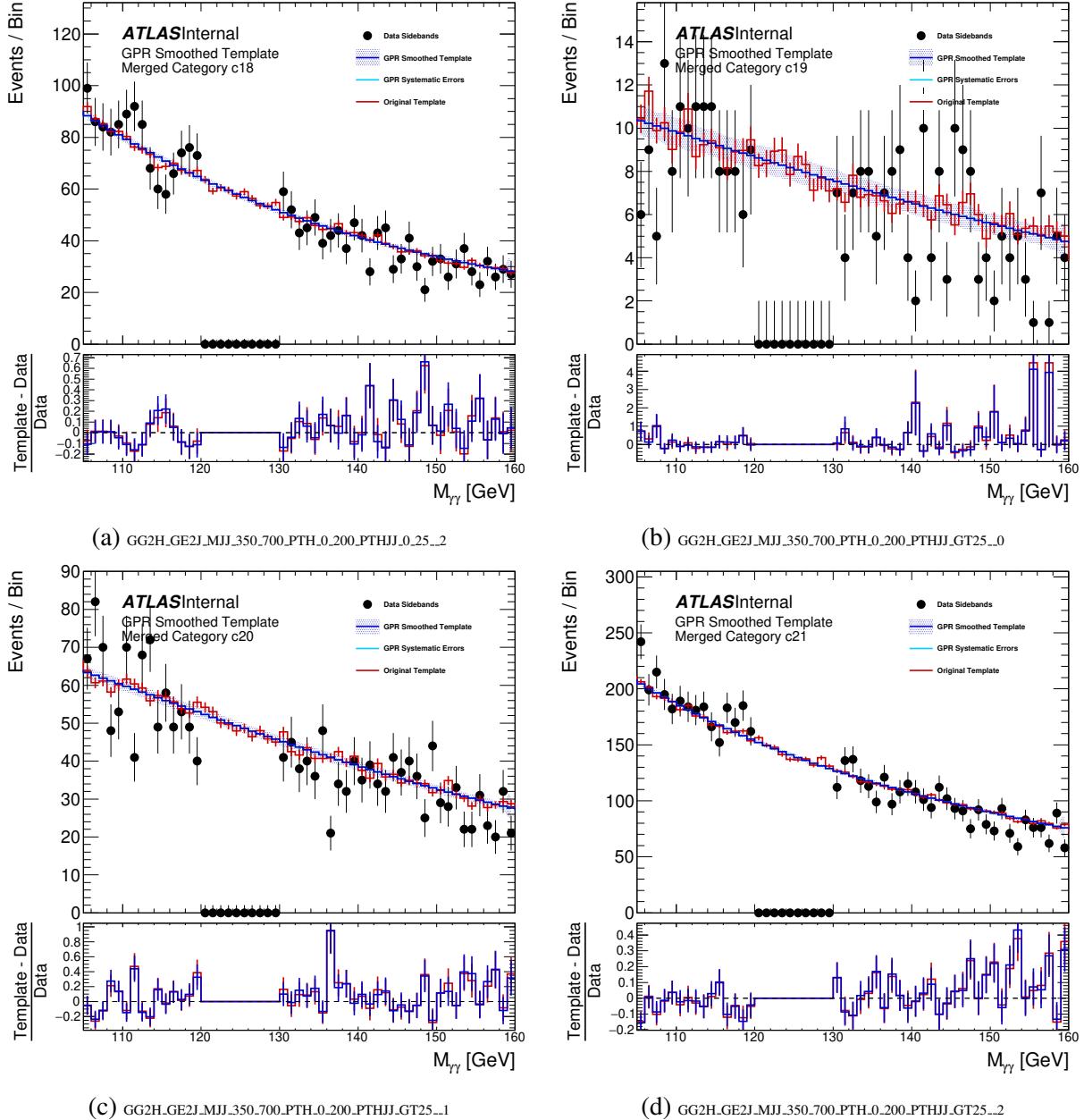


Figure A.33: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

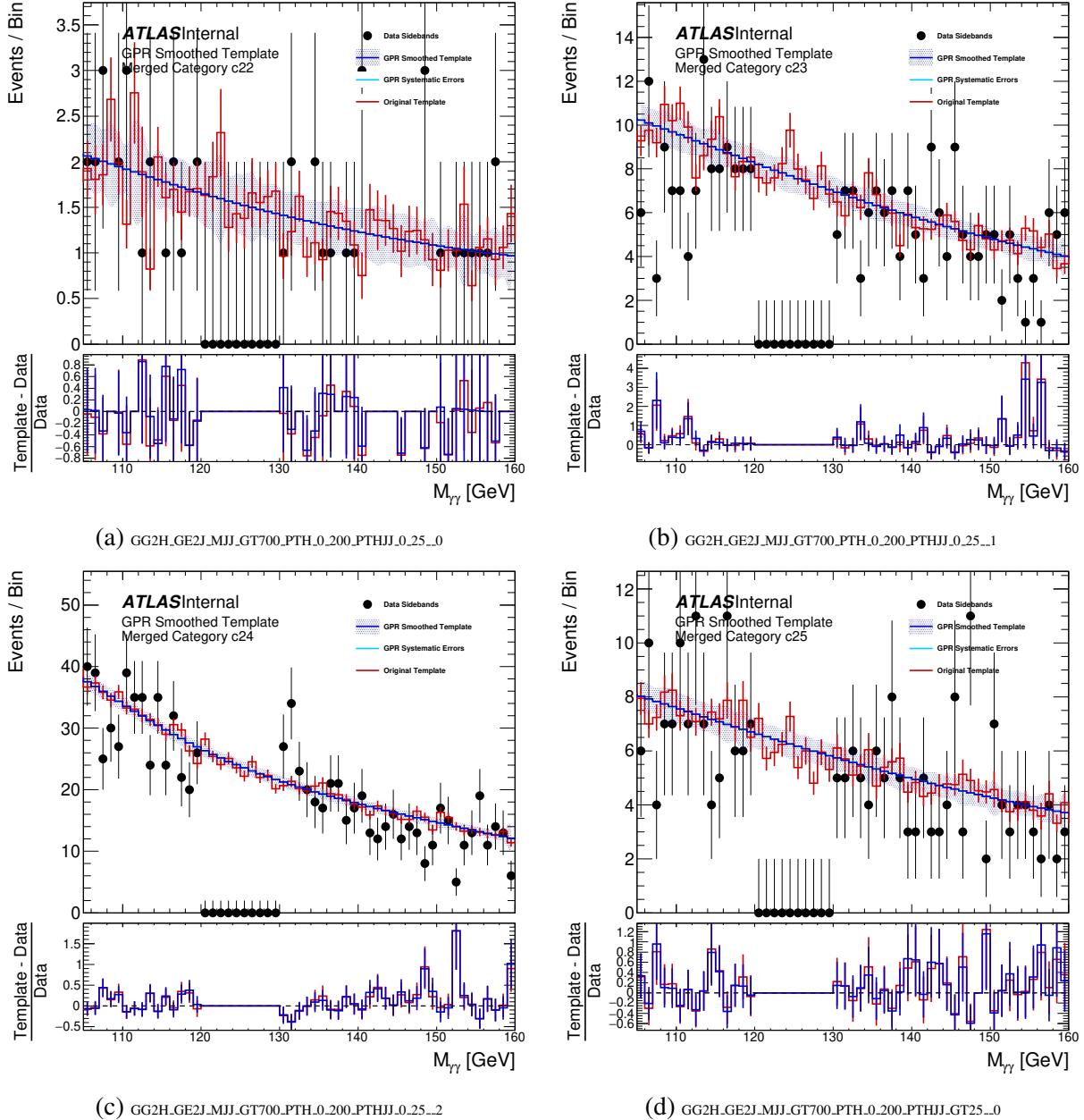


Figure A.34: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

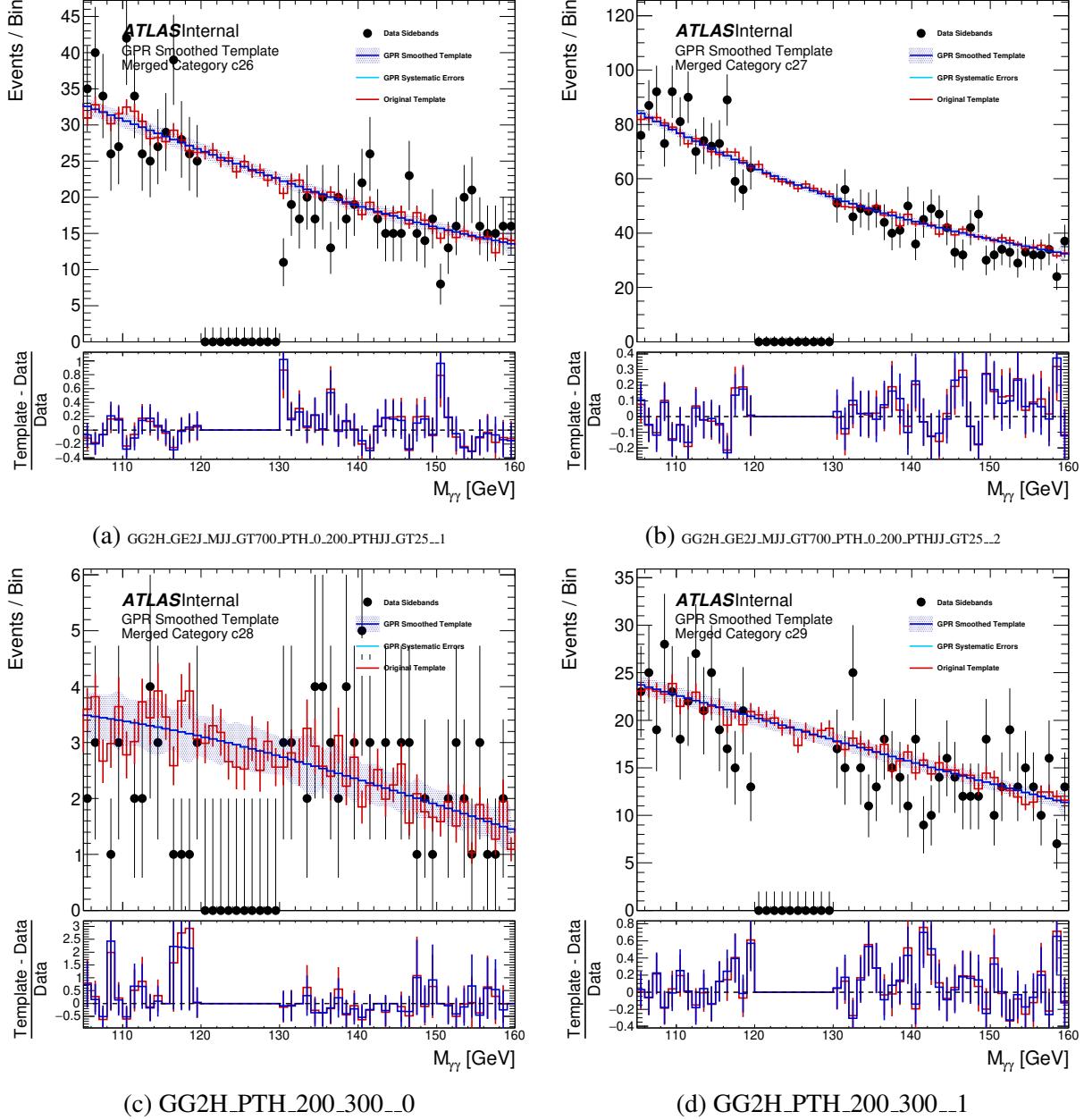


Figure A.35: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

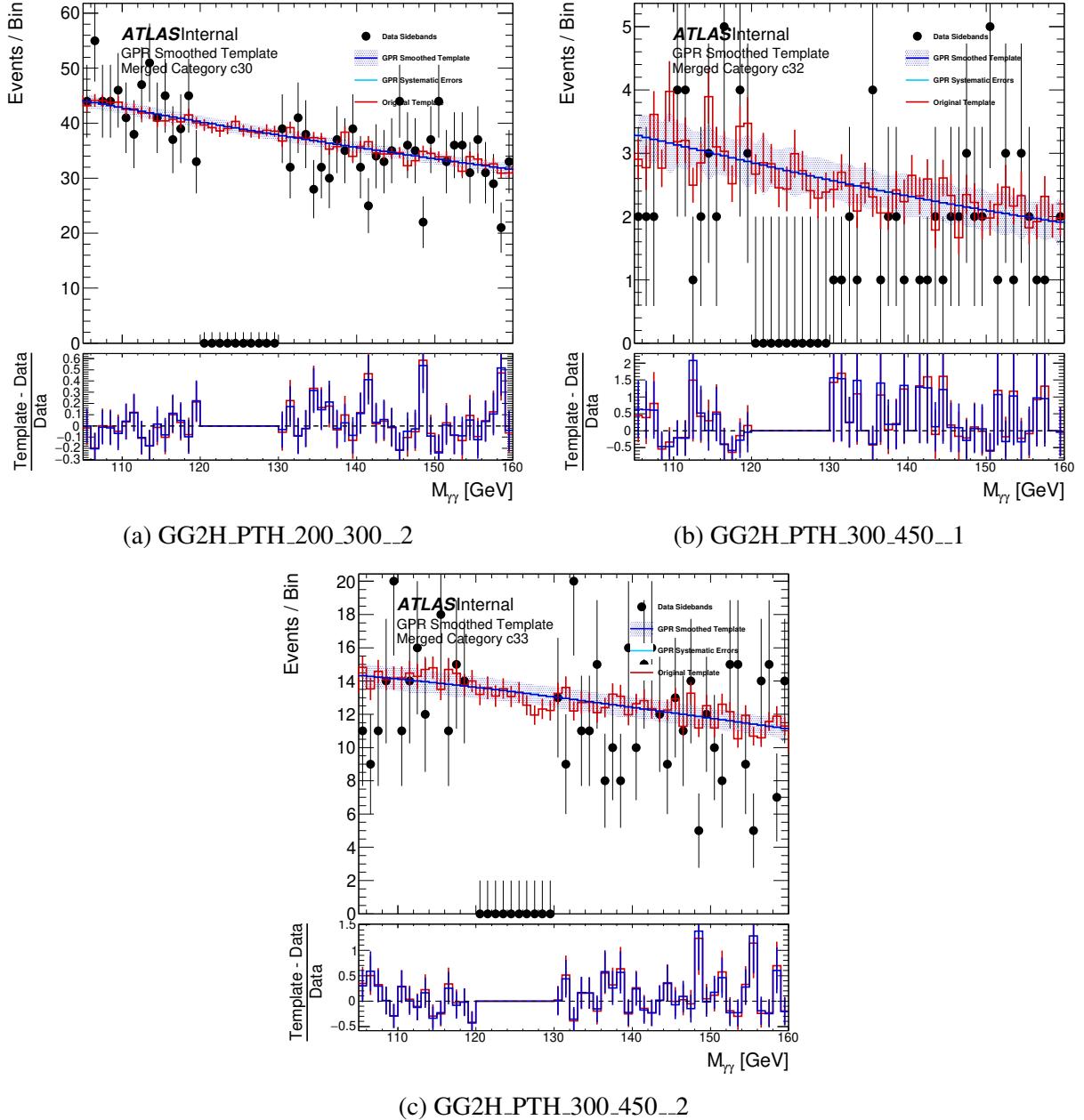
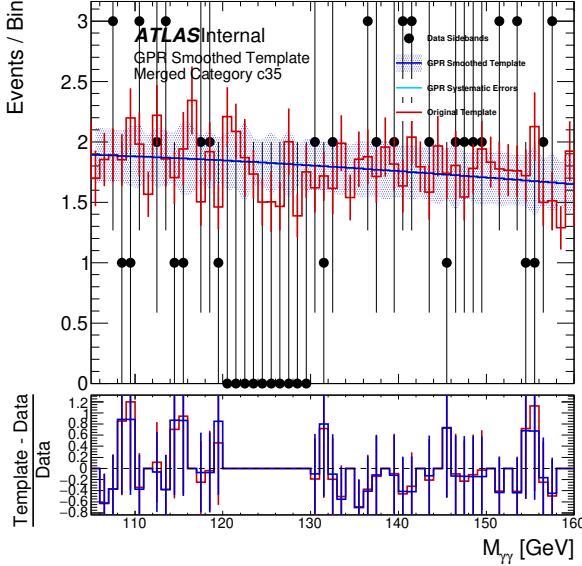


Figure A.36: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.



(a) GG2H.PTH_450_650_-1

Figure A.37: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

second is a measure of the bias using toy templates with a signal-like feature injected (i.e., understanding the GPR’s tendency to smooth away true underlying features- this does not apply to the Couplings analysis, but provides an important academic estimate of GPR’s limitations in other scenarios). Because it is assumed the underlying distribution of the data is not perfectly modelled by any existing function, a different closely-related function is used as the best-fit template function to perform the spurious signal measurement (Exponential for templates generated with ExpPoly2, PowerLaw1 for templates generated with an Exponential, ExpPoly2 for templates generated with ExpPoly3). This induces a ”true” spurious signal- if there were no such functional mismatch, the ”true” spurious signal would be zero.

First, a smoothly-falling template is produced. For all but the initial feature-width check in the injection study, we use the full Run-2 SHERPA diphoton sample, weighted by the mass of the diphoton system to produce a variety of shapes.

$$\begin{aligned}
 low &= (0 * (160 - (m_{\gamma\gamma}/1000)))/(160 - 105) + 1 \\
 med &= (0.5 * (160 - (m_{\gamma\gamma}/1000)))/(160 - 105) + 1 \\
 high &= (1 * (160 - (m_{\gamma\gamma}/1000)))/(160 - 105) + 1
 \end{aligned} \tag{A.4}$$

Each template is then weighted to have an average event weight of 0.04 (roughly the average

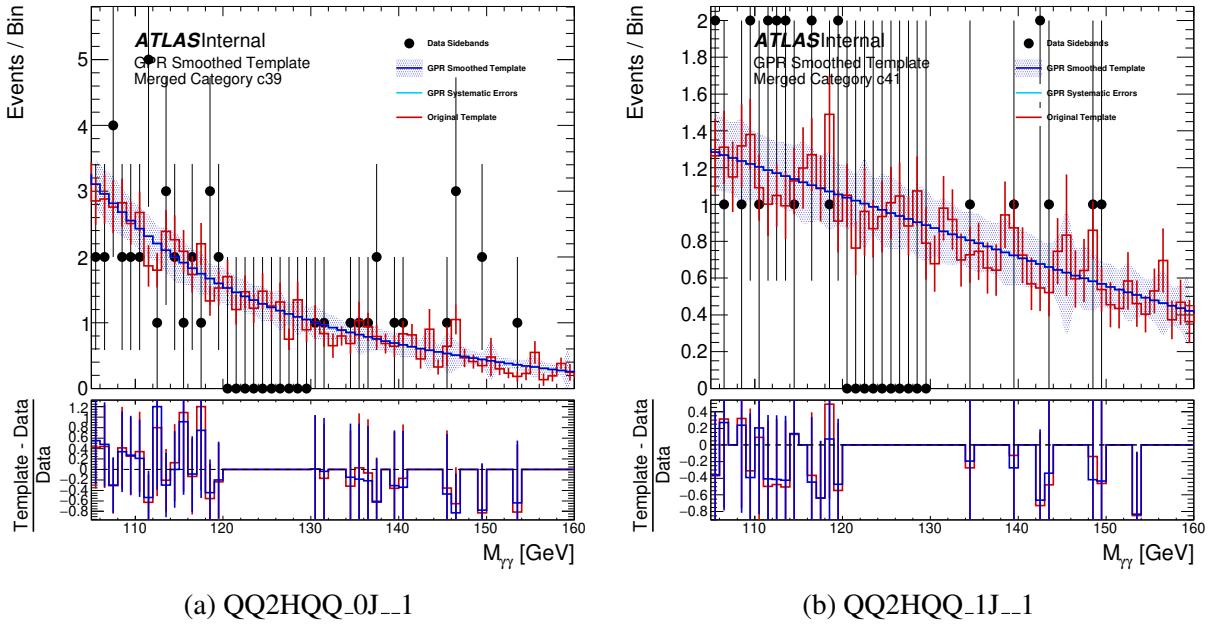
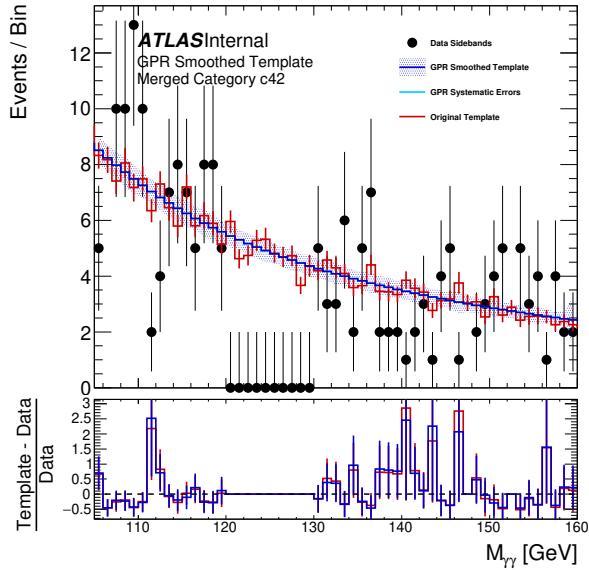


Figure A.38: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

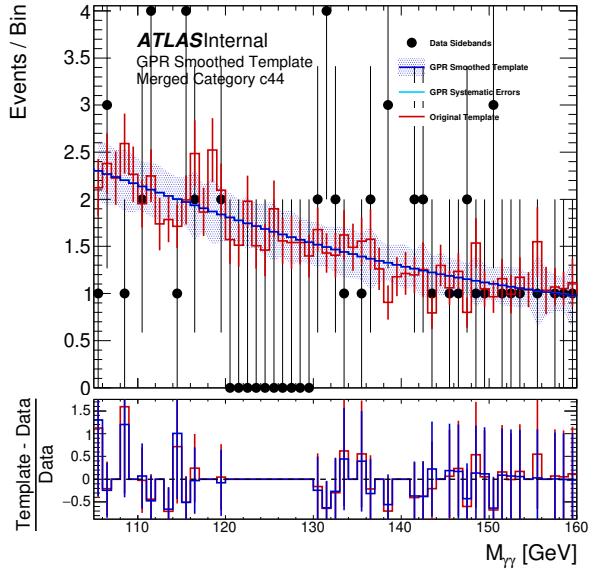
weight of the events in templates in a typical ggH category). We produce three templates in this manner, each dropping off at different rates. Templates are binned at 1 bin/0.5 GeV. We then fit a functional form to the templates, and use the functional form as a basis for drawing toys. Each toy is also manually weighted to have an average event weight of 0.04 so that the performance of the test correctly scales with toy statistics. Because statistical regimes for GPR are measured in terms of events per bin rather than total number of events, the results from this study can be generalized to the Couplings analysis.

The low-weight Sherpa template is fitted with an ExpPoly2, while the med- and high-weight templates are fitted with ExpPoly3 functions. These are the "generating" functional forms noted in subsequent tables in this section; the "fit" functional forms are the closely-related forms used to extract a spurious signal.

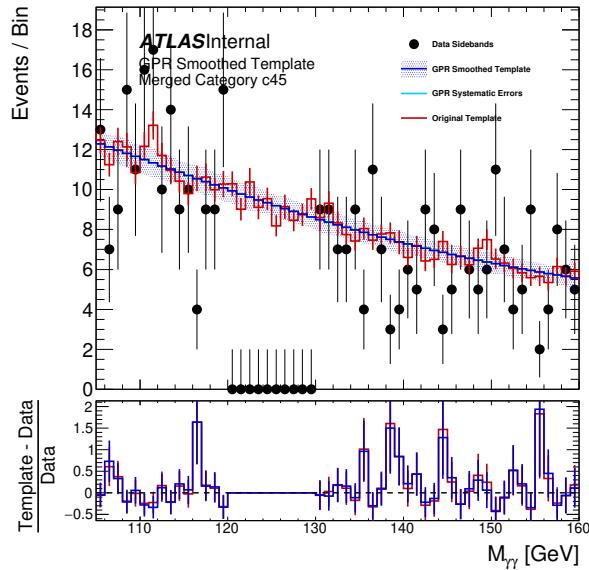
In order to visualize the bias, we record the spurious signal extracted from each toy. In the following reference distributions, the red distribution is the spurious signal distribution of the unsmoothed toy, while the blue distribution is the spurious signal distribution of the smoothed toy. To aid in visualization, we fit a Gaussian to the distribution from plotting many such points; however, numbers recorded for these tests are the sample mean and standard variation of the distribution, not those extracted from this Gaussian fit.



(a) QQ2HQQ_1J_2



(b) QQ2HQQ_GE2J_MJJ_0_60_1



(c) QQ2HQQ_GE2J_MJJ_0_60_2

Figure A.39: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

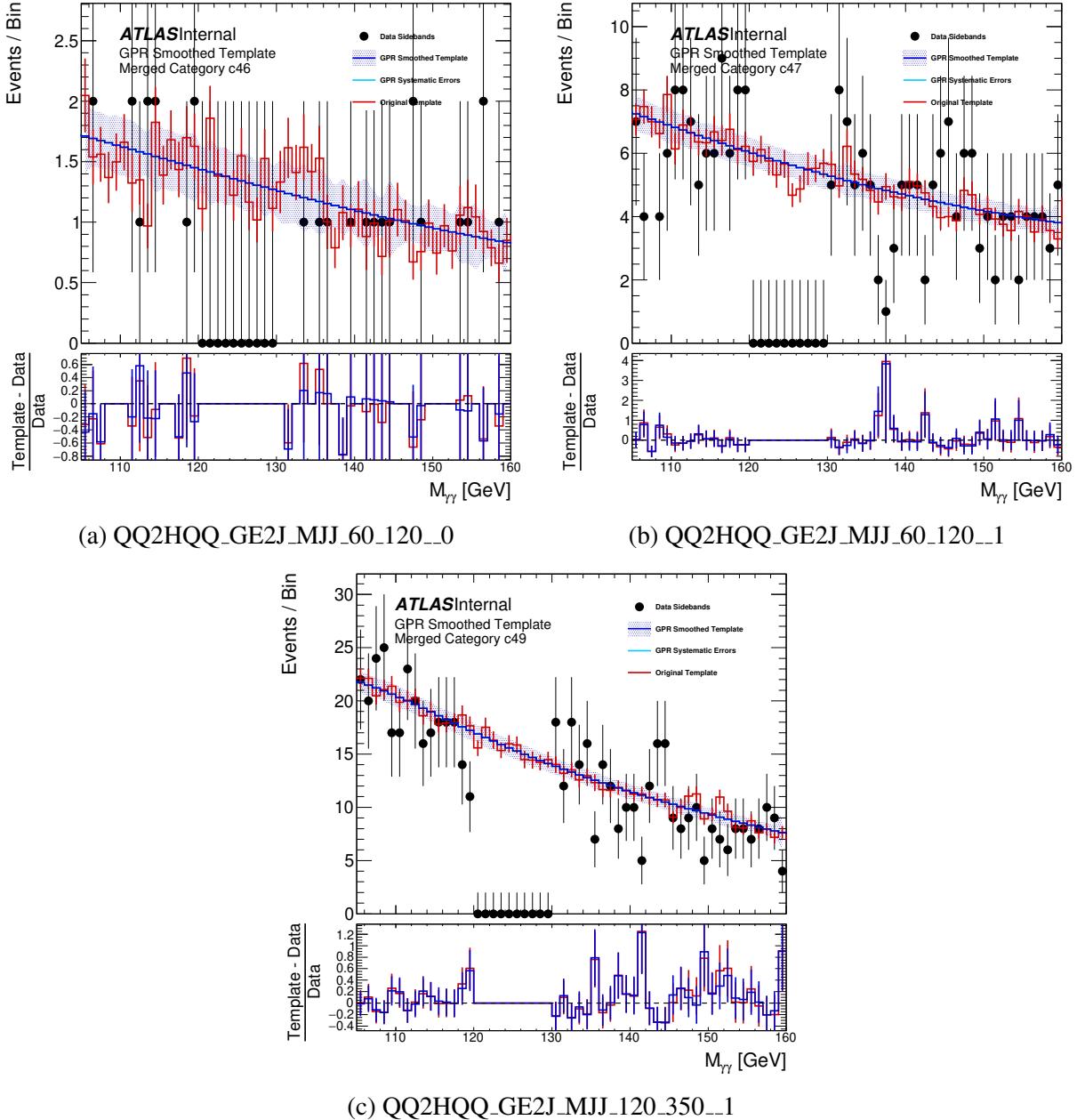


Figure A.40: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

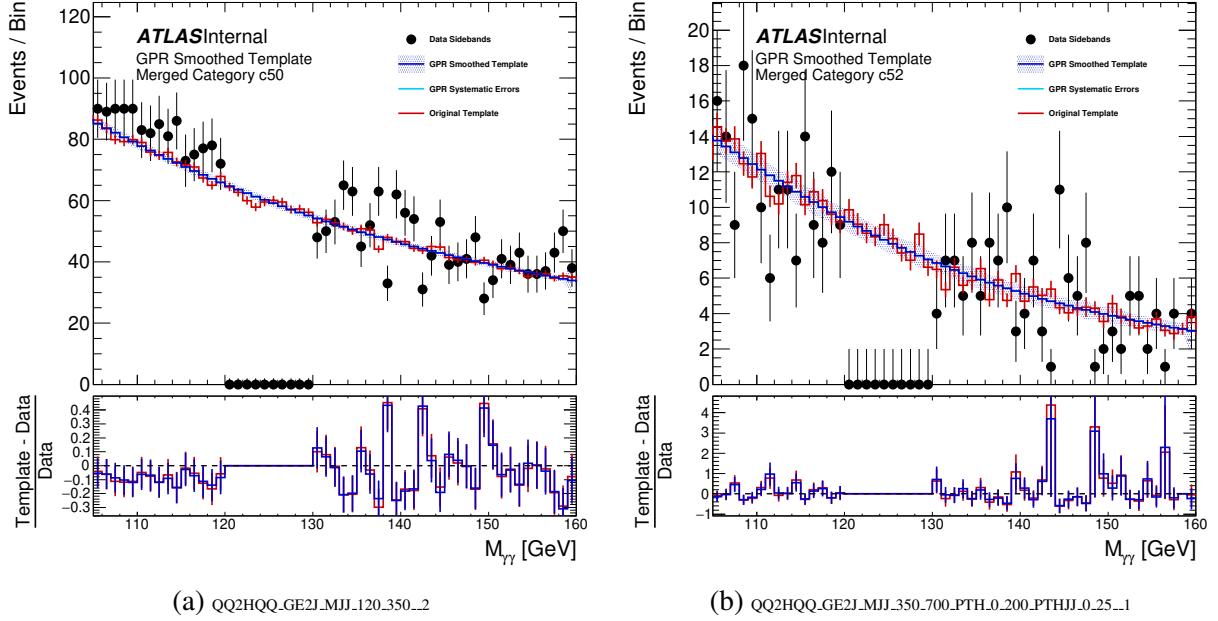


Figure A.41: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

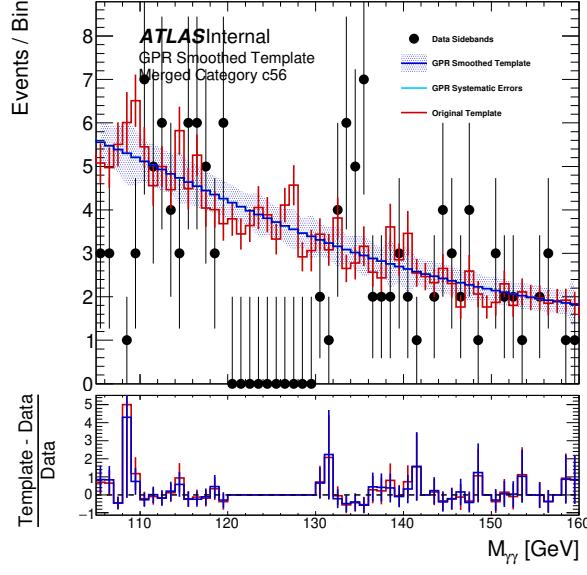
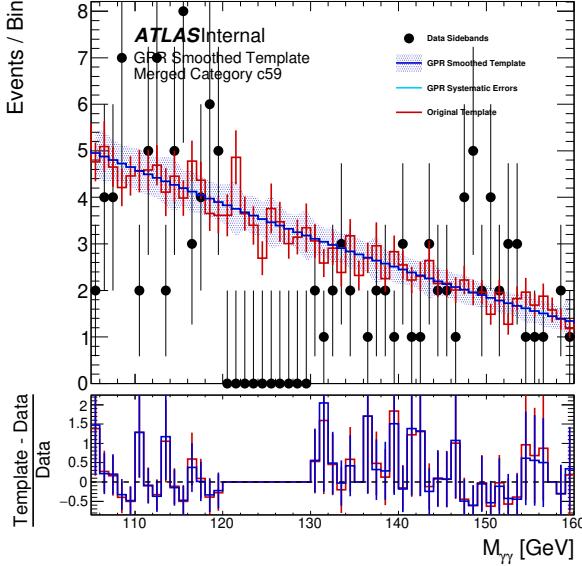


Figure A.42: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.



(a) QQ2HQQ_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_2

Figure A.43: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

A.5.1 Nominal Bias Study

In all studies, we observe that the bias does not change appreciably with template shape (that is, for a given statistics level, the bias is approximately the same for all three templates, regardless of generating functional form). However, the bias does depend on statistics: for all three shapes, the bias decreased as a function of the reference signal as the template statistics increased. We show the spurious signal extracted for 10 events through 10,000,000 events for all three shapes, and record the numeric results for all studies in a table. In this iteration of the study, we use the initial error on the templates as the noise estimate in the GPR fit step.

At this stage, the bias is shown to be consistently less than 40% of the statistical uncertainty on the unsmoothed spurious signal.

We report the results of the nominal bias study for all categories in the Table ???. Additionally, we plot the per-toy bias (that is, GP_SS - Raw_SS for each toy) for each shape and stat regime.

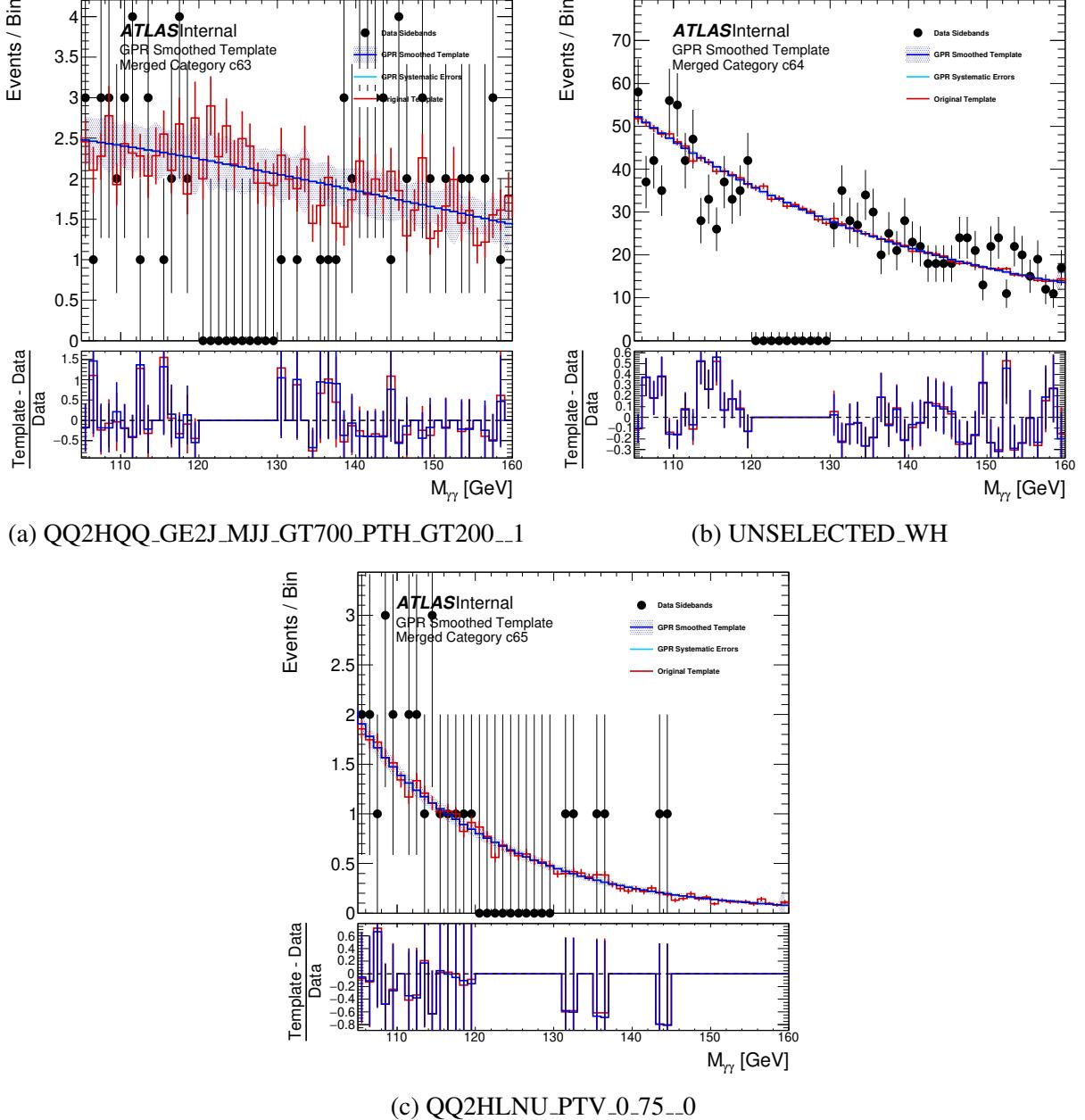


Figure A.44: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

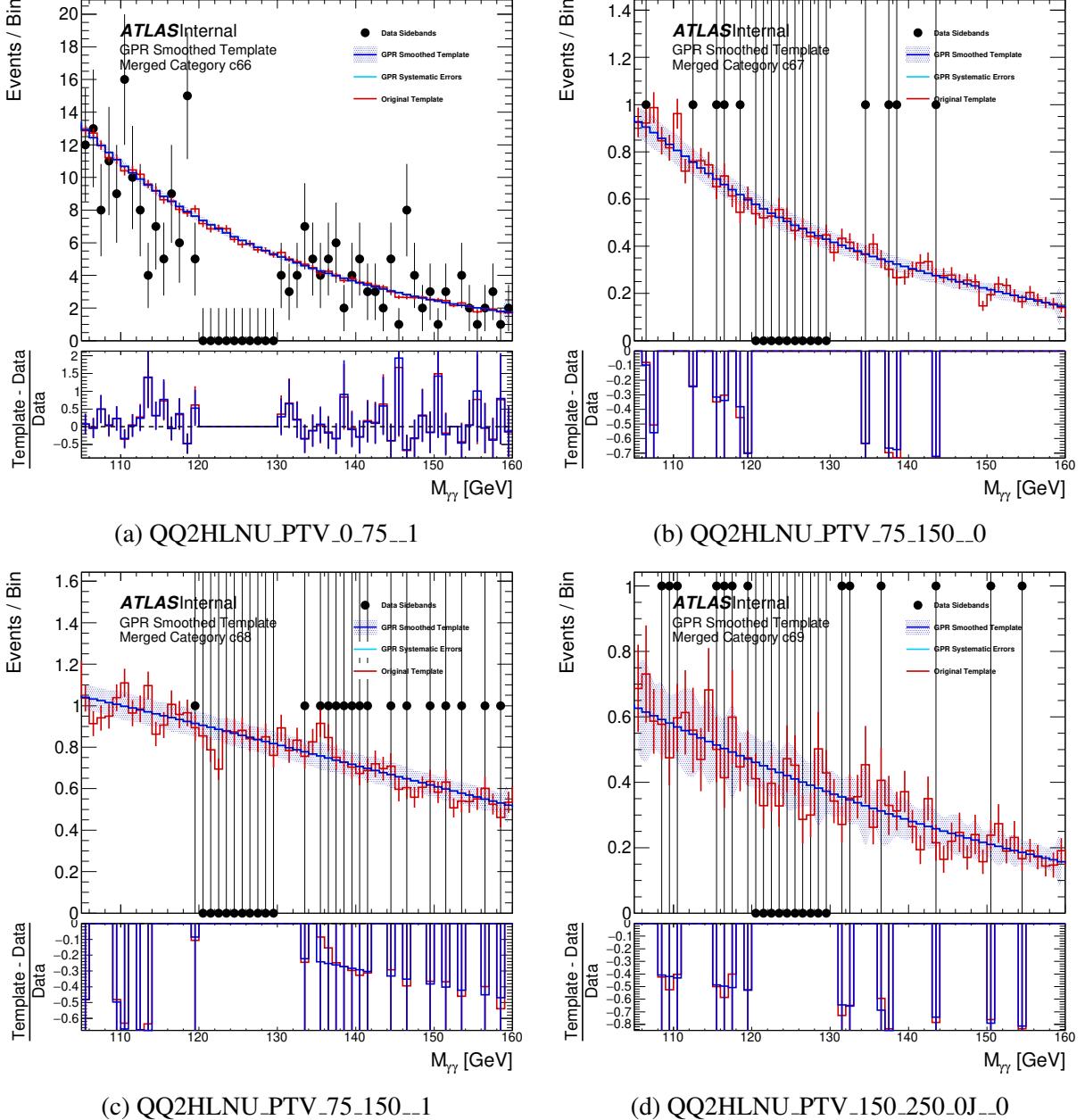


Figure A.45: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

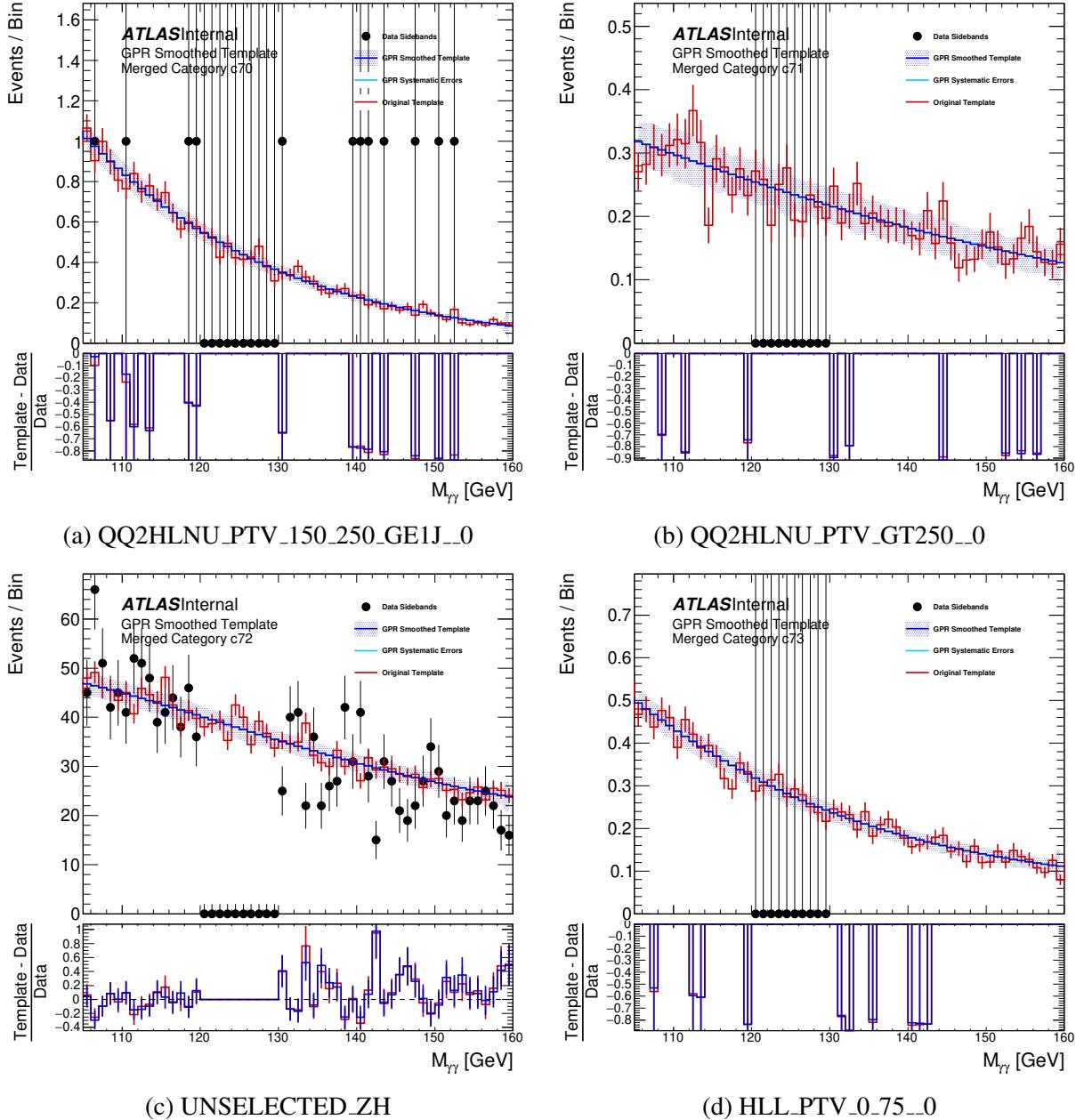


Figure A.46: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

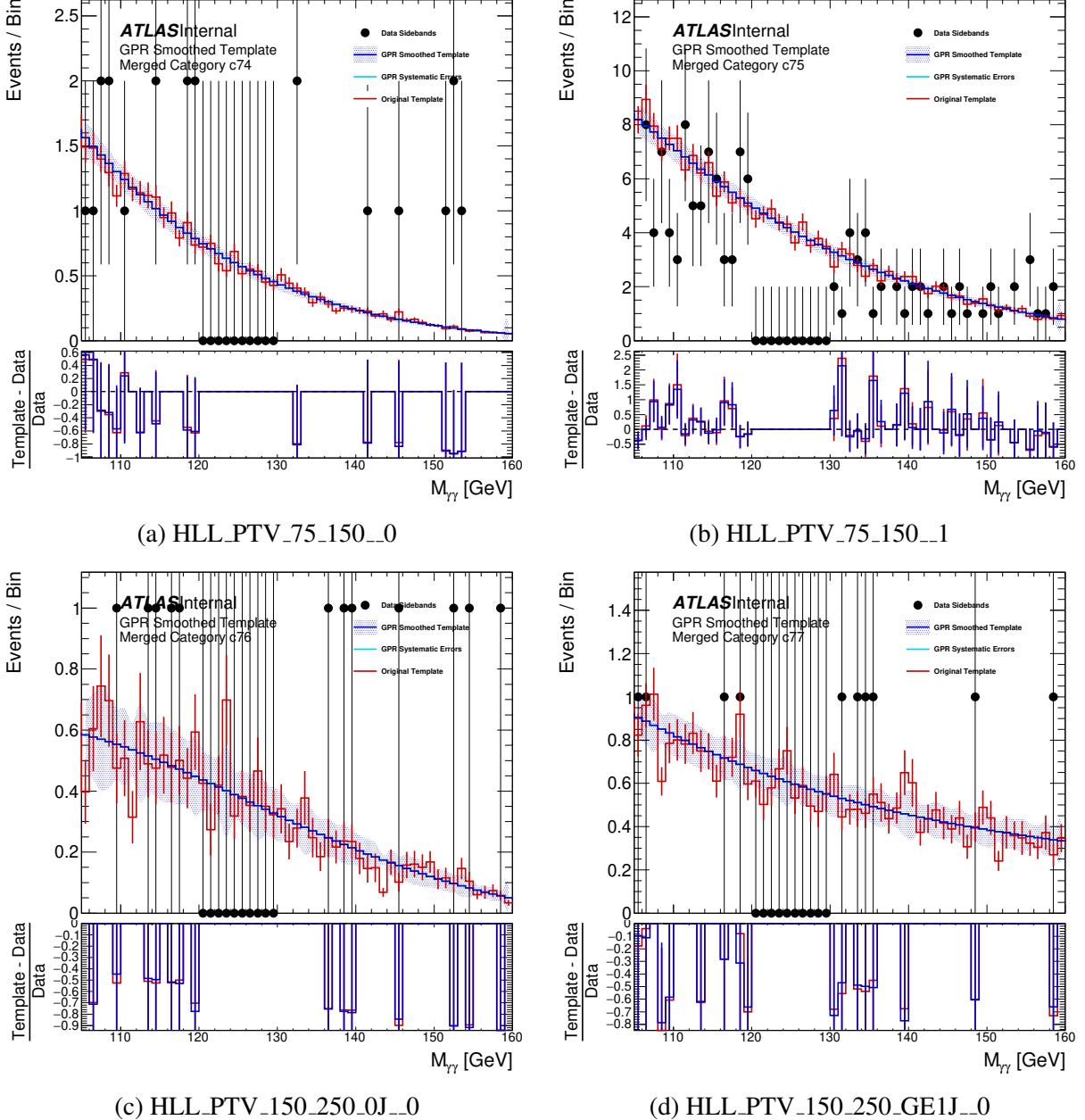


Figure A.47: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

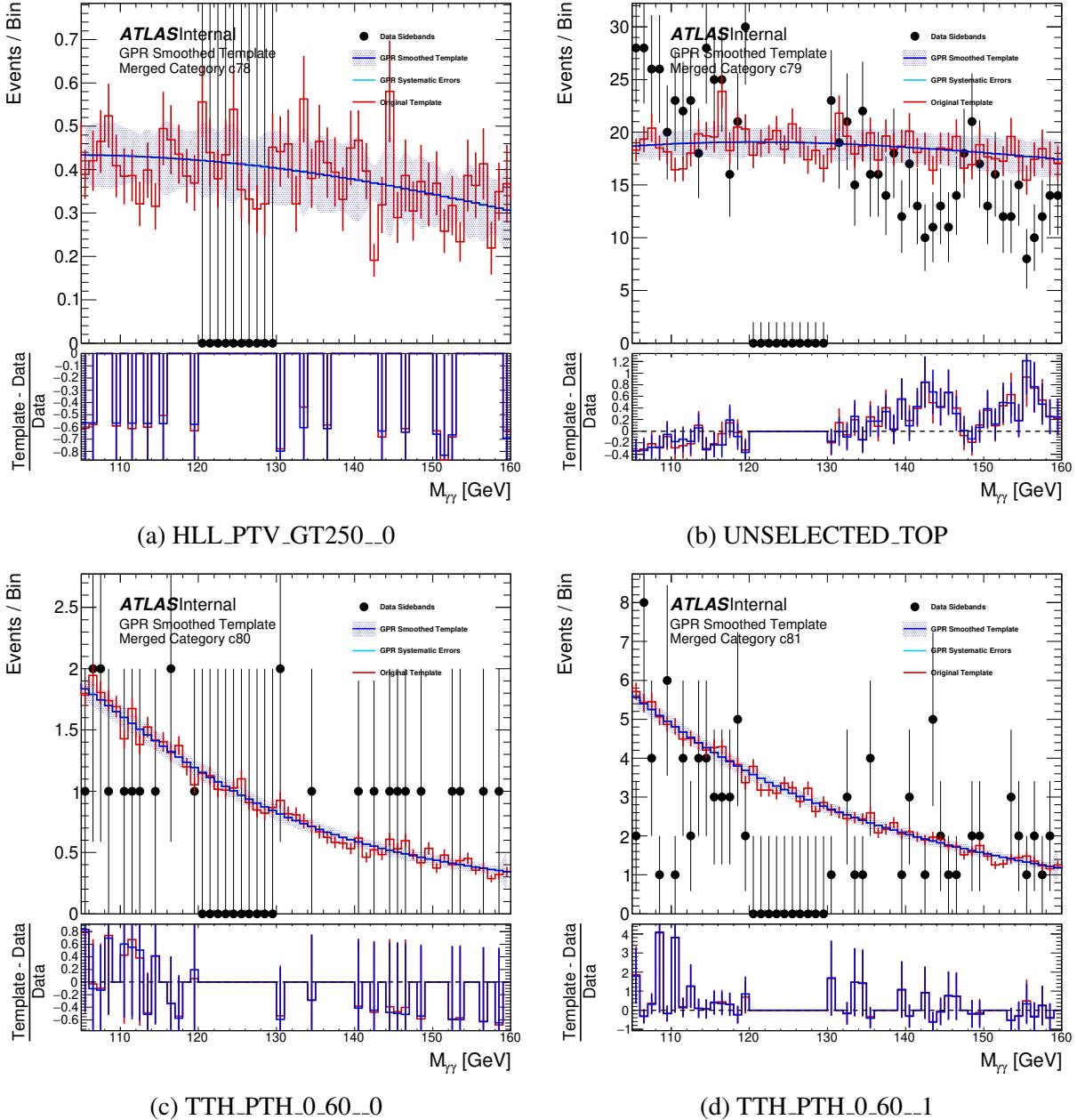


Figure A.48: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

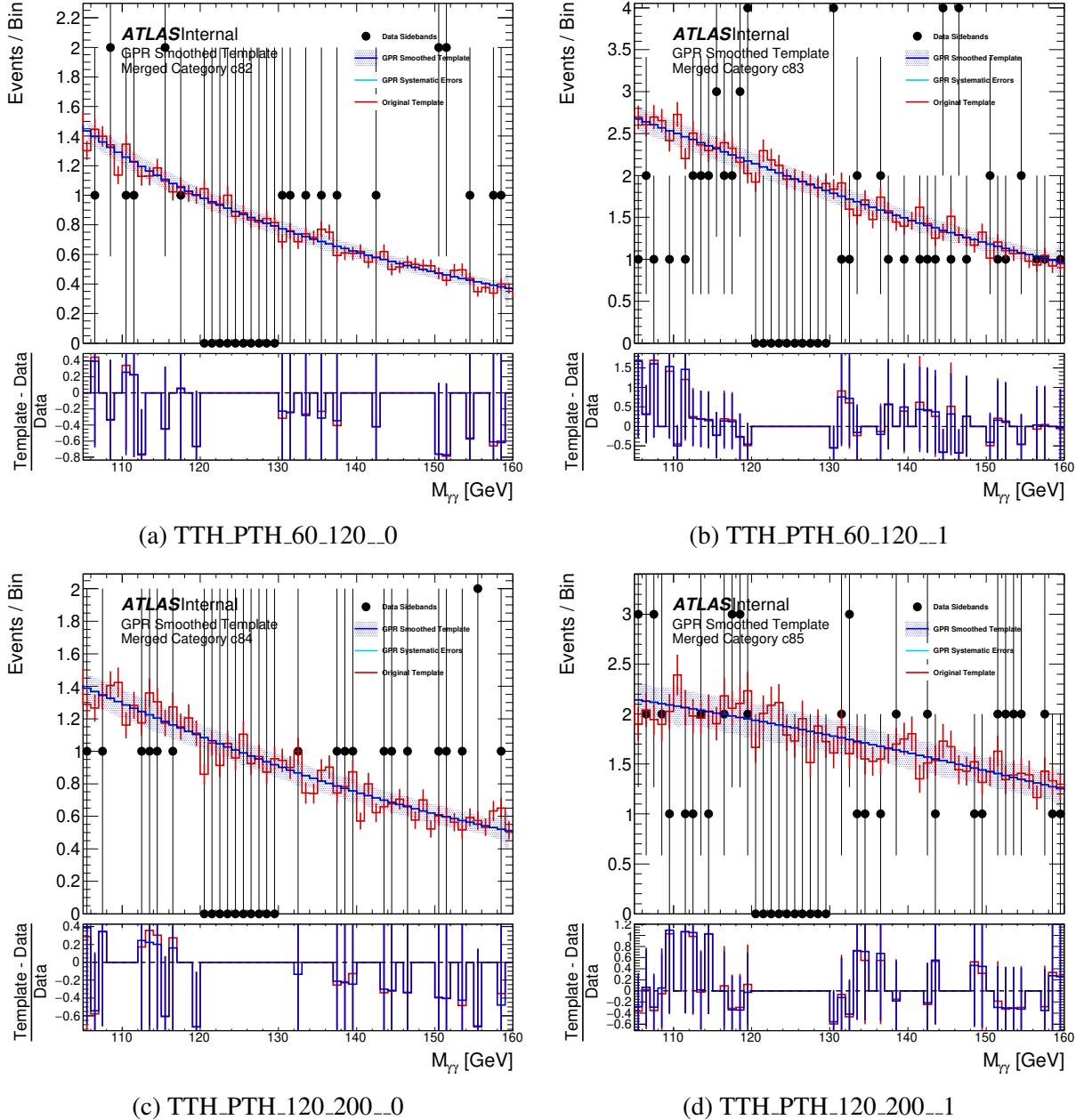


Figure A.49: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

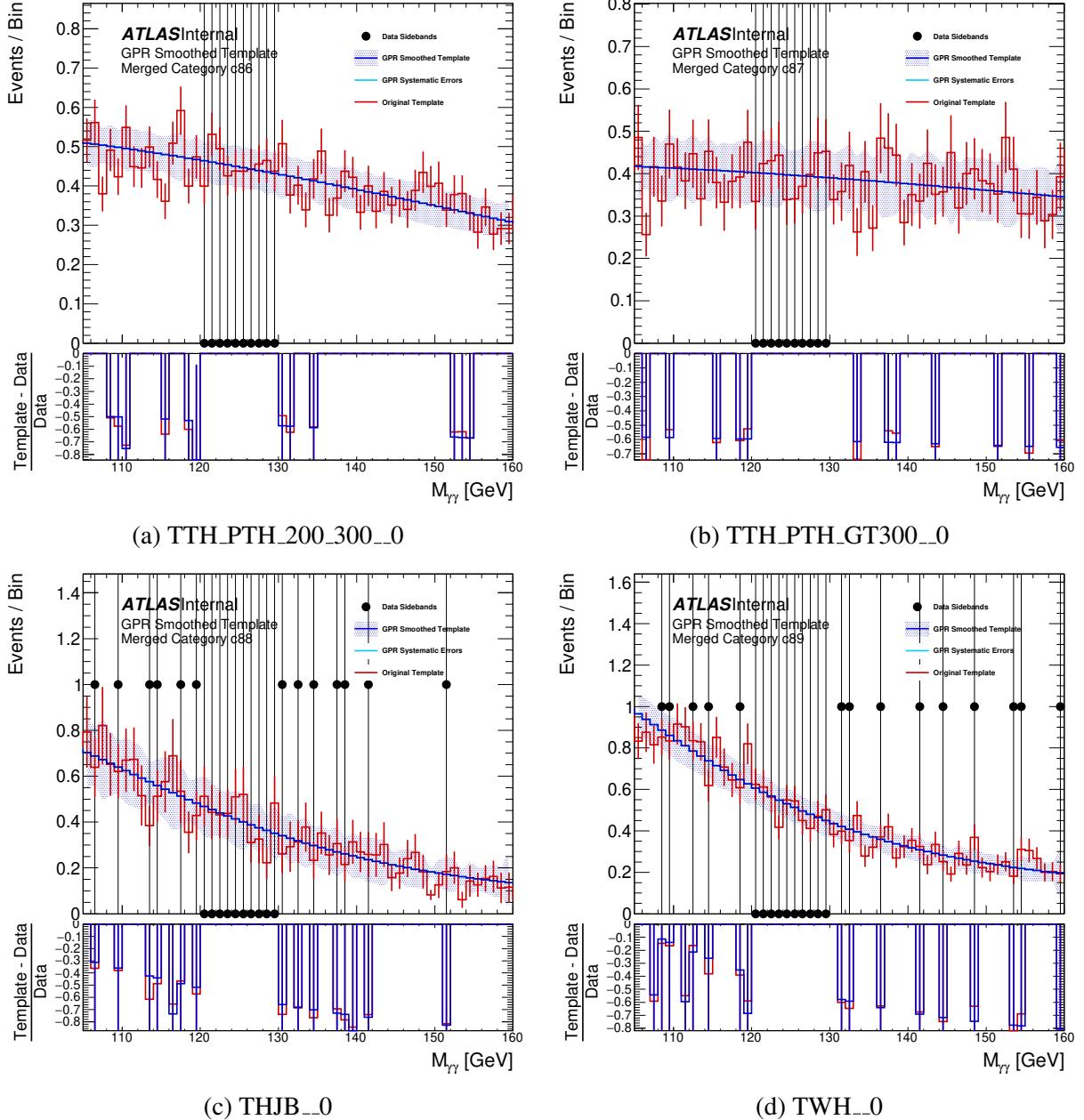


Figure A.50: The full Run 2 background templates of the labelled categories. The red shape shows the original background template, the blue shape shows the smoothed background template, and the black points show the data sidebands (for reference). The bottom panel shows the fractional difference between the smoothed and un-smoothed templates and the data sidebands.

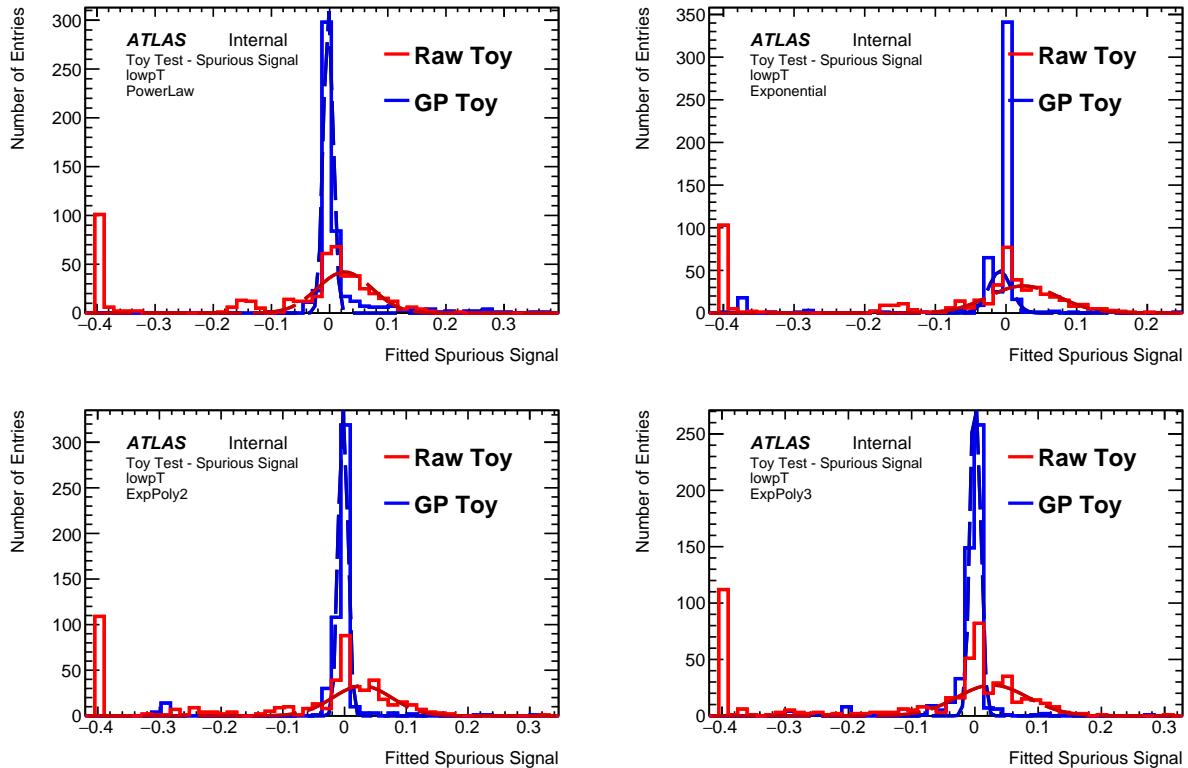


Figure A.51: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template. Each toy in this test has 10 events.

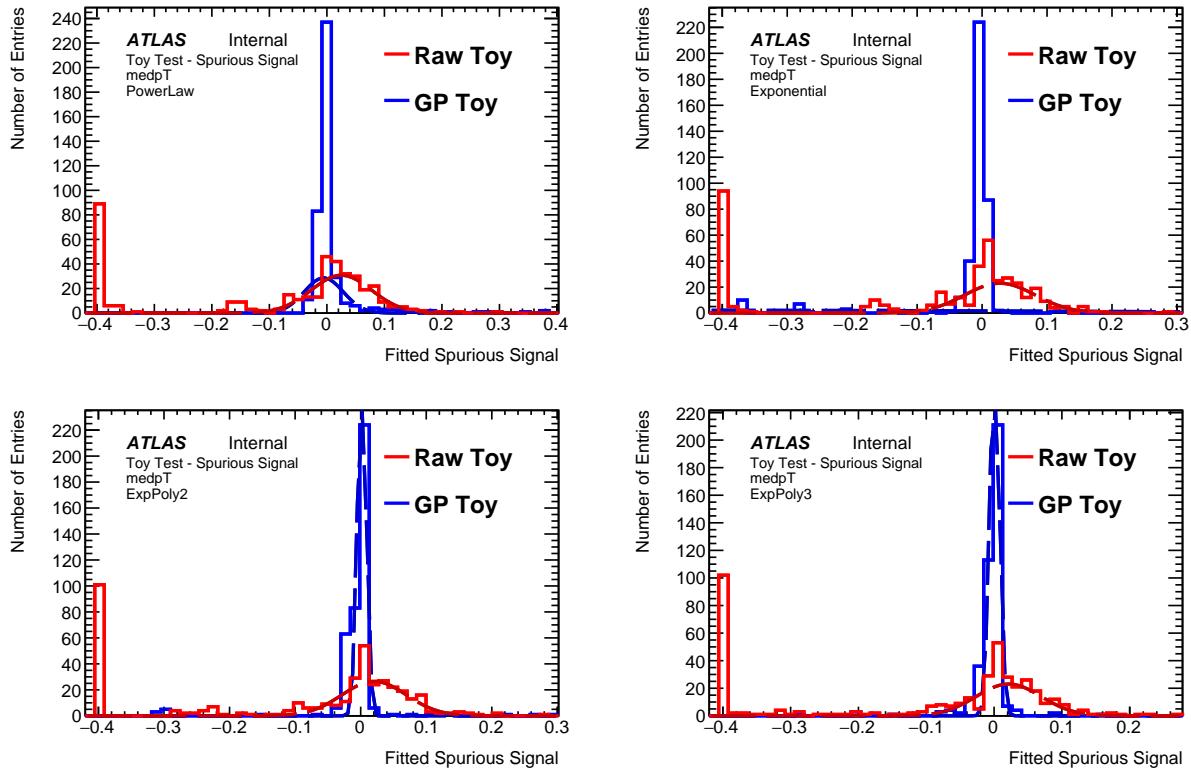


Figure A.52: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template. Each toy in this test has 10 events.

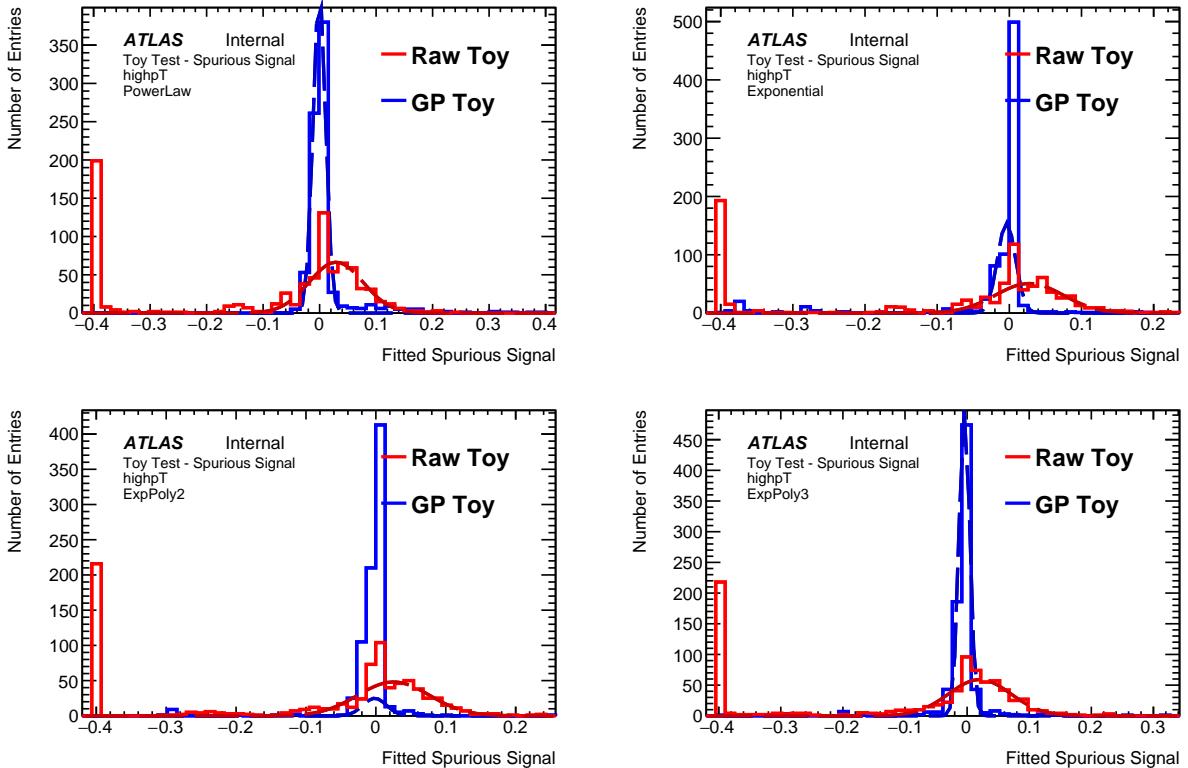


Figure A.53: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template. Each toy in this test has 10 events.

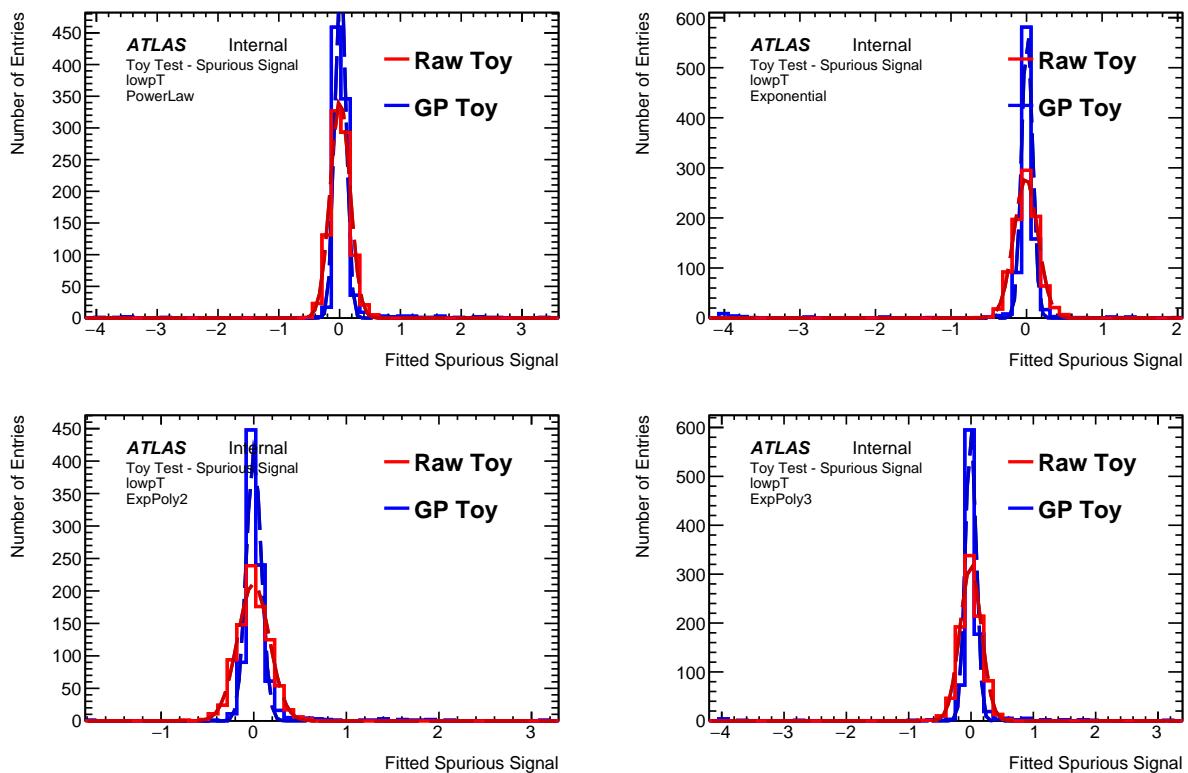


Figure A.54: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template. Each toy in this test has 100 events.

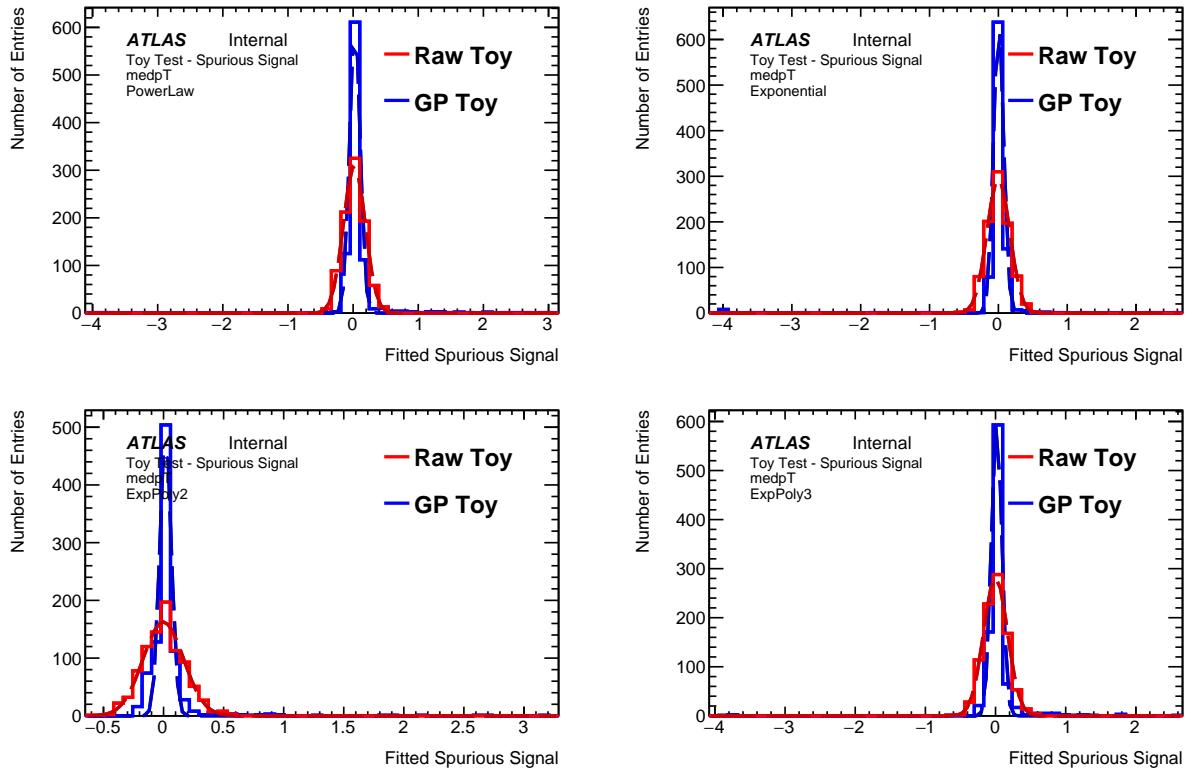


Figure A.55: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template. Each toy in this test has 100 events.

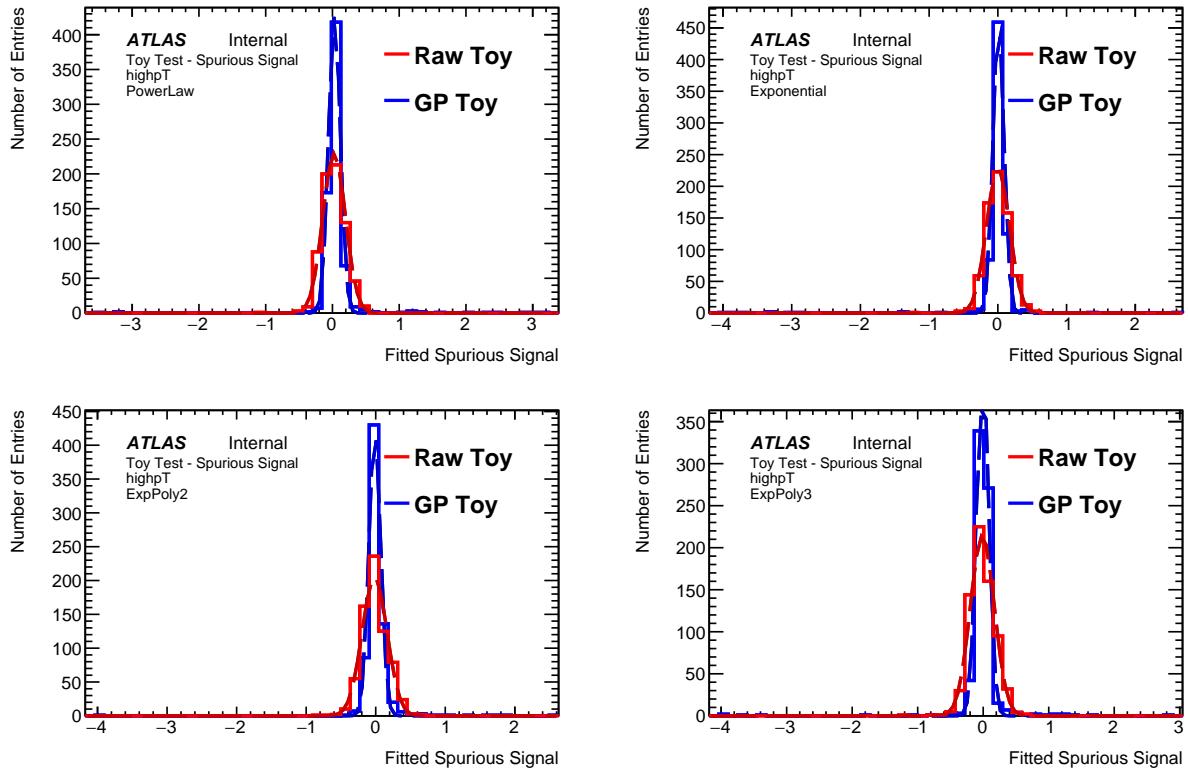


Figure A.56: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template. Each toy in this test has 100 events.

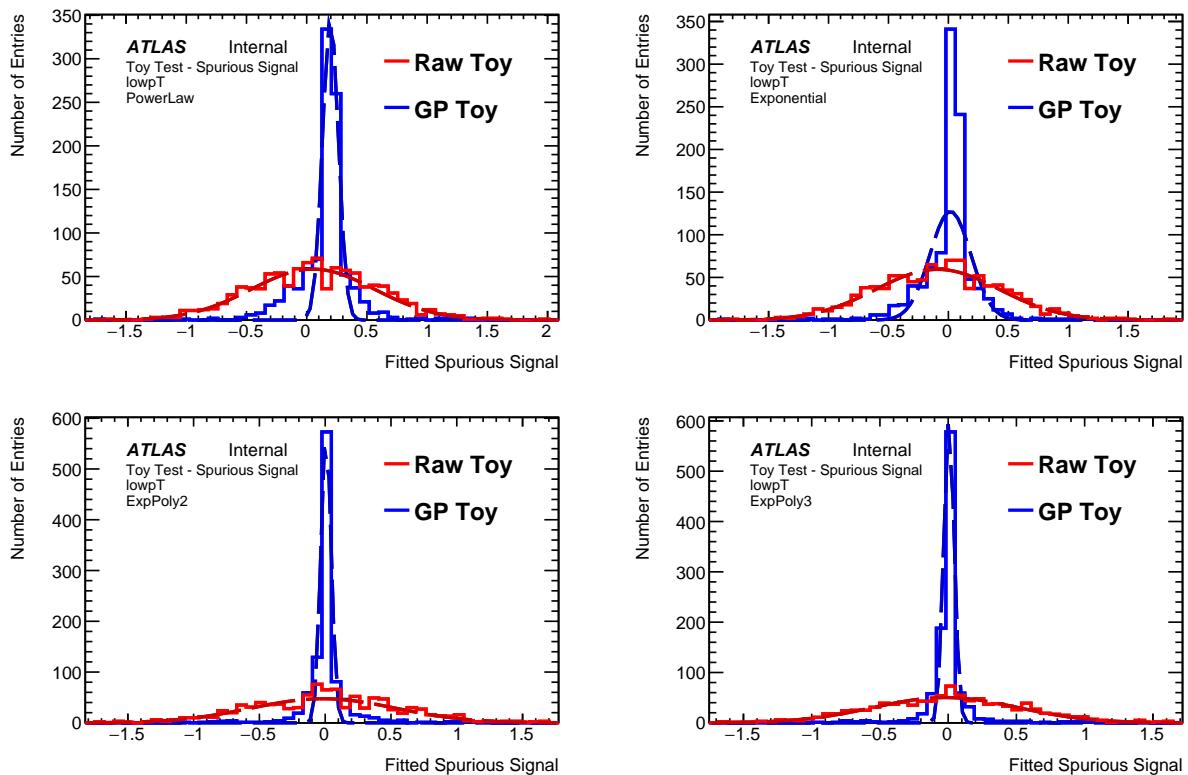


Figure A.57: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template. Each toy in this test has 1000 events.

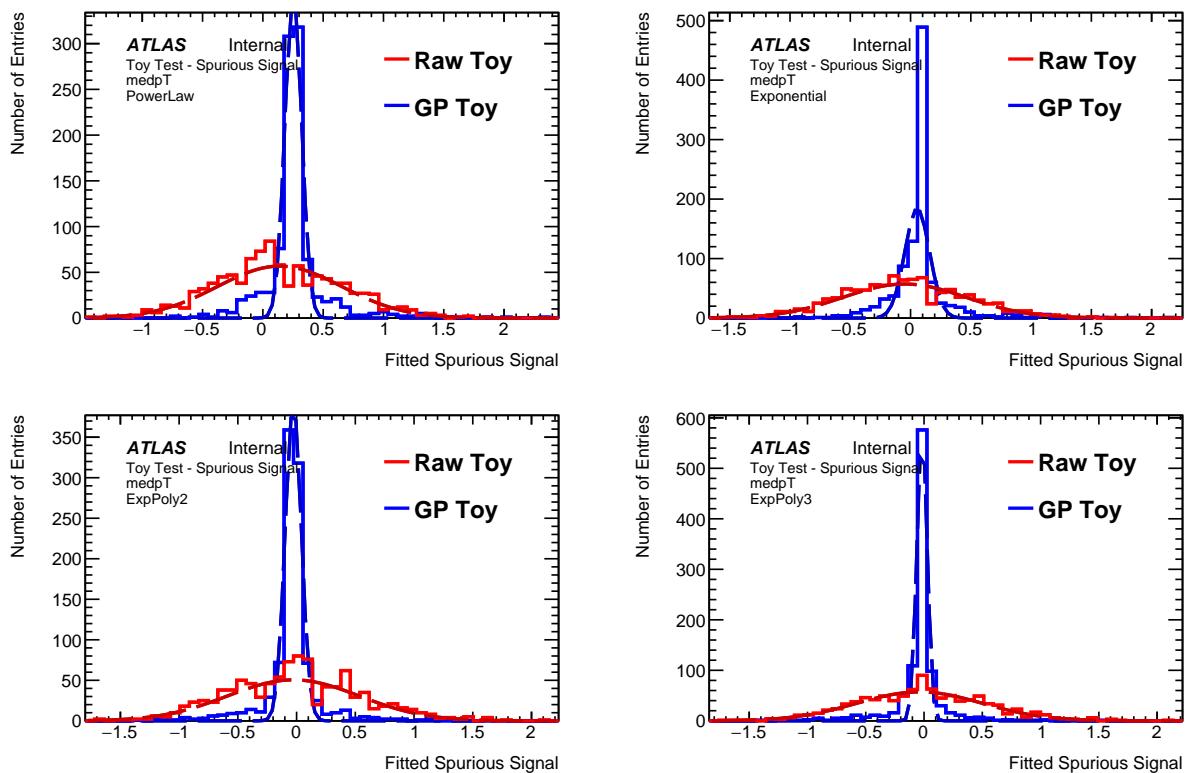


Figure A.58: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template. Each toy in this test has 1000 events.

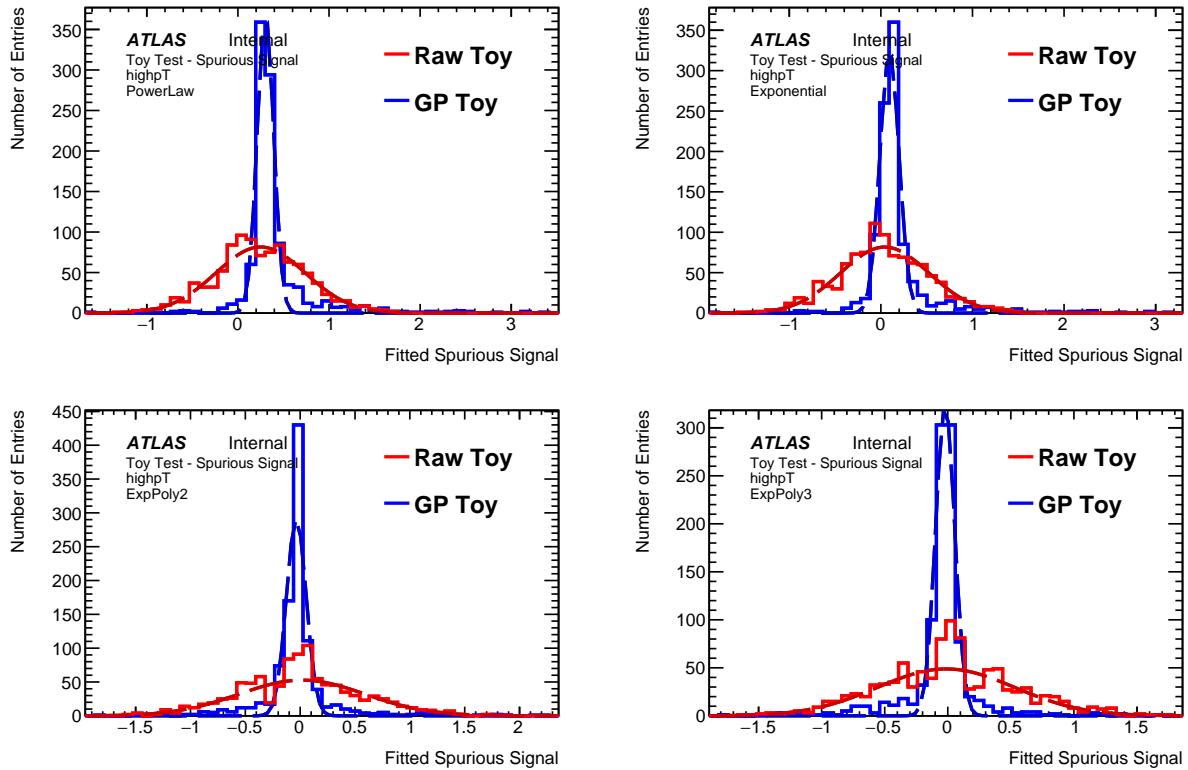


Figure A.59: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template. Each toy in this test has 1000 events.

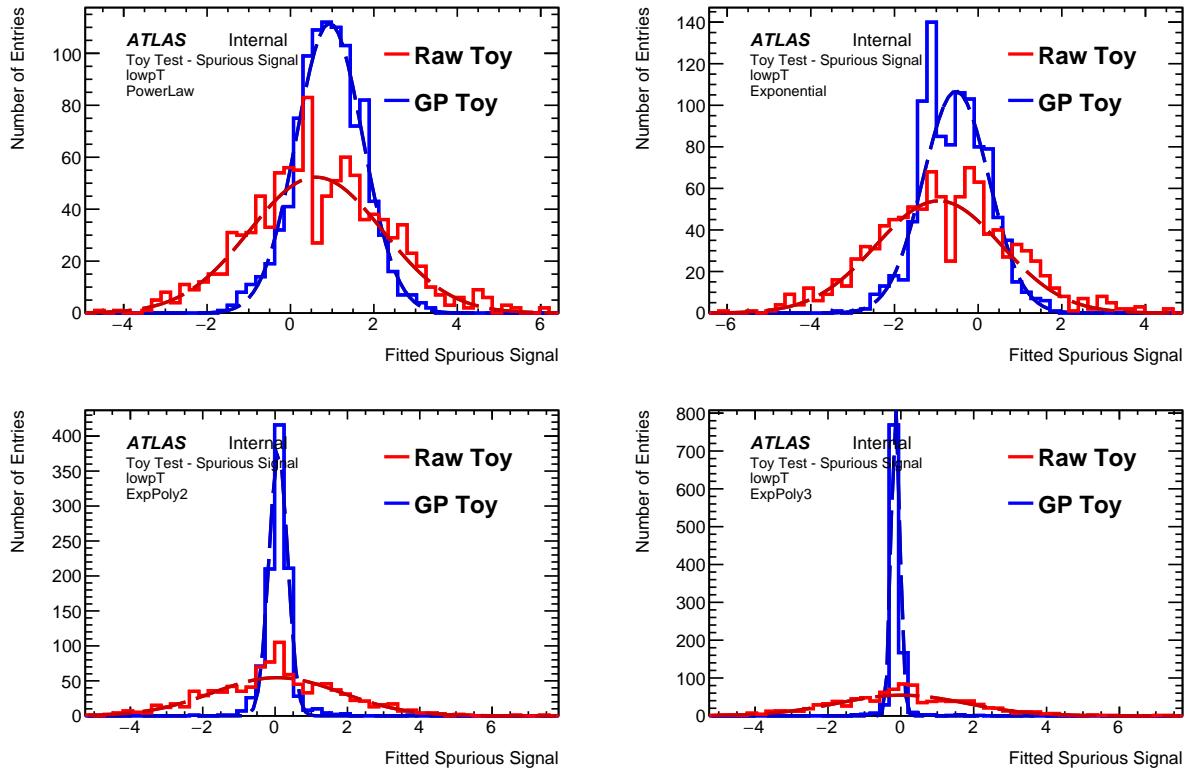


Figure A.60: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template. Each toy in this test has 10k events.

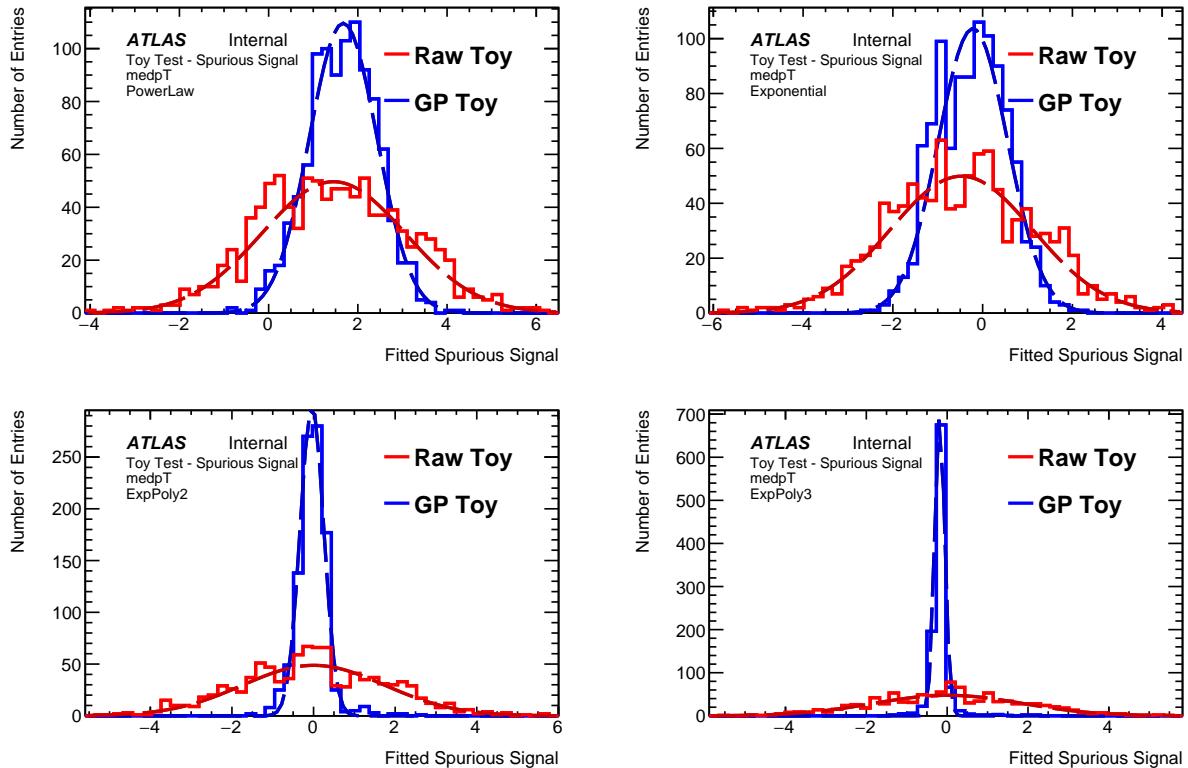


Figure A.61: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template. Each toy in this test has 10k events.

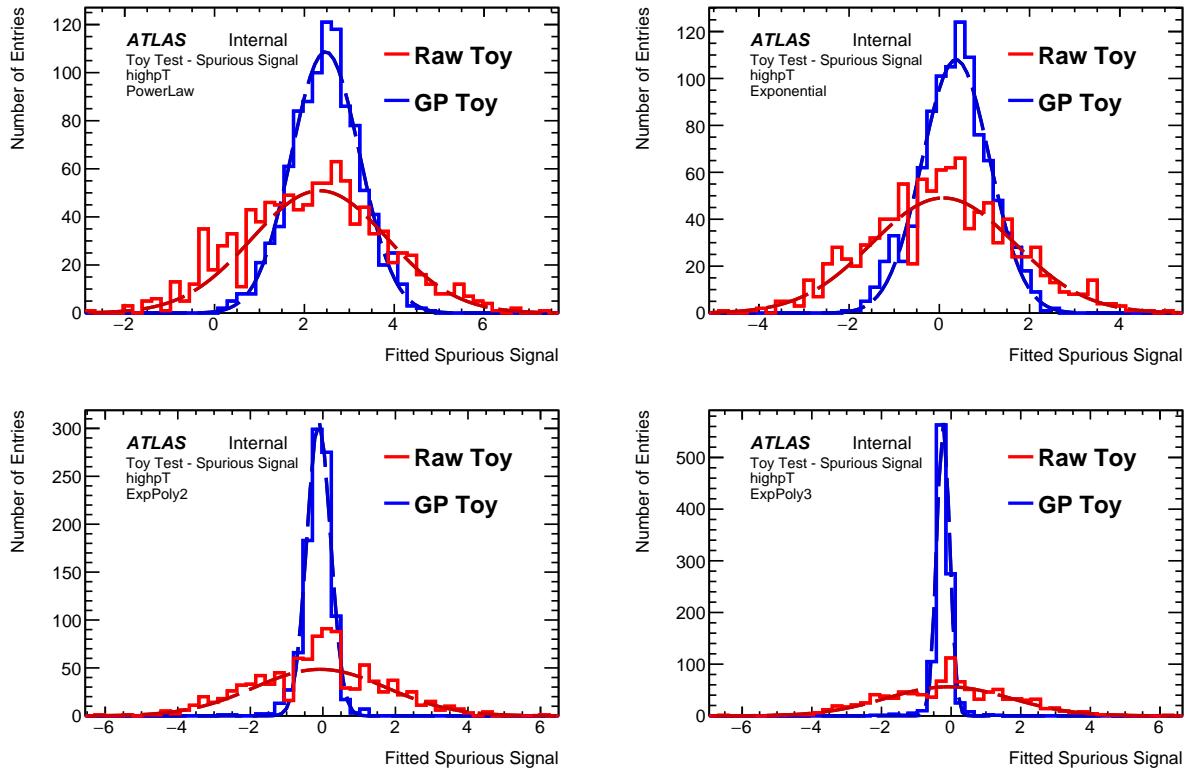


Figure A.62: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template. Each toy in this test has 10k events.

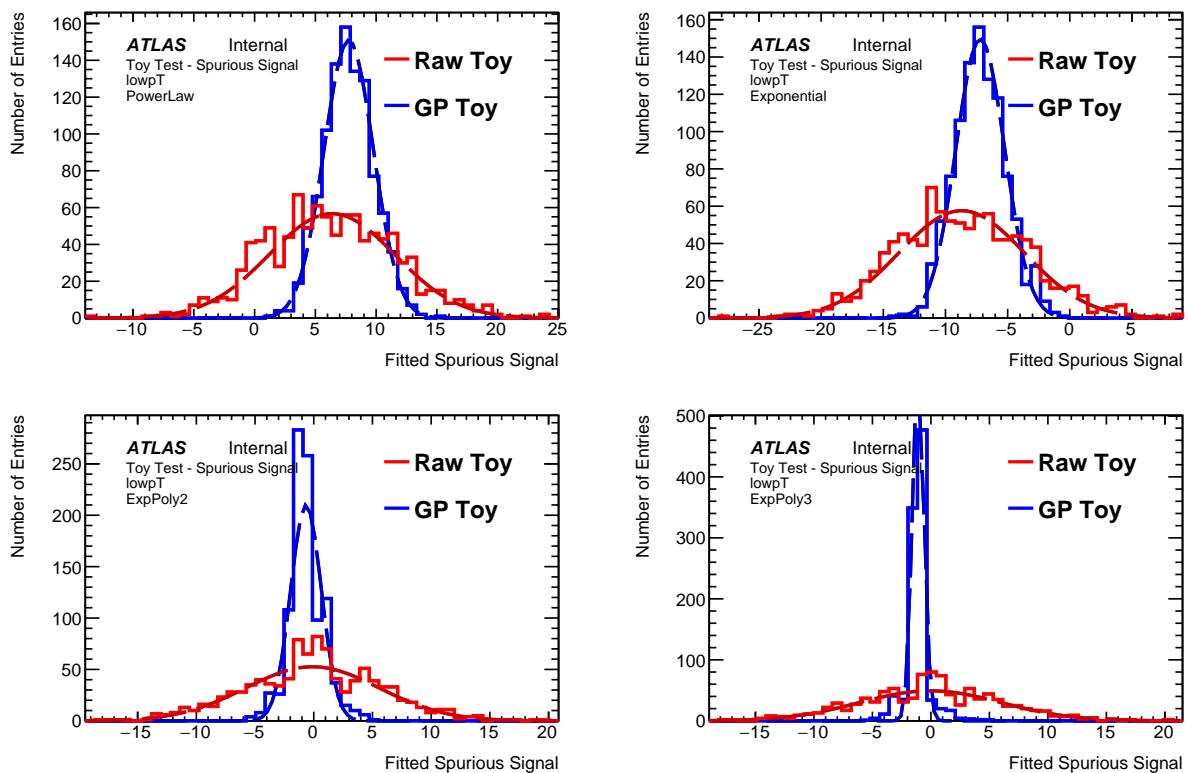


Figure A.63: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template. Each toy in this test has 100k events.

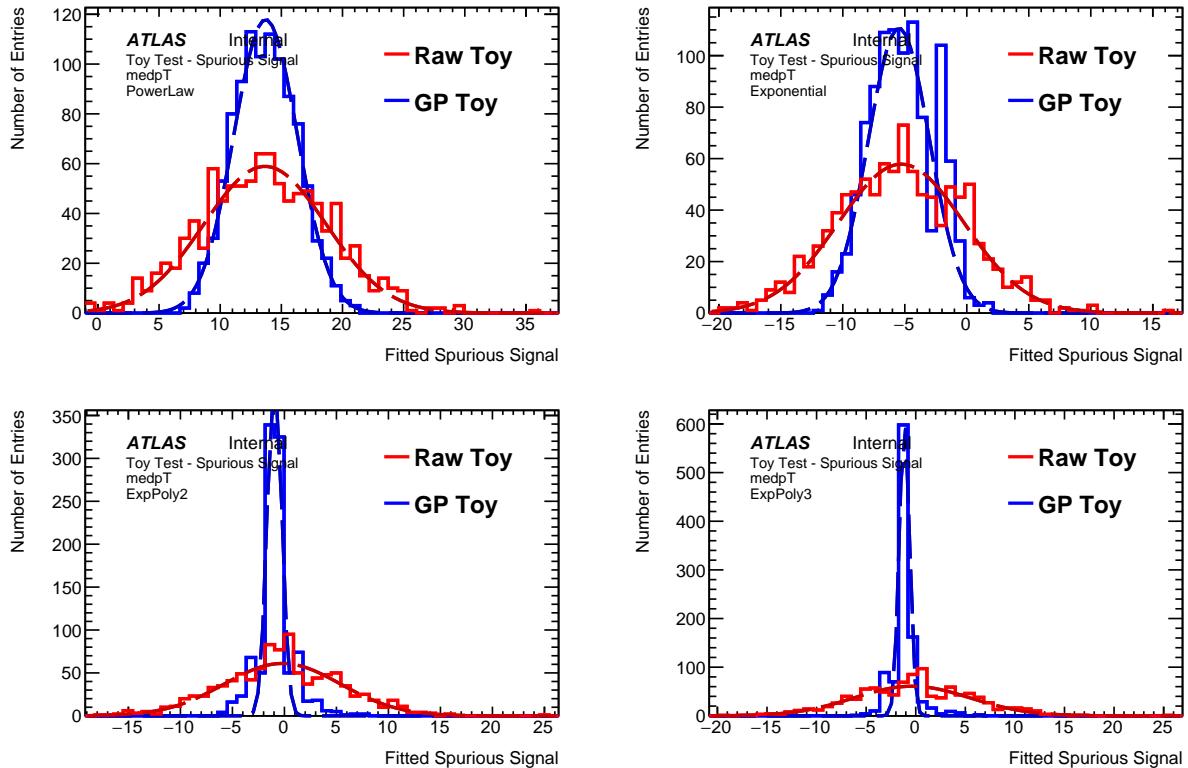


Figure A.64: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template. Each toy in this test has 100k events.

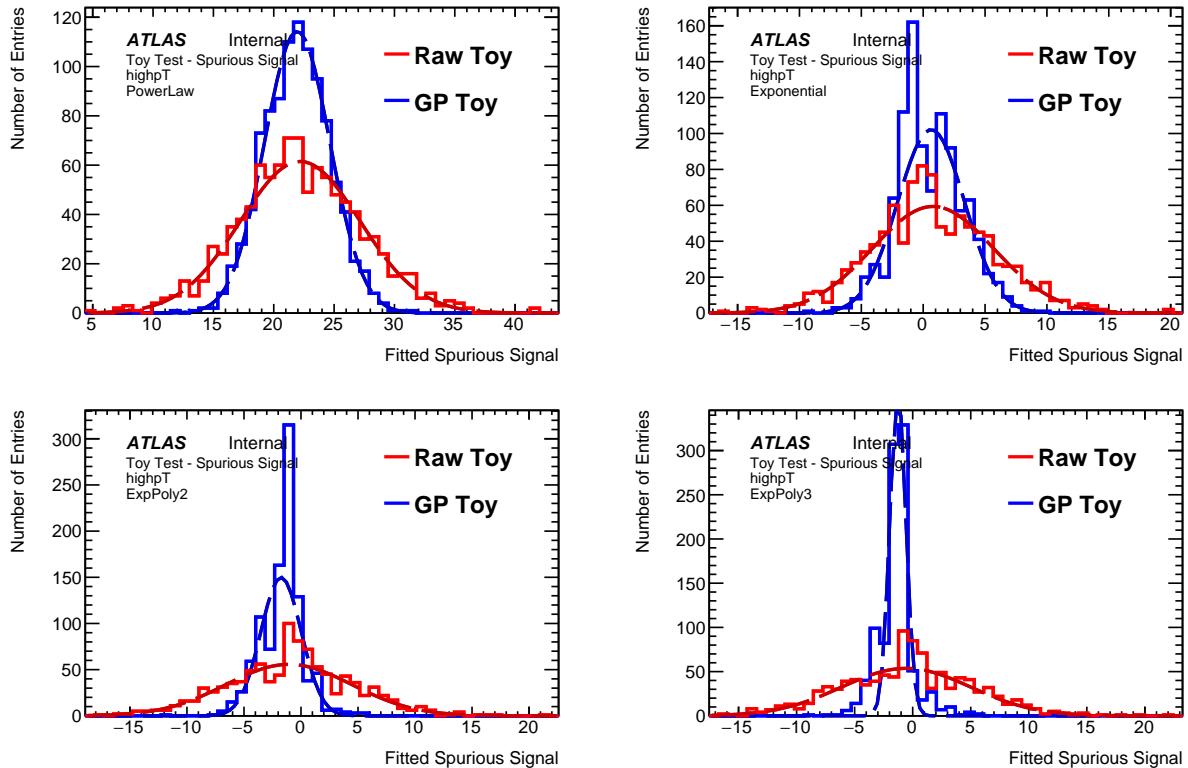


Figure A.65: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template. Each toy in this test has 100k events.

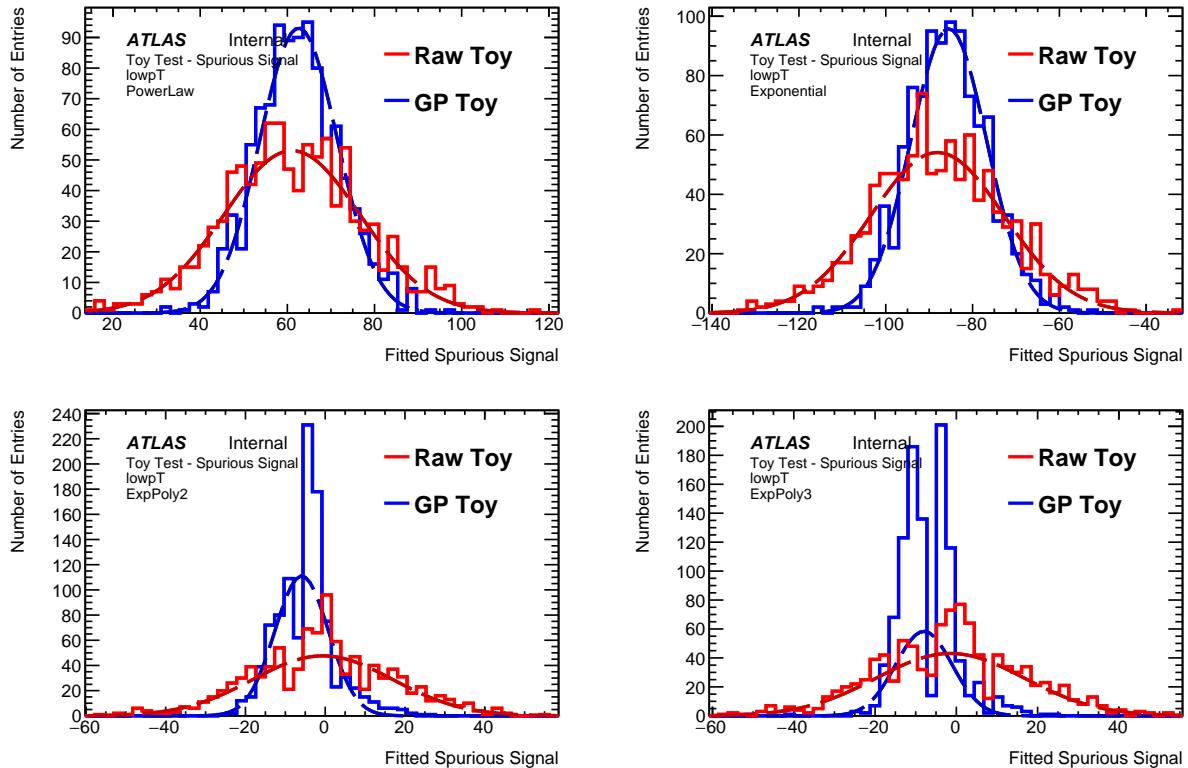


Figure A.66: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template. Each toy in this test has 1,000,000 events.

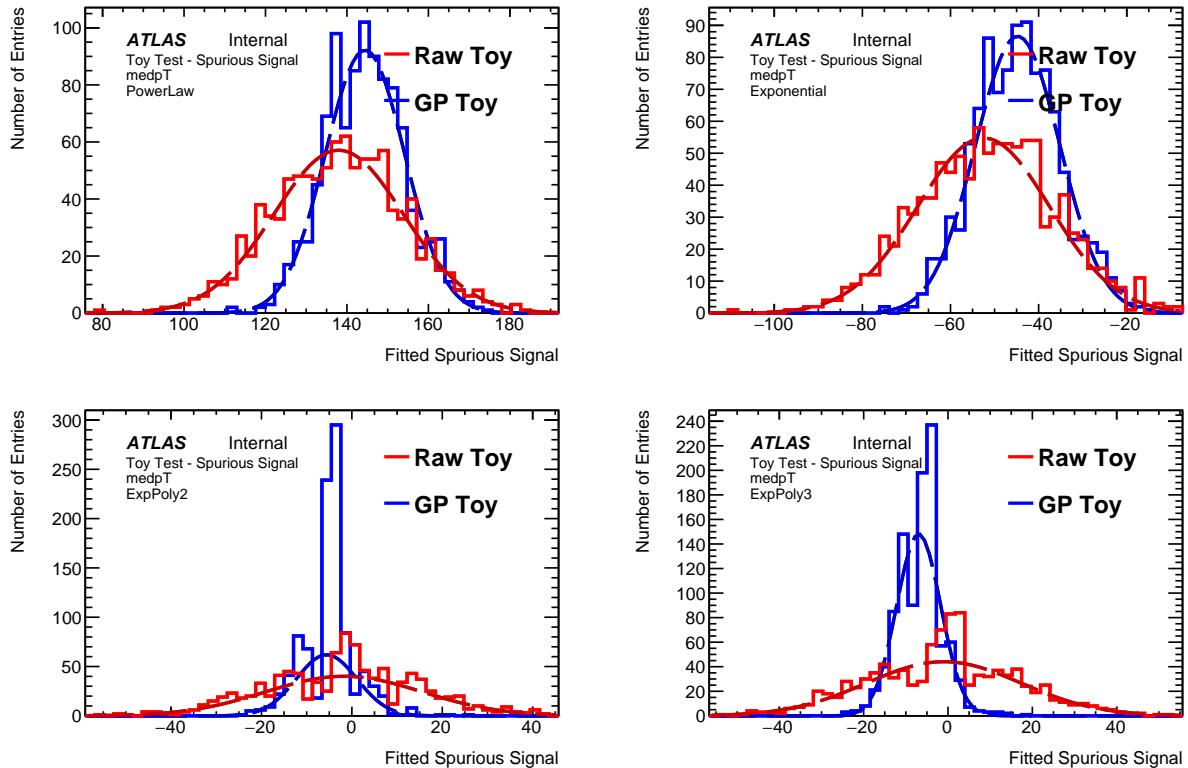


Figure A.67: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template. Each toy in this test has 1,000,000 events.

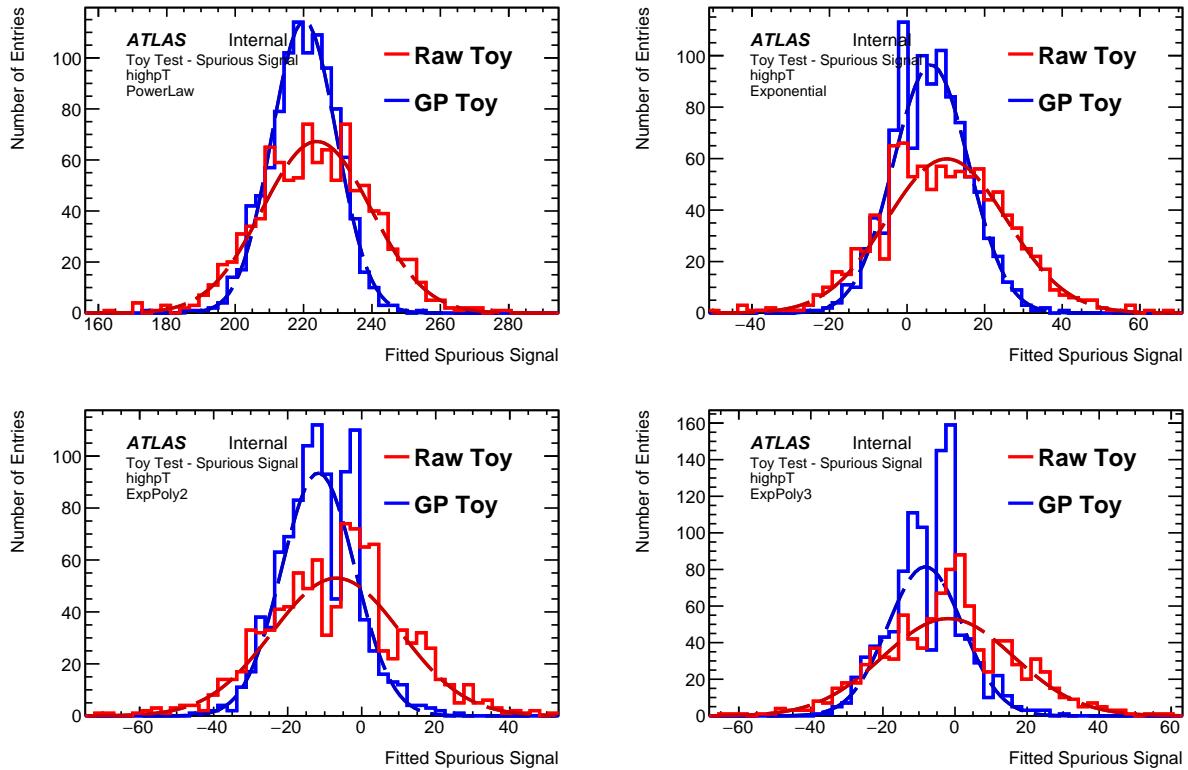


Figure A.68: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template. Each toy in this test has 1,000,000 events.

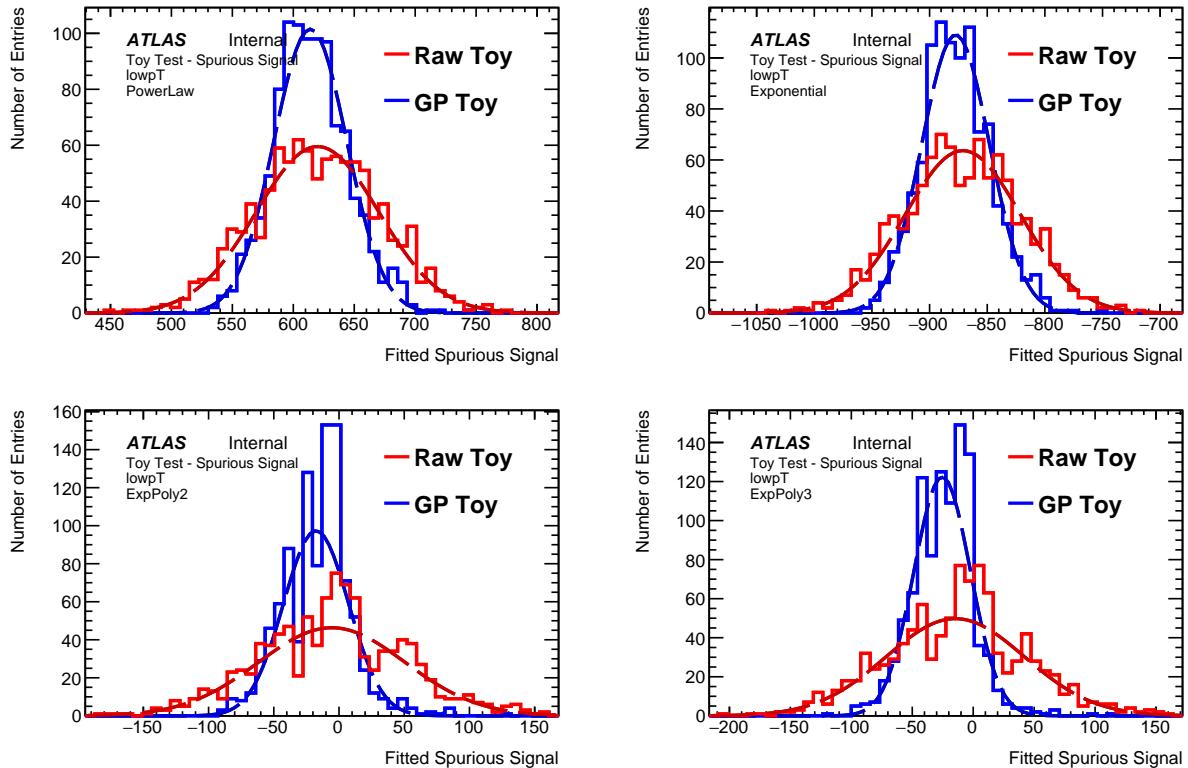


Figure A.69: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template. Each toy in this test has 10,000,000 events.

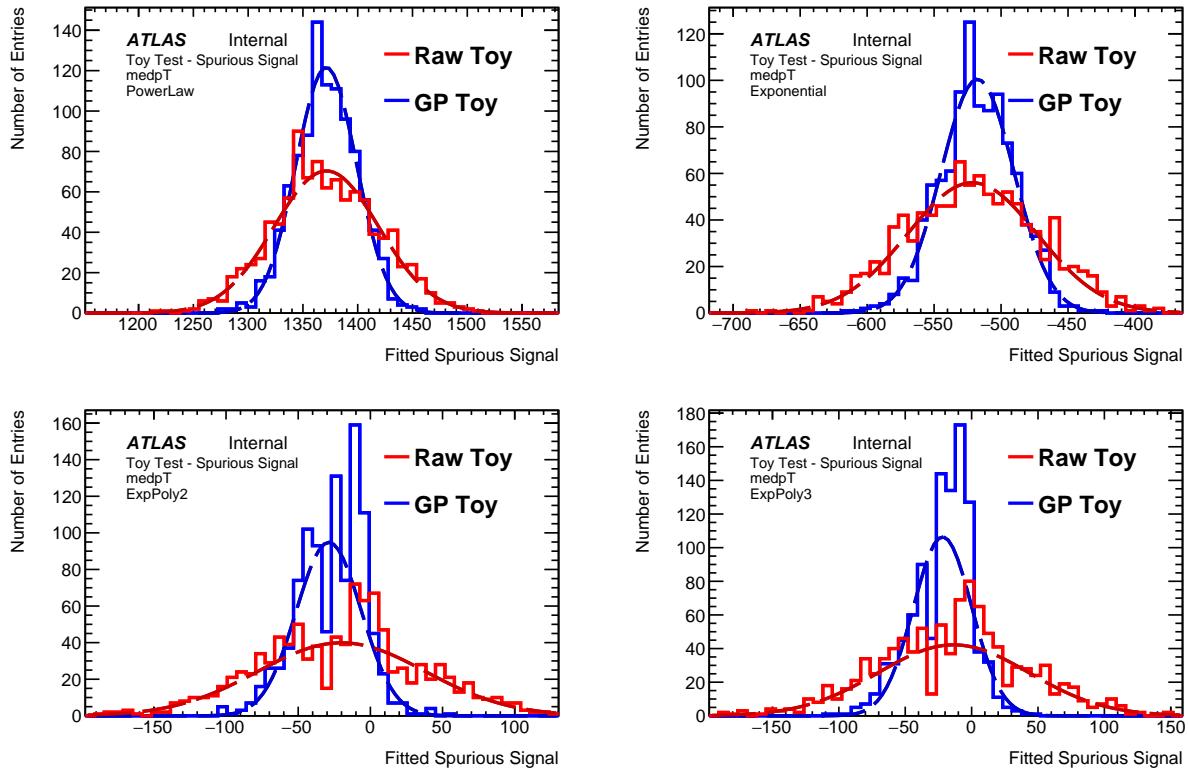


Figure A.70: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template. Each toy in this test has 10,000,000 events.

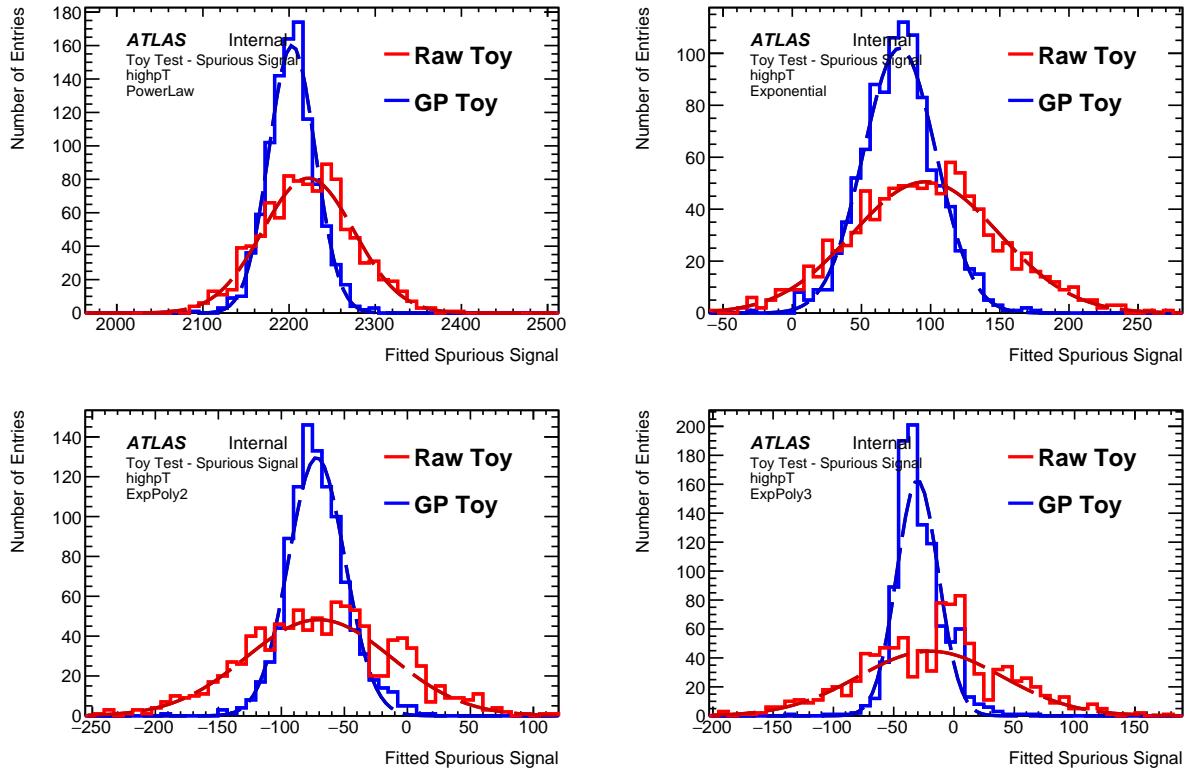


Figure A.71: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template. Each toy in this test has 10,000,000 events.

Event category	Func	$P(\chi^2)$ (%)	max S	$\frac{S}{\delta S}$ (%)	$\frac{\zeta}{\delta S}$ (%)	$\frac{S}{S_{ref}}$ (%)	$\frac{\zeta}{S_{ref}}$ (%)	MCStatUnc(%)
GG2H_0J_PTH_0_10_0	ExpPoly2	100	-69.7	-37.4	0	-8.69	0	0.50
GG2H_0J_PTH_GT10_0	ExpPoly2	100	-35.4	-9.56	0	-1.47	0	0.67
GG2H_1J_PTH_0_60	ExpPoly2	100	-36.4	-25.8	0	-5.99	0	0.70
GG2H_1J_PTH_60_120	ExpPoly2	100	33.6	27.2	0	6.41	0	0.63
GG2H_1J_PTH_120_200_0	ExpPoly2	100	1.36	7.49	0	3.31	0	0.39
GG2H_1J_PTH_120_200_1	Pow	99.7	-12.1	-45.7	-8.48	-22.4	-4.26	0.58
GG2H_GE2J_MJJ_0_350_PTH_0_60_0	ExpPoly2	100	-0.819	-2.04	0	-2.54	0	2.70
GG2H_GE2J_MJJ_0_350_PTH_0_60_1	ExpPoly2	100	19.1	21.1	0	14.4	0	0.41
GG2H_GE2J_MJJ_0_350_PTH_0_60_2	ExpPoly2	100	-42.7	-25.4	0	-7.84	0	0.43
GG2H_GE2J_MJJ_0_350_PTH_60_120_0	ExpPoly2	100	-0.136	56.9	14.8	29.4	7.56	0.39
GG2H_GE2J_MJJ_0_350_PTH_60_120_1	ExpPoly2	100	-2.76	-5.08	0	-2.6	0	15.46
GG2H_GE2J_MJJ_0_350_PTH_60_120_2	Exp	100	26.2	26.9	0	10.3	0	0.65
GG2H_GE2J_MJJ_0_350_PTH_120_200_0	ExpPoly2	100	0.222	1.29	0	0.555	0	0.46
GG2H_GE2J_MJJ_0_350_PTH_120_200_1	Pow	100	-8.76	-28	0	-11.5	0	1.14
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_0	Pow	100	0.275	6.23	0	6.39	0	16.90
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_1	Exp	100	1.18	9.74	0	7.1	0	1.62
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_2	Exp	100	-3.04	-14.9	0	-16.9	0	0.92
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_0	Exp	100	0.559	8.07	0	9.4	0	6.46
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_1	ExpPoly2	100	0.395	1.97	0	2.26	0	0.66
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_2	Exp	100	-8.92	-27.3	0	-36.2	0	0.87
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_0	Exp*	100	-0.0917	1.5	0	1.22	0	7.50
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_1	Exp	100	0.784	10.6	0	4.88	0	1.43
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_2	Pow	100	-0.896	-6.45	0	-5.65	0	11.90
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_0	Pow	100	0.809	12.9	0	13.6	0	2.60
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_1	Exp	100	2.08	15.1	0	13.4	0	1.13
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_2	Pow	100	3.39	14.8	0	19.2	0	0.84
GG2H_PTH_200_300_0	Exp*	100	0.714	18.9	0	8.7	0	0.96
GG2H_PTH_200_300_1	Exp	100	1.76	17.7	0	5.63	0	0.96
GG2H_PTH_200_300_2	Pow	100	1.07	7	0	3.99	0	6.08
GG2H_PTH_300_450_0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	1.07
GG2H_PTH_300_450_1	Pow*	100	0.137	3.88	0	1.69	0	4.80
GG2H_PTH_300_450_2	Pow	100	0.851	10.2	0	4.61	0	6.13
GG2H_PTH_450_650_0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	4.10
GG2H_PTH_450_650_1	Exp*	100	0.0262	0.906	0	1.25	0	114.44
GG2H_PTH_GT650_0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	9.32
GG2H_PTH_GT650_1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	20.50

Table A.3: The final background modelling decision and the size of spurious signal uncertainties. The reported number here is the base SS yield, without the bias uncertainty applied; the spurious signal with the bias is used in ?? and ???. In the mass range 120 GeV to 130 GeV, S is the maximum fitted spurious signal yield, δS is the associated uncertainty on the data, and S_{ref} is the expected size of Higgs signal events. The ζ is the maximum fitted spurious signal yield which is accommodate to 2σ statistical fluctuation of the background templates. The "*" in the function name indicates for which categories the "low-statistic" configuration of the SS fits (different in range and initial values) was run. Like the nominal case, we require $P(\chi^2) > 1\%$. The stat uncertainty quoted is the uncertainty on the template due to Monte Carlo statistics. The functional form is chosen using a relaxed spurious signal test applied to the unsmoothed templates.

Event category	Func	$P(\chi^2)$ (%)	max S	$\frac{S}{\delta S}$ (%)	$\frac{\zeta}{\delta S}$ (%)	$\frac{S}{S_{ref}}$ (%)	$\frac{\zeta}{S_{ref}}$ (%)	MCStatUnc(%)
QQ2HQQ_0J_0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	1.34
QQ2HQQ_0J_1	Exp*	100	-0.204	-5.8	0	-32.8	0	1.72
QQ2HQQ_1J_0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	10.01
QQ2HQQ_1J_1	Exp*	99.4	0.247	9.72	0	9.17	0	1.90
QQ2HQQ_1J_2	Pow	100	-0.67	-10.8	0	-12.5	0	-7.79
QQ2HQQ_GE2J_MJJ_0_60_0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	7.21
QQ2HQQ_GE2J_MJJ_0_60_1	Exp*	100	-0.0541	-1.71	0	-2.64	0	2.30
QQ2HQQ_GE2J_MJJ_0_60_2	Exp	100	0.221	2.64	0	4.08	0	2.64
QQ2HQQ_GE2J_MJJ_60_120_0	Exp*	100	0.0616	2.17	0	1.2	0	1.81
QQ2HQQ_GE2J_MJJ_60_120_1	Pow	100	0.216	3.46	0	3.2	0	3.37
QQ2HQQ_GE2J_MJJ_120_350_0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	3.79
QQ2HQQ_GE2J_MJJ_120_350_1	Exp	100	0.971	8.82	0	6.49	0	1.76
QQ2HQQ_GE2J_MJJ_120_350_2	Pow	100	2.75	13.2	0	9.22	0	1.07
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	-5.41
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_1	Exp	100	0.36	4.32	0	3.24	0	-4.97
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	5.20
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	6.51
QQ2HQQ_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	26.94
QQ2HQQ_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_1	Pow	100	0.802	5.55	0	1.7	0	2.26
QQ2HQQ_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	-120.22
QQ2HQQ_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	7.78
QQ2HQQ_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_2	Exp	100	0.953	19.7	0	13.1	0	1.48
QQ2HQQ_GE2J_MJJ_350_700_PTH_GT200_0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	-112.61
QQ2HQQ_GE2J_MJJ_350_700_PTH_GT200_1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	-3.41
QQ2HQQ_GE2J_MJJ_GT700_PTH_GT200_0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	-7.85
QQ2HQQ_GE2J_MJJ_GT700_PTH_GT200_1	Exp*	100	0.199	5.73	0	2.78	0	2.06
UNSELECTED_WH	Exp	100	-1.01	-6.68	0	-12.7	0	-1.37
QQ2HLNU_PTV_0.75_0	Exp*	100	0.583	2.76	0	2.52	0	2.46
QQ2HLNU_PTV_0.75_1	Exp	100	-0.152	-2.01	0	-2.44	0	-3.80
QQ2HLNU_PTV_75_150_0	Exp*	100	-0.00665	-0.351	0	-0.177	0	-14.07
QQ2HLNU_PTV_75_150_1	Exp*	100	0.122	4.78	0	9.35	0	0.82
QQ2HLNU_PTV_150_250_0J_0	Exp*	100	0.0655	4.15	0	3.54	0	3.29
QQ2HLNU_PTV_150_250_GE1J_0	Exp*	100	0.0461	2.77	0	2	0	3.13
QQ2HLNU_PTV_GT250_0	Exp*	100	0.0123	1.12	0	0.835	0	3.23
UNSELECTED_ZH	Exp	100	4.25	21.7	0	34.9	0	1.48
HLL_PTV_0.75_0	Exp*	100	-0.0253	-1.85	0	-2.85	0	6.89
HLL_PTV_75_150_0	Exp*	14.8	0.236	11.5	0	6.75	0	3.21
HLL_PTV_75_150_1	Exp	100	1.28	22.3	0	20.3	0	1.20
HLL_PTV_150_250_0J_0	Exp*	100	0.291	18.4	0	16.2	0	1.75
HLL_PTV_150_250_GE1J_0	Exp*	100	-0.00609	-0.328	0	-0.363	0	3.55
HLL_PTV_GT250_0	Exp*	100	0.0515	3.43	0	3.14	0	3.31
UNSELECTED_TOP	Exp	100	1.99	17.1	0	16	0	1.14
TTH_PTH_0_60_0	Exp*	100	-0.0813	-3.23	0	-2.5	0	-2.99
TTH_PTH_0_60_1	Exp*	100	-0.138	-2.81	0	-4.09	0	-3.02
TTH_PTH_60_120_0	Exp*	100	-0.0321	-1.2	0	-0.653	0	2.08
TTH_PTH_60_120_1	Exp*	100	0.329	8.65	0	8.21	0	0.86
TTH_PTH_120_200_0	Exp*	100	0.0593	2.35	0	1.02	0	1.59
TTH_PTH_120_200_1	Exp*	100	0.195	5.6	0	6.61	0	2.18
TTH_PTH_200_300_0	Exp*	100	0.393	2.56	0	0.884	0	2.99
TTH_PTH_GT300_0	Exp*	100	0.00755	0.562	0	0.181	0	-174.49
THJB_0	Exp*	100	0.055	3.05	0	7.09	0	1.28
TWH_0	Exp*	100	-0.0481	-2.74	0	-5.6	0	-1.76

Table A.4: The final background modelling decision and the size of spurious signal uncertainties. The reported number here is the base SS yield, without the bias uncertainty applied; the spurious signal with the bias is used in ?? and ???. In the mass range 120 GeV to 130 GeV, S is the maximum fitted spurious signal yield, δS is the associated uncertainty on the data, and S_{ref} is the expected size of Higgs signal events. The ζ is the maximum fitted spurious signal yield which is accommodate to 2σ statistical fluctuation of the background templates. The “*” in the function name indicates for which categories the “low-statistic” configuration of the SS fits (different in range and initial values) was run. Like the nominal case, we require $P(\chi^2) > 1\%$. The stat uncertainty quoted is the uncertainty on the template due to Monte Carlo statistics. The functional form is chosen using a relaxed spurious signal test applied to the unsmoothed templates.

Event category	Function	$\max(S)$	
		Nominal	Smooth temp
GG2H_0J_PTH_0_10_0	ExpPoly2	-117	-69.7
GG2H_0J_PTH_GT10_0	ExpPoly2	-199	-35.4
GG2H_1J_PTH_0_60	ExpPoly2	-67.1	-36.4
GG2H_1J_PTH_60_120	ExpPoly2	28.7	33.6
GG2H_1J_PTH_120_200_0	ExpPoly2	-1.79	1.36
GG2H_1J_PTH_120_200_1	Pow	-11.7	-12.1
GG2H_GE2J_MJJ_0_350_PTH_0_60_0	ExpPoly2	6.54	-0.819
GG2H_GE2J_MJJ_0_350_PTH_0_60_1	ExpPoly2	21.4	19.1
GG2H_GE2J_MJJ_0_350_PTH_0_60_2	ExpPoly2	-78.6	-42.7
GG2H_GE2J_MJJ_0_350_PTH_60_120_0	ExpPoly2	7.01	-0.136
GG2H_GE2J_MJJ_0_350_PTH_60_120_1	ExpPoly2	7.04	-2.76
GG2H_GE2J_MJJ_0_350_PTH_60_120_2	Exp	59.8	26.2
GG2H_GE2J_MJJ_0_350_PTH_120_200_0	ExpPoly2	7.8	0.222
GG2H_GE2J_MJJ_0_350_PTH_120_200_1	Pow	-15.4	-8.76
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_0	Pow	-2.83	0.275
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_1	Exp	-1.53	1.18
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_2	Exp	-5.05	-3.04
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_0	Exp	2.08	0.559
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_1	ExpPoly2	7.73	0.395
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_2	Exp	-17.6	-8.92
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_0	Exp*	1.13	-0.0917
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_1	Exp	4.76	0.784
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_2	Pow	-2.44	-0.896
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_0	Pow	-1.69	0.809
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_1	Exp	1.82	2.08
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_2	Pow	5.52	3.39
GG2H_PTH_200_300_0	Exp*	0.8	0.714
GG2H_PTH_200_300_1	Exp	4.11	1.76
GG2H_PTH_200_300_2	Pow	2.62	1.07
GG2H_PTH_300_450_0	Exp*	0.34	N/A
GG2H_PTH_300_450_1	Pow*	-0.81	0.137
GG2H_PTH_300_450_2	Pow	-3.49	0.851
GG2H_PTH_450_650_0	Exp*	-0.67	N/A
GG2H_PTH_450_650_1	Exp*	-0.96	0.0262

Table A.5: Comparison of the SS test (function and systematic uncertainty assigned) with nominal un-smoothed templates and smoothed ones. The functional form is chosen using a relaxed spurious signal test applied to the unsmoothed templates.

Event category	Function	$\max(S)$	
		Nominal	Smooth temp
GG2H_PTH_GT650_0	Exp*	0.63	N/A
GG2H_PTH_GT650_1	Exp*	-0.36	N/A
QQ2HQQ_0J_0	Exp*	-0.68	N/A
QQ2HQQ_0J_1	Exp*	-0.33	-0.204
QQ2HQQ_1J_0	Exp*	-0.53	N/A
QQ2HQQ_1J_1	Exp*	0.44	0.247
QQ2HQQ_1J_2	Pow	-1.35	-0.67
QQ2HQQ_GE2J_MJJ_0_60_0	Exp*	0.64	N/A
QQ2HQQ_GE2J_MJJ_0_60_1	Exp*	-0.39	-0.0541
QQ2HQQ_GE2J_MJJ_0_60_2	Exp	-1.51	0.221
QQ2HQQ_GE2J_MJJ_60_120_0	Exp*	0.66	0.0616
QQ2HQQ_GE2J_MJJ_60_120_1	Pow	-2.35	0.216
QQ2HQQ_GE2J_MJJ_120_350_0	Exp*	-0.6	N/A
QQ2HQQ_GE2J_MJJ_120_350_1	Exp	1.13	0.971
QQ2HQQ_GE2J_MJJ_120_350_2	Pow	-7.49	2.75
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_0	Exp*	-0.25	N/A
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_1	Exp	1.66	0.36
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_0	Exp*	0.38	N/A
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_1	Exp*	-1.06	N/A
QQ2HQQ_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_0	Exp*	-1.46	N/A
QQ2HQQ_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_1	Pow	-2.2	0.802
QQ2HQQ_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_0	Exp*	1.25	N/A
QQ2HQQ_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_1	Exp*	-0.45	N/A
QQ2HQQ_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_2	Exp	1.69	0.953
QQ2HQQ_GE2J_MJJ_350_700_PTH_GT200_0	Exp*	-0.31	N/A
QQ2HQQ_GE2J_MJJ_350_700_PTH_GT200_1	Exp*	-0.38	N/A
QQ2HQQ_GE2J_MJJ_GT700_PTH_GT200_0	Exp*	1.24	N/A
QQ2HQQ_GE2J_MJJ_GT700_PTH_GT200_1	Exp*	1.89	0.199
UNSELECTED_WH	Exp	-2.69	-1.01
QQ2HLNU_PTV_0_75_0	Exp*	-0.14	0.583
QQ2HLNU_PTV_0_75_1	Exp	-0.69	-0.152
QQ2HLNU_PTV_75_150_0	Exp*	1.37	-0.00665
QQ2HLNU_PTV_75_150_1	Exp*	-0.42	0.122
QQ2HLNU_PTV_150_250_0J_0	Exp*	-0.4	0.0655
QQ2HLNU_PTV_150_250_GE1J_0	Exp*	-0.18	0.0461
QQ2HLNU_PTV_GT250_0	Exp*	-0.08	0.0123
UNSELECTED_ZH	Exp	9.43	4.25
HLL_PTV_0_75_0	Exp*	-0.06	-0.0253
HLL_PTV_75_150_0	Exp*	-0.16	0.236
HLL_PTV_75_150_1	Exp	1.11	1.28
HLL_PTV_150_250_0J_0	Exp*	0.4	0.291
HLL_PTV_150_250_GE1J_0	Exp*	1.14	-0.00609
HLL_PTV_GT250_0	Exp*	-0.34	0.0515
UNSELECTED_TOP	Exp	2.02	1.99
TTH_PTH_0_60_0	Exp*	-0.15	-0.0813
TTH_PTH_0_60_1	Exp*	-0.75	-0.138
TTH_PTH_60_120_0	Exp*	0.1	-0.0321
TTH_PTH_60_120_1	Exp*	0.57	0.329
TTH_PTH_120_200_0	Exp*	-0.36	0.0593
TTH_PTH_120_200_1	Exp*	0.68	0.195
TTH_PTH_200_300_0	Exp*	0.17	0.393
TTH_PTH_GT300_0	Exp*	0.13	0.00755
THJB_0	Exp*	0.3	0.055
TWH_0	Exp*	0.17	-0.0481

Table A.6: Comparison of the SS test (function `h78` systematic uncertainty assigned) with nominal un-smoothed templates and smoothed ones. The functional form is chosen using a relaxed spurious signal test applied to the unsmoothed templates.

Event category	Func	$P(\chi^2)$ (%)	max S	$\frac{S}{\delta S}$ (%)	$\frac{S}{S_{ref}}$ (%)	MCStatUnc(%)
GG2H_0J_PTH_0_10_0	ExpPoly2	100	-69.7	-37.4	-8.69	0.50
GG2H_0J_PTH_GT10_0	ExpPoly2	100	-35.4	-9.56	-1.47	0.67
GG2H_1J_PTH_0_60	ExpPoly2	100	-36.4	-25.8	-5.99	0.70
GG2H_1J_PTH_60_120	ExpPoly2	100	33.6	27.2	6.41	0.63
GG2H_1J_PTH_120_200_0	ExpPoly2	100	1.37	7.53	3.32	0.39
GG2H_1J_PTH_120_200_1	ExpPoly2	100	-4.97	-17.4	-9.11	0.58
GG2H_GE2J_MJJ_0_350_PTH_0_60_0	ExpPoly2	100	-0.784	-1.98	-2.05	2.70
GG2H_GE2J_MJJ_0_350_PTH_0_60_1	Bern3	100	-14.8	-17.2	-11.1	0.41
GG2H_GE2J_MJJ_0_350_PTH_0_60_2	ExpPoly2	100	-42.7	-25.4	-7.84	0.43
GG2H_GE2J_MJJ_0_350_PTH_60_120_0	ExpPoly2	100	-0.136	-0.601	-0.394	0.39
GG2H_GE2J_MJJ_0_350_PTH_60_120_1	ExpPoly2	100	-2.76	-5.08	-2.6	15.46
GG2H_GE2J_MJJ_0_350_PTH_60_120_2	ExpPoly2	100	-19.6	-20.5	-7.69	0.65
GG2H_GE2J_MJJ_0_350_PTH_120_200_0	ExpPoly2	100	0.222	1.29	0.555	0.46
GG2H_GE2J_MJJ_0_350_PTH_120_200_1	ExpPoly2	100	-5.24	-15.7	-6.82	1.14
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_0	Exp	100	0.166	3.75	3.87	16.90
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_1	Exp	100	1.18	9.74	7.1	1.62
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_2	Exp	100	-3.04	-14.9	-16.9	0.92
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_0	Exp	100	0.559	8.07	9.4	6.46
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_1	ExpPoly2	100	0.395	1.97	2.26	0.66
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_2	ExpPoly2	100	-3.14	-8.71	-12.8	0.87
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_0	Pow*	100	0.0505	1.5	1.22	7.50
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_1	Exp	100	0.784	10.6	4.88	1.43
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_2	Pow	100	-0.896	-6.45	-5.65	11.90
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_0	Exp	100	0.28	4.4	5.02	2.60
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_1	Exp	100	2.08	15.1	13.4	1.13
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_2	Pow	100	3.39	14.8	19.2	0.84
GG2H_PTH_200_300_0	Exp*	100	0.714	18.9	8.7	0.96
GG2H_PTH_200_300_1	Exp	100	1.76	17.7	5.63	0.96
GG2H_PTH_200_300_2	Exp	100	-0.0951	-0.622	-0.359	6.08
GG2H_PTH_300_450_0	N/A	N/A	N/A	N/A	N/A	1.07
GG2H_PTH_300_450_1	Exp*	100	0.0274	0.743	0.372	4.80
GG2H_PTH_300_450_2	Exp	100	0.491	5.91	2.94	6.13
GG2H_PTH_450_650_0	N/A	N/A	N/A	N/A	N/A	4.10
GG2H_PTH_450_650_1	Exp*	100	0.0297	1.02	1.42	114.44
GG2H_PTH_GT650_0	N/A	N/A	N/A	N/A	N/A	9.32
GG2H_PTH_GT650_1	N/A	N/A	N/A	N/A	N/A	20.50

Table A.7: The final background modelling decision and the size of spurious signal uncertainties. The reported number here is the base SS yield, without the bias uncertainty applied; the spurious signal with the bias is used in ?? and ???. In the mass range 120 GeV to 130 GeV, S is the maximum fitted spurious signal yield, δS is the associated uncertainty on the data, and S_{ref} is the expected size of Higgs signal events. The ζ is the maximum fitted spurious signal yield which is accommodate to 2σ statistical fluctuation of the background templates. The “*” in the function name indicates for which categories the “low-statistic” configuration of the SS fits (different in range and initial values) was run. Like the nominal case, we require $P(\chi^2) > 1\%$. The stat uncertainty quoted is the uncertainty on the template due to Monte Carlo statistics. The functional form is chosen using a non-relaxed spurious signal test applied to the smoothed templates.

Event category	Func	$P(\chi^2)$ (%)	max S	$\frac{S}{\delta S}$ (%)	$\frac{S}{S_{ref}}$ (%)	MCStatUnc(%)
QQ2HQQ_0J_0	N/A	N/A	N/A	N/A	N/A	1.34
QQ2HQQ_0J_1	Exp*	100	-0.204	-5.8	-32.8	1.72
QQ2HQQ_1J_0	N/A	N/A	N/A	N/A	N/A	10.01
QQ2HQQ_1J_1	Exp*	100	0.247	9.72	9.17	1.90
QQ2HQQ_1J_2	Pow	100	-0.67	-10.8	-12.5	-7.79
QQ2HQQ_GE2J_MJJ_0.60_0	N/A	N/A	N/A	N/A	N/A	7.21
QQ2HQQ_GE2J_MJJ_0.60_1	Exp*	100	-0.0541	-1.71	-2.64	2.30
QQ2HQQ_GE2J_MJJ_0.60_2	Exp	100	0.221	2.64	4.08	2.64
QQ2HQQ_GE2J_MJJ_60_120_0	Exp*	100	0.0616	2.17	1.2	1.81
QQ2HQQ_GE2J_MJJ_60_120_1	Exp	100	-0.203	-3.47	-3.01	3.37
QQ2HQQ_GE2J_MJJ_120_350_0	N/A	N/A	N/A	N/A	N/A	3.79
QQ2HQQ_GE2J_MJJ_120_350_1	Exp	100	0.971	8.82	6.49	1.76
QQ2HQQ_GE2J_MJJ_120_350_2	Pow	100	2.75	13.2	9.22	1.07
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_0	N/A	N/A	N/A	N/A	N/A	-5.41
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_1	Exp	100	0.36	4.32	3.24	-4.97
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_0	N/A	N/A	N/A	N/A	N/A	5.20
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_1	N/A	N/A	N/A	N/A	N/A	6.51
QQ2HQQ_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_0	N/A	N/A	N/A	N/A	N/A	26.94
QQ2HQQ_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_1	Exp	100	0.315	5.55	1.7	2.26
QQ2HQQ_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_0	N/A	N/A	N/A	N/A	N/A	-120.22
QQ2HQQ_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_1	N/A	N/A	N/A	N/A	N/A	7.78
QQ2HQQ_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_2	Exp	100	0.953	19.7	13.1	1.48
QQ2HQQ_GE2J_MJJ_350_700_PTH_GT200_0	N/A	N/A	N/A	N/A	N/A	-112.61
QQ2HQQ_GE2J_MJJ_350_700_PTH_GT200_1	N/A	N/A	N/A	N/A	N/A	-3.41
QQ2HQQ_GE2J_MJJ_GT700_PTH_GT200_0	N/A	N/A	N/A	N/A	N/A	-7.85
QQ2HQQ_GE2J_MJJ_GT700_PTH_GT200_1	Exp*	100	0.199	5.73	2.78	2.06
UNSELECTED_WH	Exp	100	-1.01	-6.68	-12.7	-1.37
QQ2HLNU_PTV_0_75_0	Exp*	100	0.583	2.76	2.52	2.46
QQ2HLNU_PTV_0_75_1	Exp	100	-0.152	-2.01	-2.44	-3.80
QQ2HLNU_PTV_75_150_0	Exp*	100	-0.00665	-0.351	-0.177	-14.07
QQ2HLNU_PTV_75_150_1	Exp*	100	0.122	4.78	9.35	0.82
QQ2HLNU_PTV_150_250_0J_0	Exp*	100	0.0655	4.15	3.54	3.29
QQ2HLNU_PTV_150_250_GE1J_0	Exp*	100	0.0461	2.77	2	3.13
QQ2HLNU_PTV_GT250_0	Exp*	100	0.0123	1.12	0.835	3.23
UNSELECTED_ZH	ExpPoly2	100	2.18	10.5	18.2	1.48
HLL_PTV_0_75_0	Exp*	100	-0.0253	-1.85	-2.85	6.89
HLL_PTV_75_150_0	Exp*	14.8	0.236	11.5	6.75	3.21
HLL_PTV_75_150_1	ExpPoly2*	100	-0.0891	-1.58	-2.12	1.20
HLL_PTV_150_250_0J_0	Exp*	72.3	0.291	18.4	16.2	1.75
HLL_PTV_150_250_GE1J_0	Pow*	100	-0.00609	-0.328	-0.363	3.55
HLL_PTV_GT250_0	Exp*	100	0.0515	3.43	3.14	3.31
UNSELECTED_TOP	Exp	100	1.99	17.1	16	1.14
TTH_PTH_0.60_0	Exp*	100	-0.0813	-3.23	-2.5	-2.99
TTH_PTH_0.60_1	Exp*	100	-0.138	-2.81	-4.09	-3.02
TTH_PTH_60_120_0	Exp*	100	-0.0321	-1.2	-0.653	2.08
TTH_PTH_60_120_1	Exp*	100	0.329	8.65	8.21	0.86
TTH_PTH_120_200_0	Exp*	100	0.0593	2.35	1.02	1.59
TTH_PTH_120_200_1	Exp*	100	0.195	5.6	6.61	2.18
TTH_PTH_200_300_0	Exp*	100	0.0393	2.56	0.884	2.99
TTH_PTH_GT300_0	Exp*	100	0.00755	0.562	0.181	-174.49
THJB_0	Exp*	100	0.055	3.05	7.09	1.28
TWH_0	Exp*	100	-0.0481	-2.74	-5.6	-1.76

Table A.8: The final background modelling decision and the size of spurious signal uncertainties. The reported number here is the base SS yield, without the bias uncertainty applied; the spurious signal with the bias is used in ?? and ???. In the mass range 120 GeV to 130 GeV, S is the maximum fitted spurious signal yield, δS is the associated uncertainty on the data, and S_{ref} is the expected size of Higgs signal events. The ζ is the maximum fitted spurious signal yield which is accommodate to 2σ statistical fluctuation of the background templates. The “*” in the function name indicates for which categories the “low-statistic” configuration of the SS fits (different in range and initial values) was run. Like the nominal case, we require $P(\chi^2) > 1\%$. The stat uncertainty quoted is the uncertainty on the template due to Monte Carlo statistics. The functional form is chosen using a non-relaxed spurious signal test applied to the smoothed templates.

Event category	$\max(S)$		$\max(S)$	
	Nominal	Smooth temp	Nominal	Smooth temp
GG2H_0J_PTH_0_10_0	ExpPoly2	ExpPoly2	-117	-69.7
GG2H_0J_PTH_GT10_0	ExpPoly2	ExpPoly2	-199	-35.4
GG2H_1J_PTH_0_60	ExpPoly2	ExpPoly2	-67.1	-36.4
GG2H_1J_PTH_60_120	ExpPoly2	ExpPoly2	28.7	33.6
GG2H_1J_PTH_120_200_0	ExpPoly2	ExpPoly2	-1.79	1.37
GG2H_1J_PTH_120_200_1	Pow	ExpPoly2	-11.7	-4.97
GG2H_GE2J_MJJ_0_350_PTH_0_60_0	ExpPoly2	ExpPoly2	6.54	-0.784
GG2H_GE2J_MJJ_0_350_PTH_0_60_1	ExpPoly2	Bern3	21.4	-14.8
GG2H_GE2J_MJJ_0_350_PTH_0_60_2	ExpPoly2	ExpPoly2	-78.6	-42.7
GG2H_GE2J_MJJ_0_350_PTH_60_120_0	ExpPoly2	ExpPoly2	7.01	-0.136
GG2H_GE2J_MJJ_0_350_PTH_60_120_1	ExpPoly2	ExpPoly2	7.04	-2.76
GG2H_GE2J_MJJ_0_350_PTH_60_120_2	Exp	ExpPoly2	59.8	-19.6
GG2H_GE2J_MJJ_0_350_PTH_120_200_0	ExpPoly2	ExpPoly2	7.8	0.222
GG2H_GE2J_MJJ_0_350_PTH_120_200_1	Pow	ExpPoly2	-15.4	-5.24
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_0	Pow	Exp	-2.83	0.166
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_1	Exp	Exp	-1.53	1.18
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_2	Exp	Exp	-5.05	-3.04
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_0	Exp	Exp	2.08	0.559
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_1	ExpPoly2	ExpPoly2	7.73	0.395
GG2H_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_2	Exp	ExpPoly2	-17.6	-3.14
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_0	Exp*	Pow*	1.13	0.0505
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_1	Exp	Exp	4.76	0.784
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_0_25_2	Pow	Pow	-2.44	-0.896
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_0	Pow	Exp	-1.69	0.28
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_1	Exp	Exp	1.82	2.08
GG2H_GE2J_MJJ_GT700_PTH_0_200_PTHJJ_GT25_2	Pow	Pow	5.52	3.39
GG2H_PTH_200_300_0	Exp*	Exp*	0.8	0.714
GG2H_PTH_200_300_1	Exp	Exp	4.11	1.76
GG2H_PTH_200_300_2	Pow	Exp	2.62	-0.0951
GG2H_PTH_300_450_0	Exp*	N/A	0.34	N/A
GG2H_PTH_300_450_1	Pow*	Exp*	-0.81	0.0274
GG2H_PTH_300_450_2	Pow	Exp	-3.49	0.491
GG2H_PTH_450_650_0	Exp*	N/A	-0.67	N/A
GG2H_PTH_450_650_1	Exp*	Exp*	-0.96	0.0297
GG2H_PTH_GT650_0	Exp*	N/A	0.63	N/A
GG2H_PTH_GT650_1	Exp*	N/A	-0.36	N/A

Table A.9: Comparison of the SS test (function and systematic uncertainty assigned) with nominal un-smoothed templates and smoothed ones. The functional form is chosen using a non-relaxed spurious signal test applied to the smoothed templates.

Event category	$\max(S)$		$\max(S)$	
	Nominal	Smooth temp	Nominal	Smooth temp
QQ2HQQ_0J_0	Exp*	N/A	-0.68	N/A
QQ2HQQ_0J_1	Exp*	Exp*	-0.33	-0.204
QQ2HQQ_1J_0	Exp*	N/A	-0.53	N/A
QQ2HQQ_1J_1	Exp*	Exp*	0.44	0.247
QQ2HQQ_1J_2	Pow	Pow	-1.35	-0.67
QQ2HQQ_GE2J_MJJ_0_60_0	Exp*	N/A	0.64	N/A
QQ2HQQ_GE2J_MJJ_0_60_1	Exp*	Exp*	-0.39	-0.0541
QQ2HQQ_GE2J_MJJ_0_60_2	Exp	Exp	-1.51	0.221
QQ2HQQ_GE2J_MJJ_60_120_0	Exp*	Exp*	0.66	0.0616
QQ2HQQ_GE2J_MJJ_60_120_1	Pow	Exp	-2.35	-0.203
QQ2HQQ_GE2J_MJJ_120_350_0	Exp*	N/A	-0.6	N/A
QQ2HQQ_GE2J_MJJ_120_350_1	Exp	Exp	1.13	0.971
QQ2HQQ_GE2J_MJJ_120_350_2	Pow	Pow	-7.49	2.75
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_0	Exp*	N/A	-0.25	N/A
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_0_25_1	Exp	Exp	1.66	0.36
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_0	Exp*	N/A	0.38	N/A
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200_PTHJJ_GT25_1	Exp*	N/A	-1.06	N/A
QQ2HQQ_GE2J_MJJ_G700_PTH_0_200_PTHJJ_0_25_0	Exp*	N/A	-1.46	N/A
QQ2HQQ_GE2J_MJJ_G700_PTH_0_200_PTHJJ_0_25_1	Pow	Exp	-2.2	0.315
QQ2HQQ_GE2J_MJJ_G700_PTH_0_200_PTHJJ_GT25_0	Exp*	N/A	1.25	N/A
QQ2HQQ_GE2J_MJJ_G700_PTH_0_200_PTHJJ_GT25_1	Exp*	N/A	-0.45	N/A
QQ2HQQ_GE2J_MJJ_G700_PTH_0_200_PTHJJ_GT25_2	Exp	Exp	1.69	0.953
QQ2HQQ_GE2J_MJJ_350_700_PTH_GT200_0	Exp*	N/A	-0.31	N/A
QQ2HQQ_GE2J_MJJ_350_700_PTH_GT200_1	Exp*	N/A	-0.38	N/A
QQ2HQQ_GE2J_MJJ_G700_PTH_GT200_0	Exp*	N/A	1.24	N/A
QQ2HQQ_GE2J_MJJ_G700_PTH_GT200_1	Exp*	Exp*	1.89	0.199
UNSELECTED_WH	Exp	Exp	-2.69	-1.01
QQ2HLNU_PTV_0_75_0	Exp*	Exp*	-0.14	0.583
QQ2HLNU_PTV_0_75_1	Exp	Exp	-0.69	-0.152
QQ2HLNU_PTV_75_150_0	Exp*	Exp*	1.37	-0.00665
QQ2HLNU_PTV_75_150_1	Exp*	Exp*	-0.42	0.122
QQ2HLNU_PTV_150_250_0J_0	Exp*	Exp*	-0.4	0.0655
QQ2HLNU_PTV_150_250_GE1J_0	Exp*	Exp*	-0.18	0.0461
QQ2HLNU_PTV_GT250_0	Exp*	Exp*	-0.08	0.0123
UNSELECTED_ZH	Exp	ExpPoly2	9.43	2.18
HLL_PTV_0_75_0	Exp*	Exp*	-0.06	-0.0253
HLL_PTV_75_150_0	Exp*	Exp*	-0.16	0.236
HLL_PTV_75_150_1	Exp	ExpPoly2*	1.11	-0.0891
HLL_PTV_150_250_0J_0	Exp*	Exp*	0.4	0.291
HLL_PTV_150_250_GE1J_0	Exp*	Pow*	1.14	-0.00609
HLL_PTV_GT250_0	Exp*	Exp*	-0.34	0.0515
UNSELECTED_TOP	Exp	Exp	2.02	1.99
TTH_PTH_0_60_0	Exp*	Exp*	-0.15	-0.0813
TTH_PTH_0_60_1	Exp*	Exp*	-0.75	-0.138
TTH_PTH_60_120_0	Exp*	Exp*	0.1	-0.0321
TTH_PTH_60_120_1	Exp*	Exp*	0.57	0.329
TTH_PTH_120_200_0	Exp*	Exp*	-0.36	0.0593
TTH_PTH_120_200_1	Exp*	Exp*	0.68	0.195
TTH_PTH_200_300_0	Exp*	Exp*	0.17	0.0393
TTH_PTH_GT300_0	Exp*	Exp*	0.13	0.00755
THJB_0	Exp*	Exp*	0.3	0.055
TWH_0	Exp*	Exp*	0.17	-0.0481

Table A.10: Comparison of the SS test (function and systematic uncertainty assigned) with nominal un-smoothed templates and smoothed ones. The functional form is chosen using a non-relaxed spurious signal test applied to the smoothed templates.

Nominal	N_sig	Unweighted Events	Bkg events weighted	Mean SS_raw	Mean SS_GPR	GP-Raw: Bias, Mean	Sigma Toy	Sigma GPR	Bias/sigma(SS_raw)
Fit: Exp									
Generating: Exp2									
low10	0	10^1	0.4	-0.09	-0.02	0.07	0.18	0.08	0.37
low100	0	10^2	4	-0.00	-0.06	-0.06	0.16	0.58	-0.37
low1k	0	10^3	4×10^1	-0.08	0.02	0.10	0.52	0.22	0.19
low10k	0	10^4	4×10^2	-0.81	-0.57	0.24	1.62	0.79	0.15
low100k	0	10^5	4×10^3	-8.66	-7.18	1.48	5.22	2.06	0.28
low1M	0	10^6	4×10^4	-87.70	-85.61	2.09	15.67	9.12	0.13
low10M	0	10^7	4×10^5	-871.75	-876.58	-4.83	50.32	29.77	-0.10
Fit: Exp2									
Generating: Exp3									
med10	0	10^1	0.4	-0.10	-0.01	0.09	0.19	0.05	0.49
med100	0	10^2	4	-0.00	0.04	0.04	0.17	0.24	0.22
med1k	0	10^3	4×10^1	-0.01	-0.04	-0.03	0.56	0.23	-0.05
med10k	0	10^4	4×10^2	-0.01	-0.03	-0.01	1.78	0.39	-0.01
med100k	0	10^5	4×10^3	-0.27	-0.89	-0.62	5.68	1.64	-0.11
med1M	0	10^6	4×10^4	-1.95	-4.96	-3.01	16.90	5.69	-0.18
med10M	0	10^7	4×10^5	-20.43	-26.76	-6.33	55.29	23.06	-0.11
Fit: Exp2									
Generating: Exp3									
high10	0	10^1	0.4	-0.10	-0.01	0.10	0.19	0.04	0.50
high100	0	10^2	4	-0.01	0.02	0.03	0.18	0.25	0.17
high1k	0	10^3	4×10^1	0.00	-0.03	-0.03	0.55	0.32	-0.06
high10k	0	10^4	4×10^2	-0.13	-0.14	-0.00	1.74	0.39	-0.00
high100k	0	10^5	4×10^3	-0.91	-1.63	-0.71	5.47	1.86	-0.13
high1M	0	10^6	4×10^4	-6.72	-11.33	-4.61	17.92	10.22	-0.26
high10M	0	10^7	4×10^5	-67.99	-71.01	-3.02	58.73	24.19	-0.05

Table A.11: Spurious signal means and widths for the three test functional-form distributions for a range of different template statistics.

To determine how to reduce the bias further, we note a further set of tests performed in an earlier iteration of this method [?], evaluating the difference in GP fit bias when different functional priors were used as the GP mean. The Exponential function, Linear function, and a flat line were all tested. These tests were performed using templates constructed from Power Law (Fig. ??), ExpPoly2 (Fig. ??), and Bernstein 5 (Fig. ??) functions. Again, different levels of statistics in the templates were tested. Overall, the choice of GP mean does not seem to affect the GP fit behavior significantly, other than for templates with less than 10 effective MC events per bin. The unit of the y-axis in these plots is the percentage disagreement between the smoothed and the unsmoothed template, similar to a ratio plot.

In the lower statistics templates, most of the differences in fitting bias are again confined to the extremities of the template range, as expected, and can be reduced by padding the template with dummy bins. Furthermore, the use of the χ^2 method to automatically select the flat prior in lower-statistics regions means that, in regions where the choice of prior matters, we are automatically choosing the one that best fits the data. Thus, the choice of prior does not appear to introduce a significant bias in the fit result.

A.5.2 Extended Templates

Some disagreement between the nominal template and the GP template is observed at the edges of the plots. To reduce this, larger templates can be provided to the GP- that is, we perform the GP fit on templates in the mass range 100-165 GeV, but perform the spurious signal test on the smoothed and unsmoothed template in the range 105-160 GeV, thus relegating the edge effects to the outer 5 bins. We show that extending the templates in this manner reduces the bias observed from 40% to 25% of the statistical uncertainty on the spurious signal.

To approximate this extension, we can also perform a linear fit padding on the templates, extending the templates by 5 bins on either side based on the slope of the first and last 5 bins.

We report the results of the padded-template bias study for all categories in the Table ???. Additionally, we plot the per-toy bias (that is, GP_SS - Raw_SS for each toy) for each shape and stat regime.

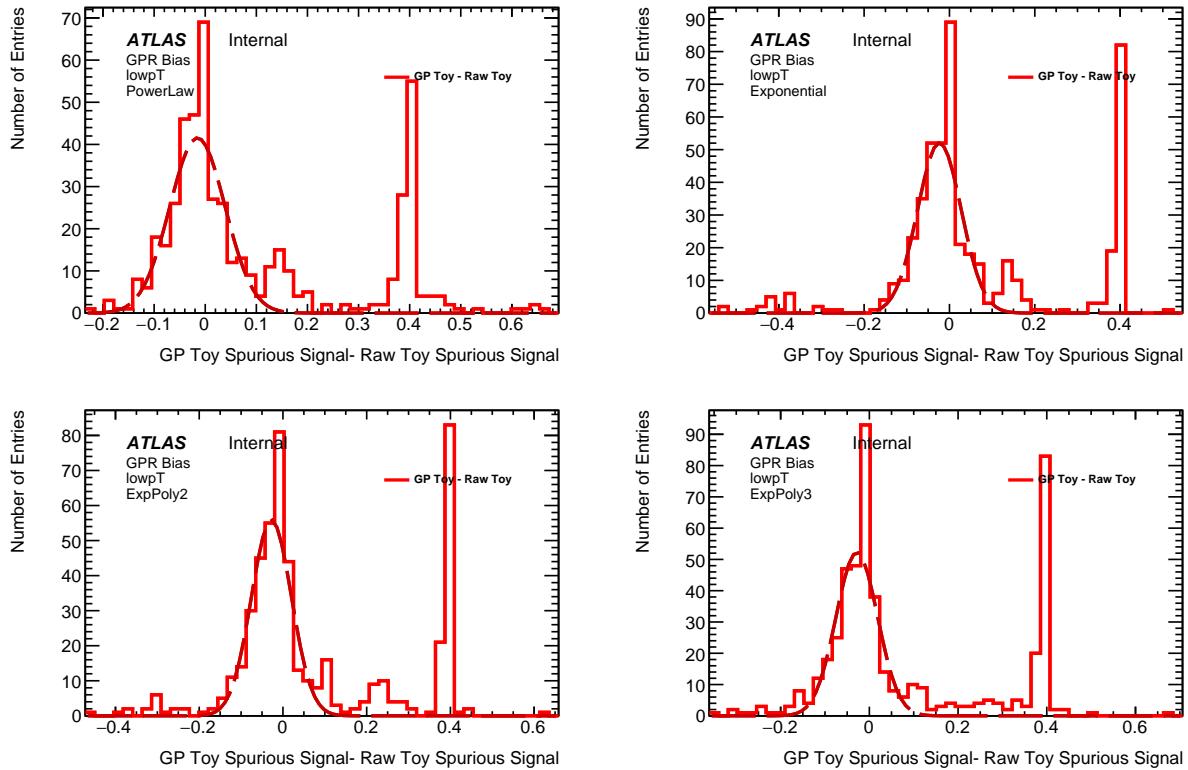


Figure A.72: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template. Each toy in this test has 10 events.

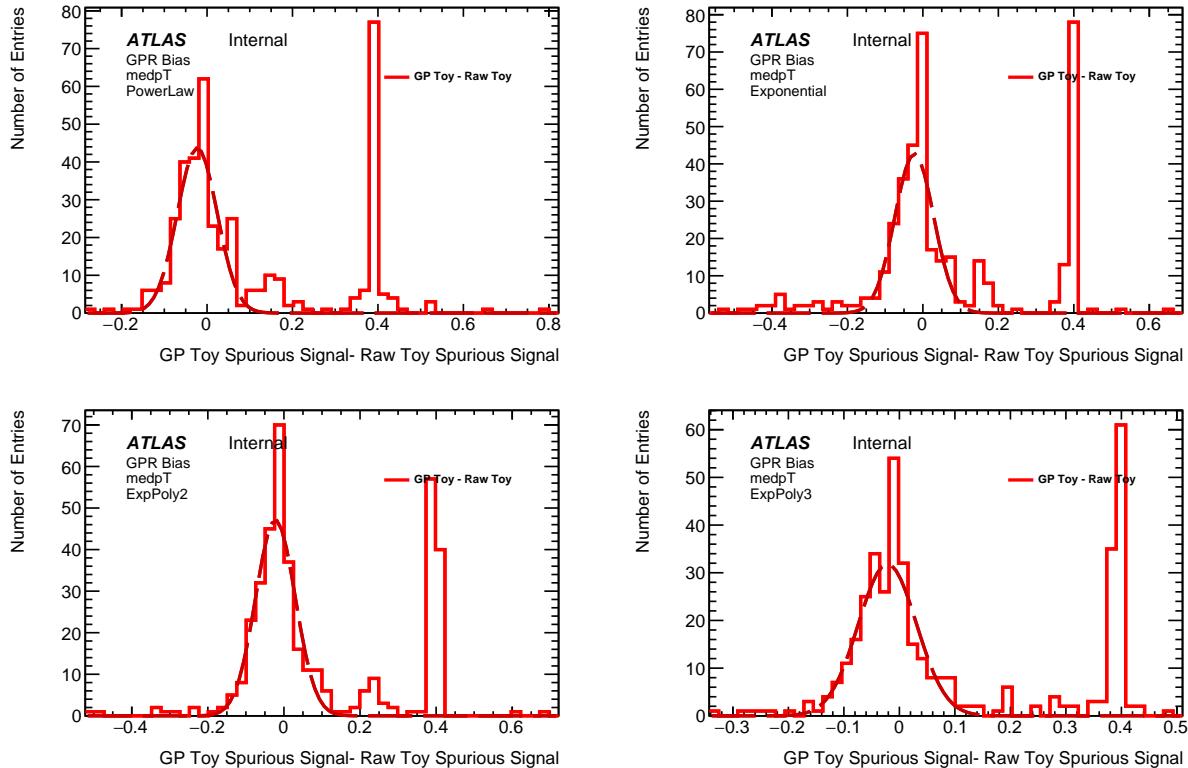


Figure A.73: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template. Each toy in this test has 10 events.

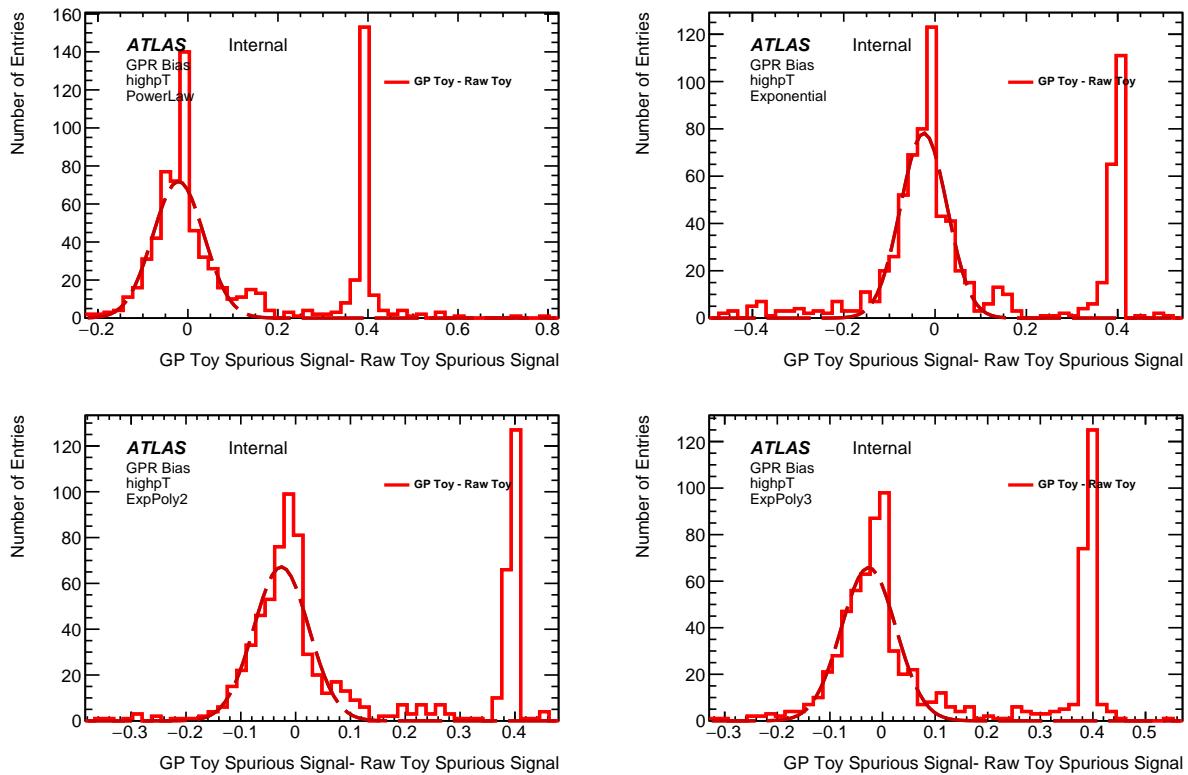


Figure A.74: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template. Each toy in this test has 10 events.

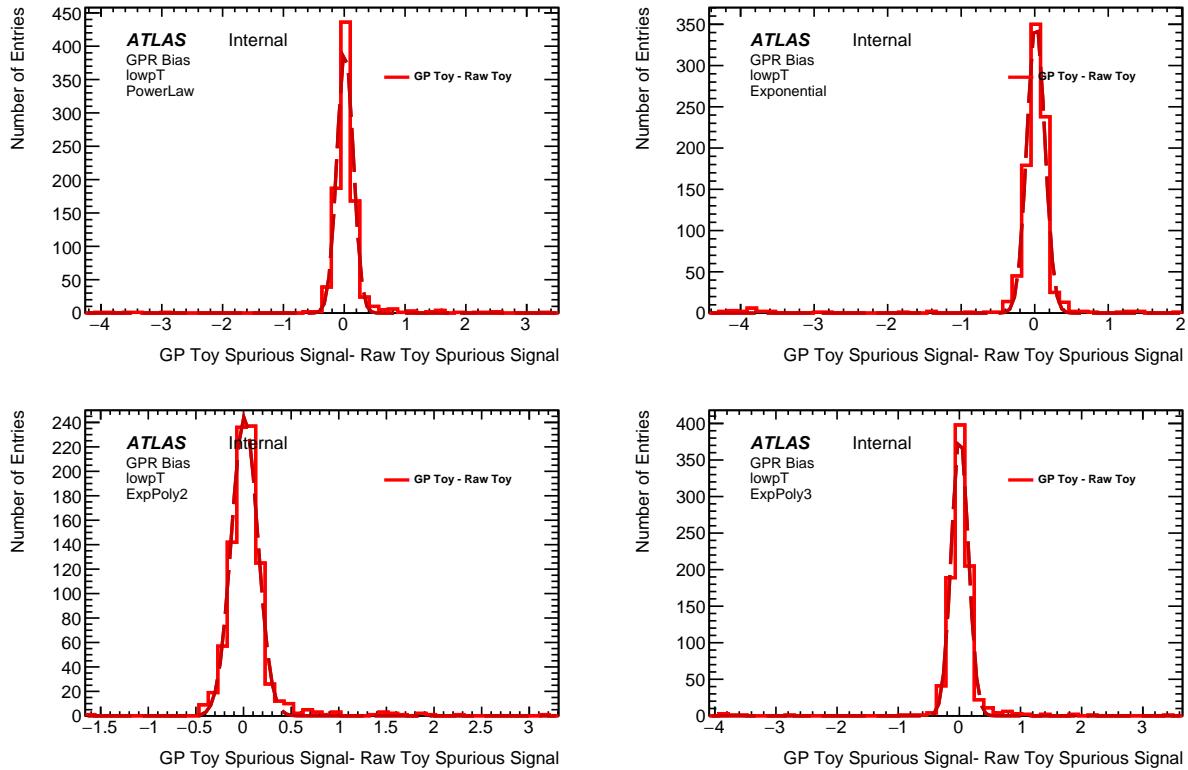


Figure A.75: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template. Each toy in this test has 100 events.

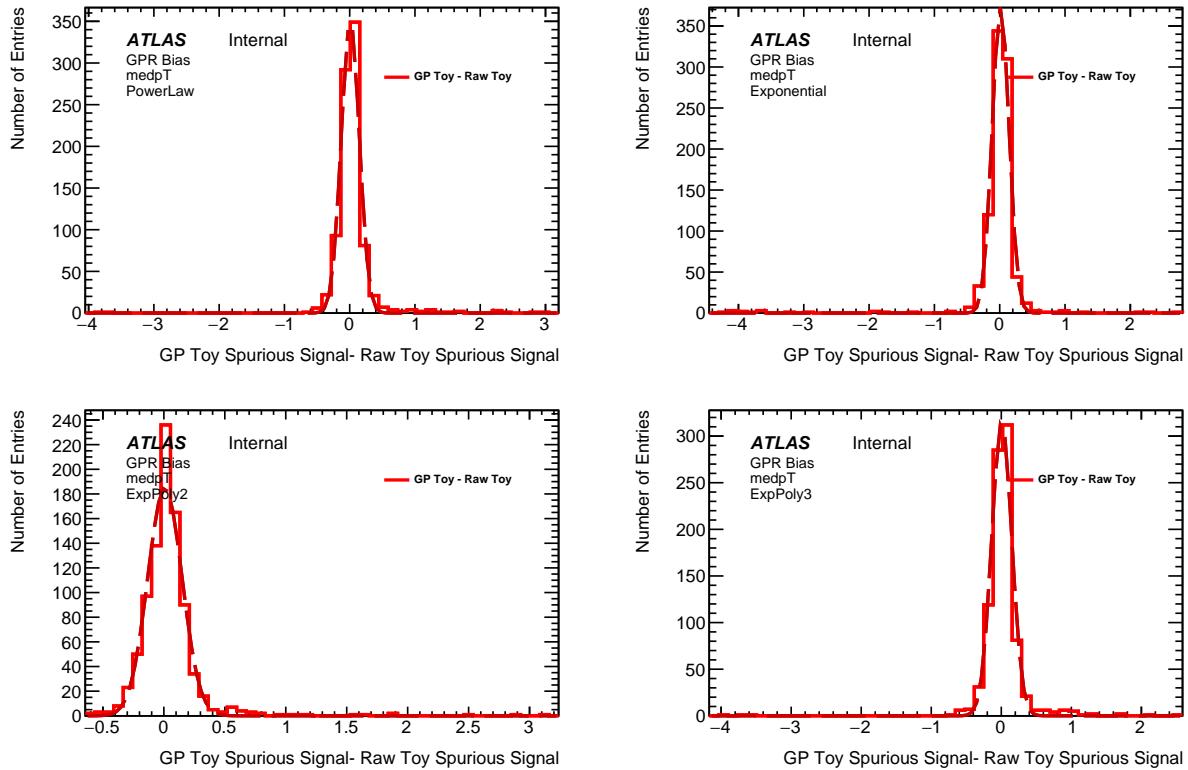


Figure A.76: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template. Each toy in this test has 100 events.

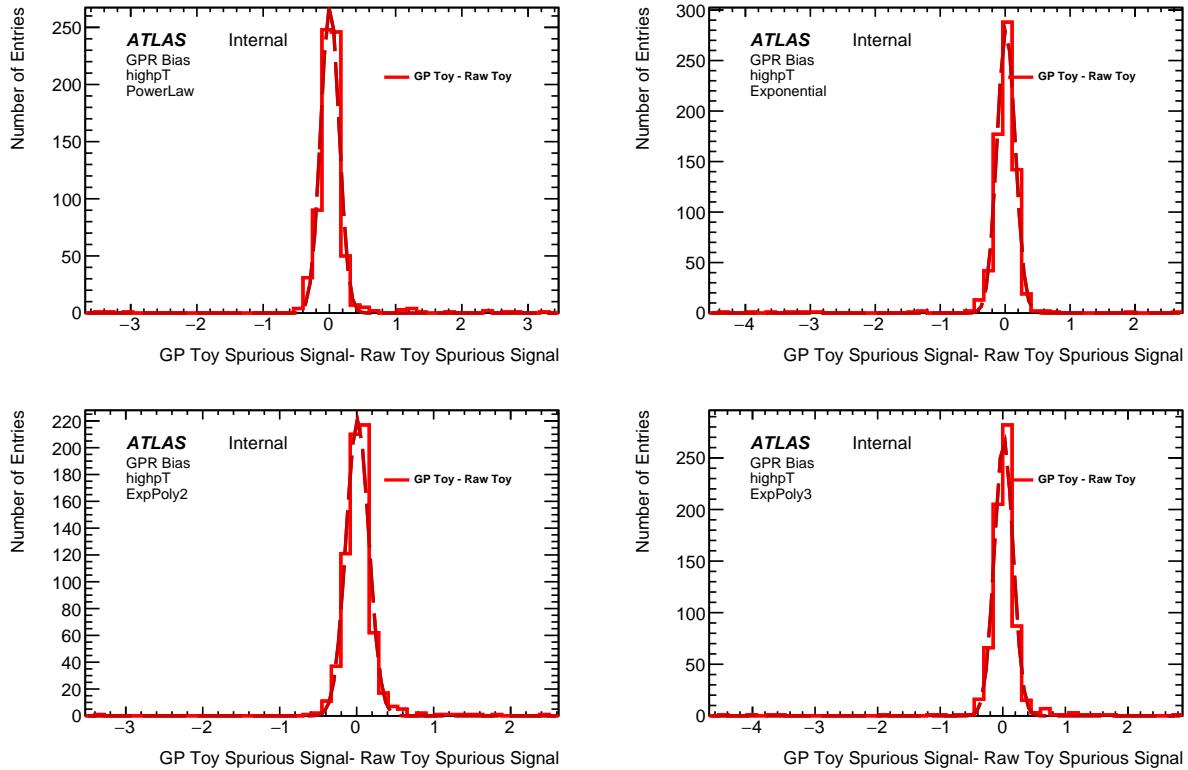


Figure A.77: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template. Each toy in this test has 100 events.

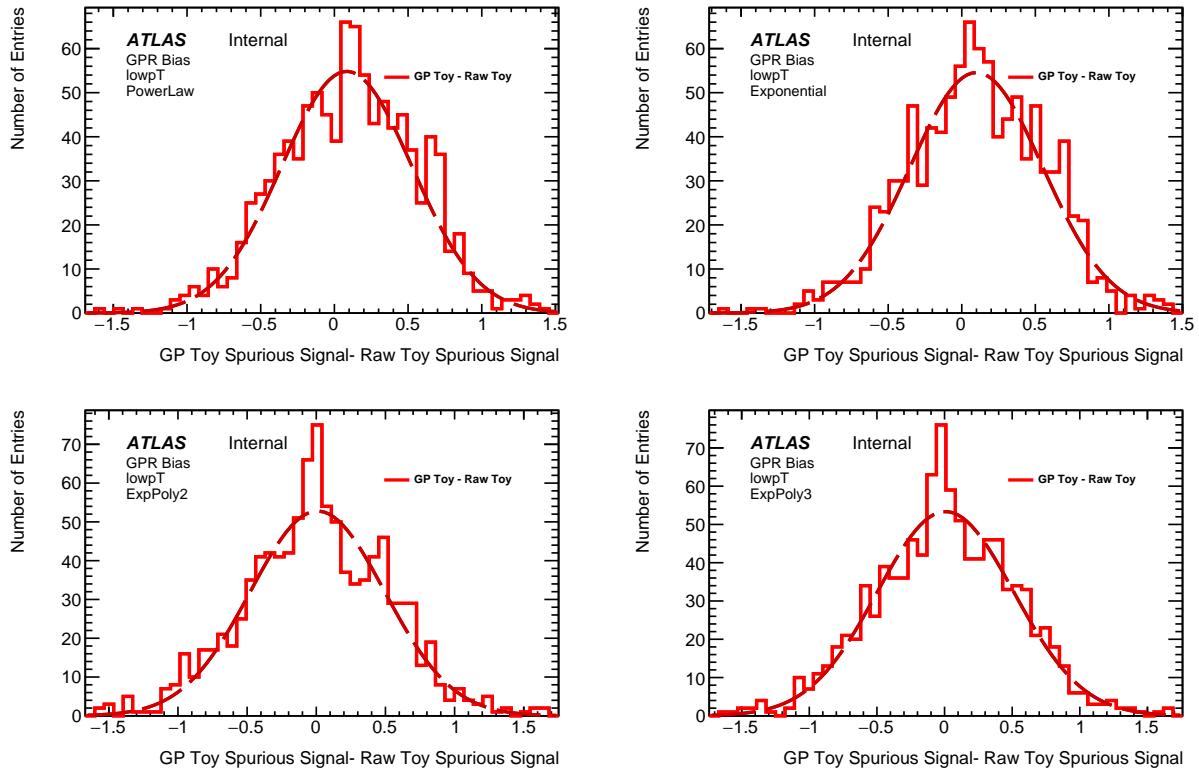


Figure A.78: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template. Each toy in this test has 1000 events.

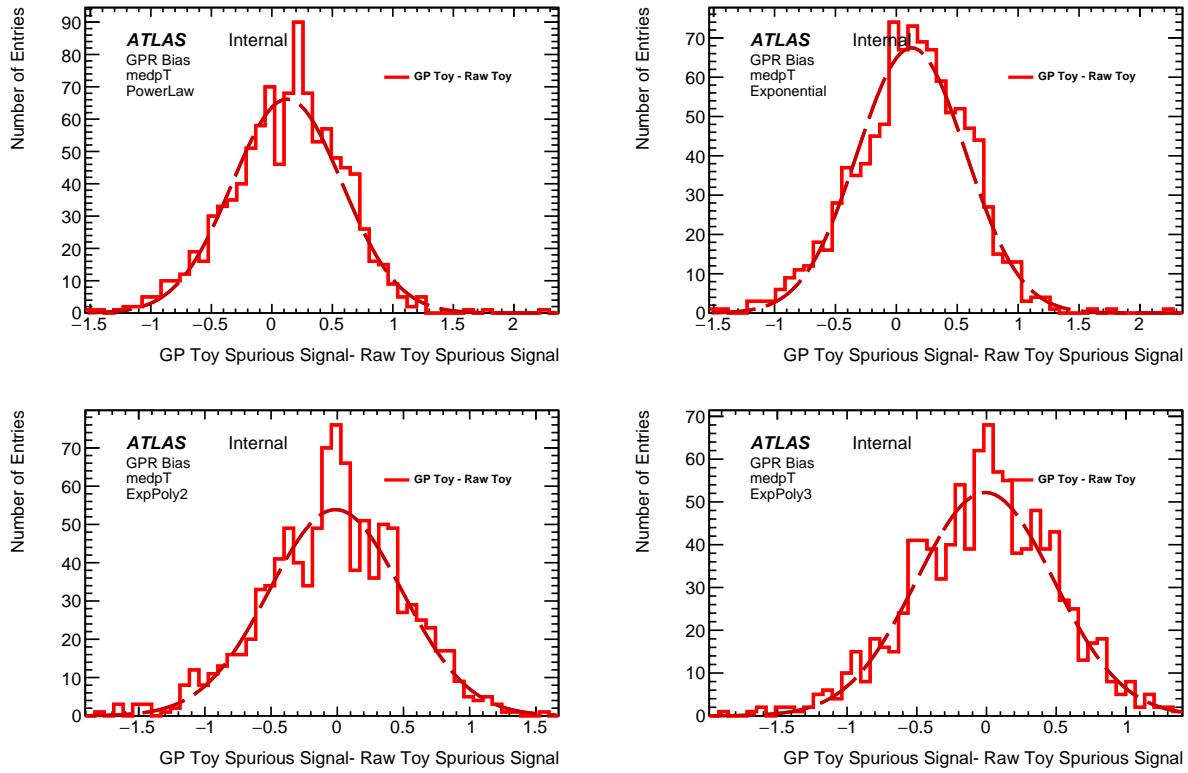


Figure A.79: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template. Each toy in this test has 1000 events.

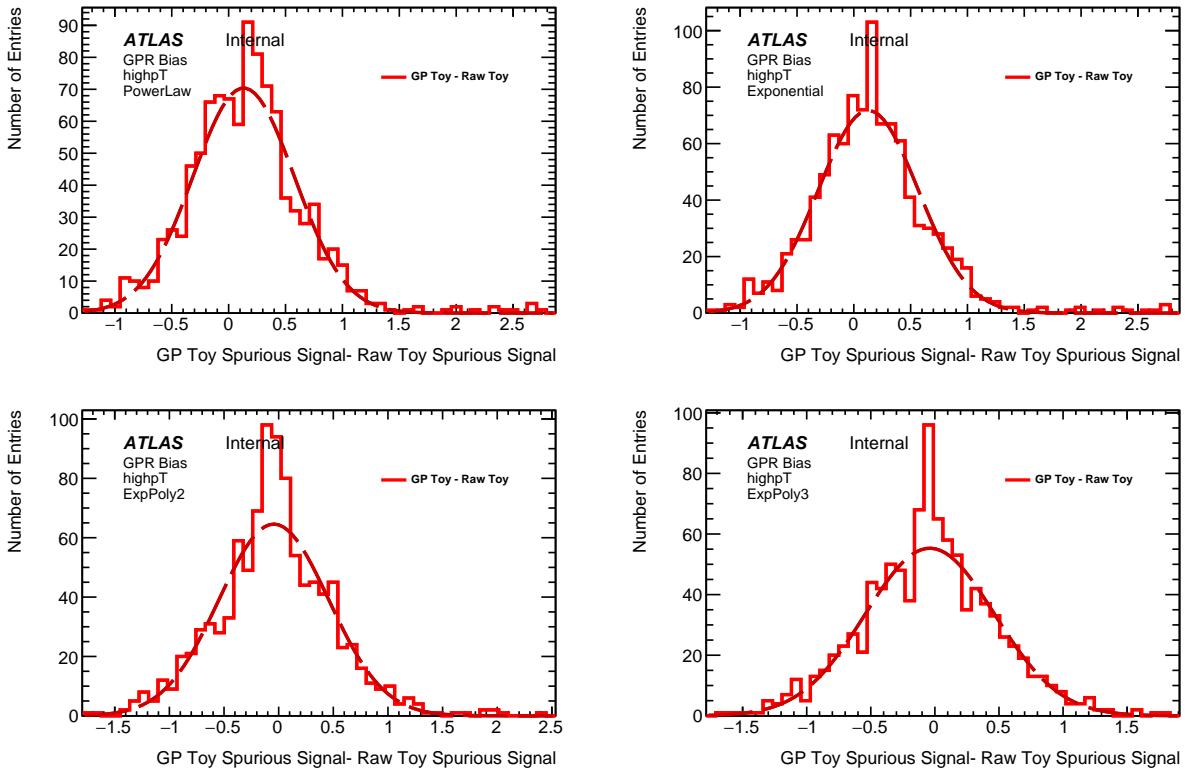


Figure A.80: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template. Each toy in this test has 1000 events.

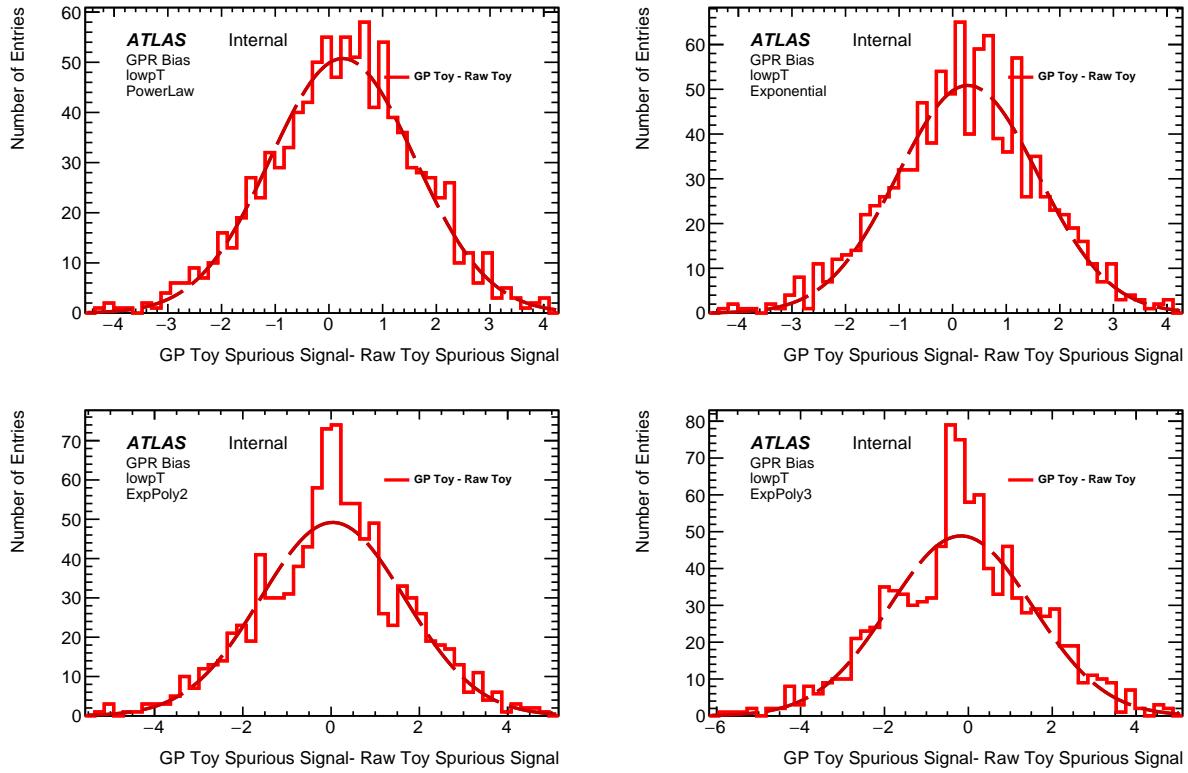


Figure A.81: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template. Each toy in this test has 10k events.

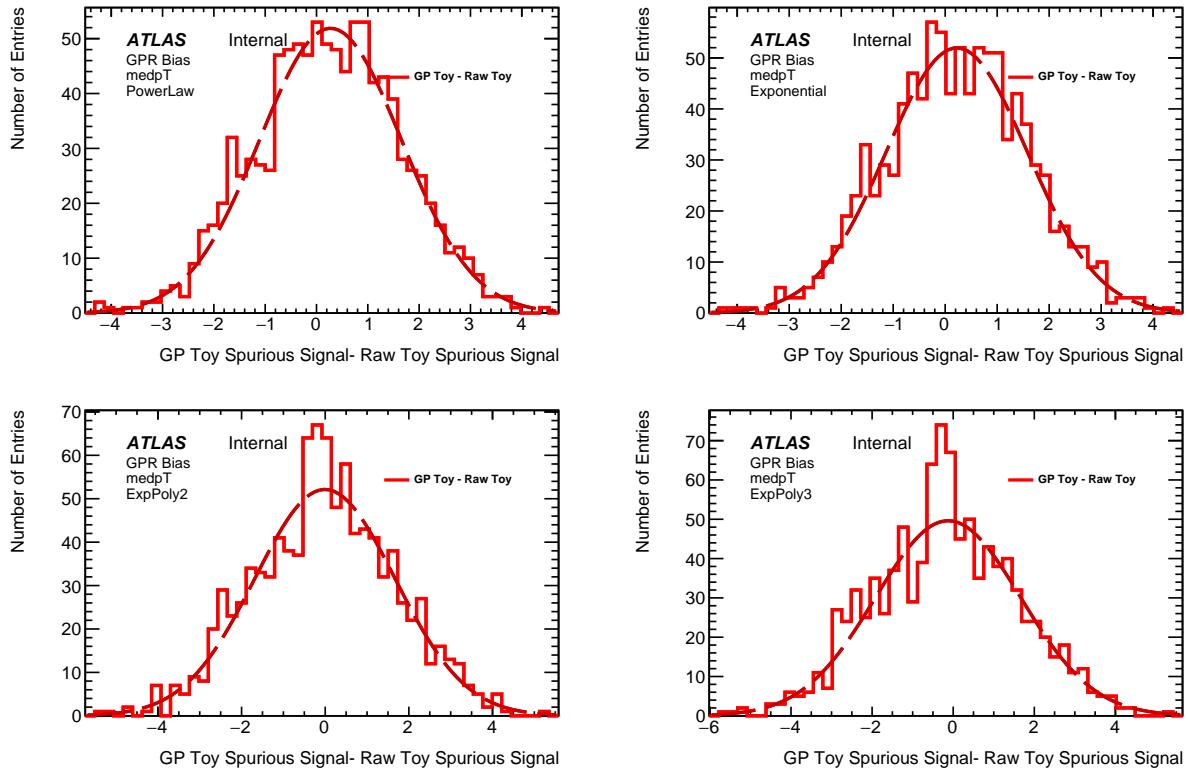


Figure A.82: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template. Each toy in this test has 10k events.

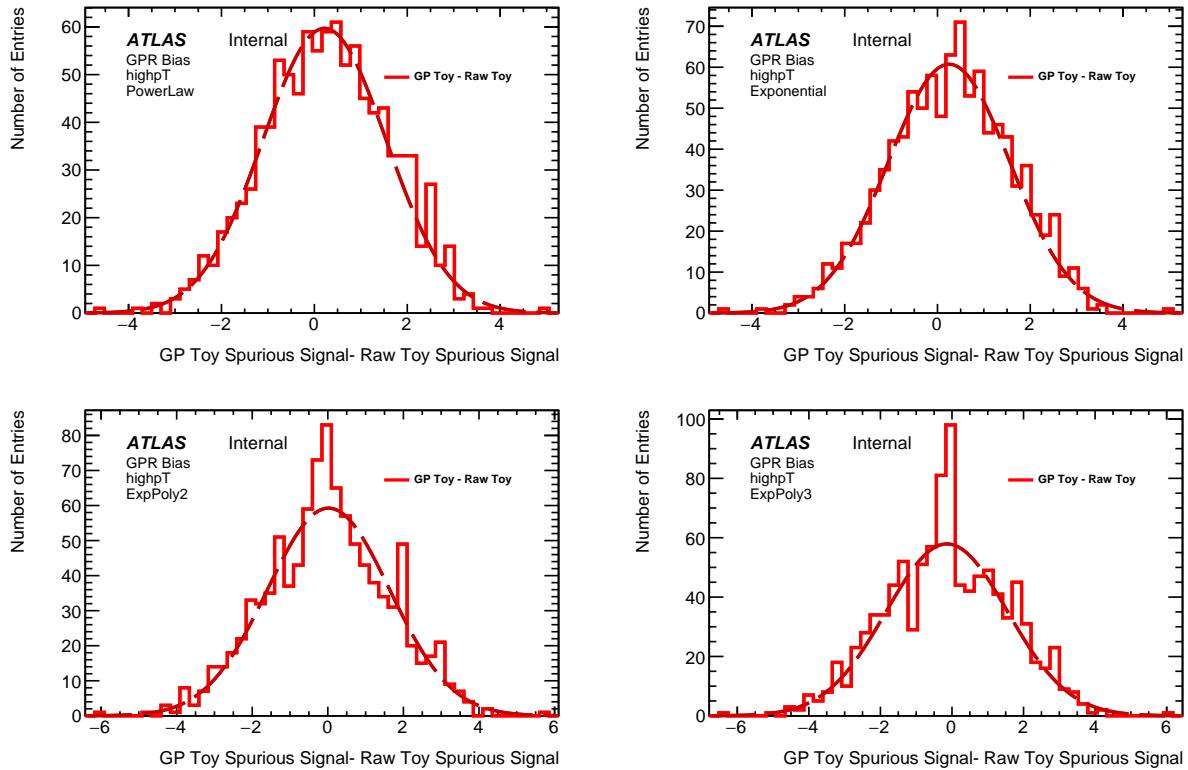


Figure A.83: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template. Each toy in this test has 10k events.

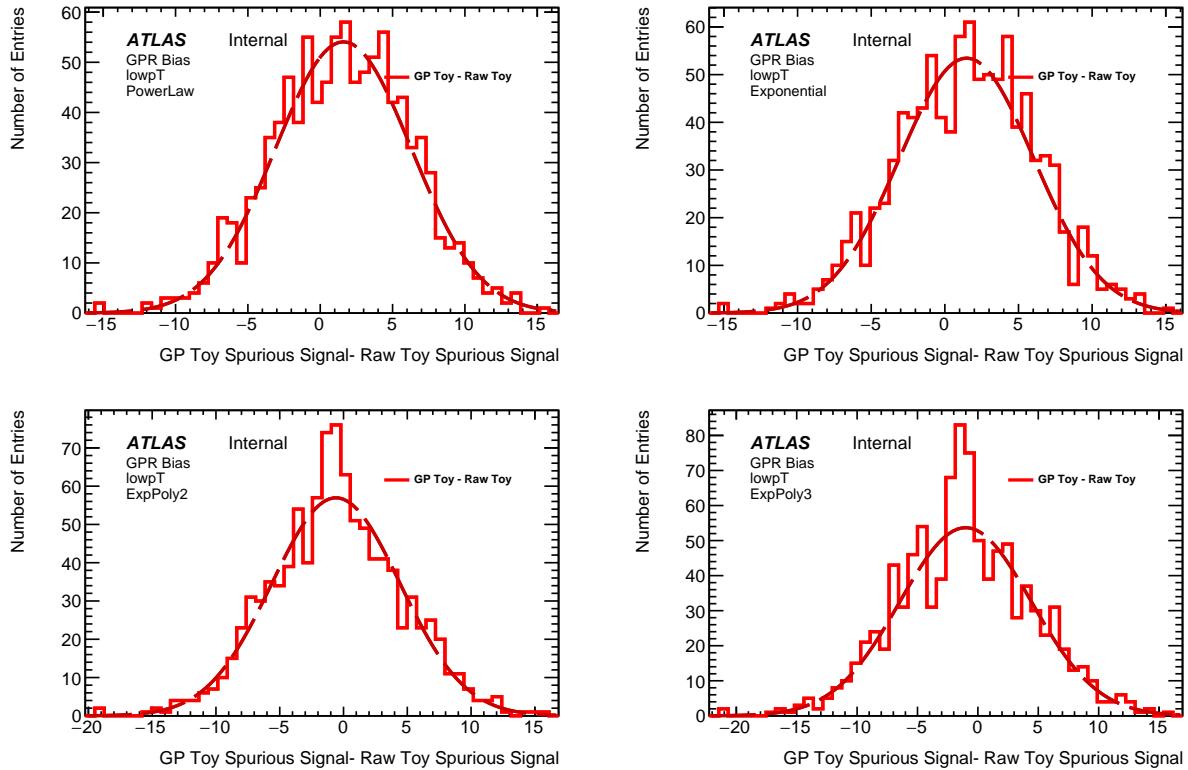


Figure A.84: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template. Each toy in this test has 100k events.

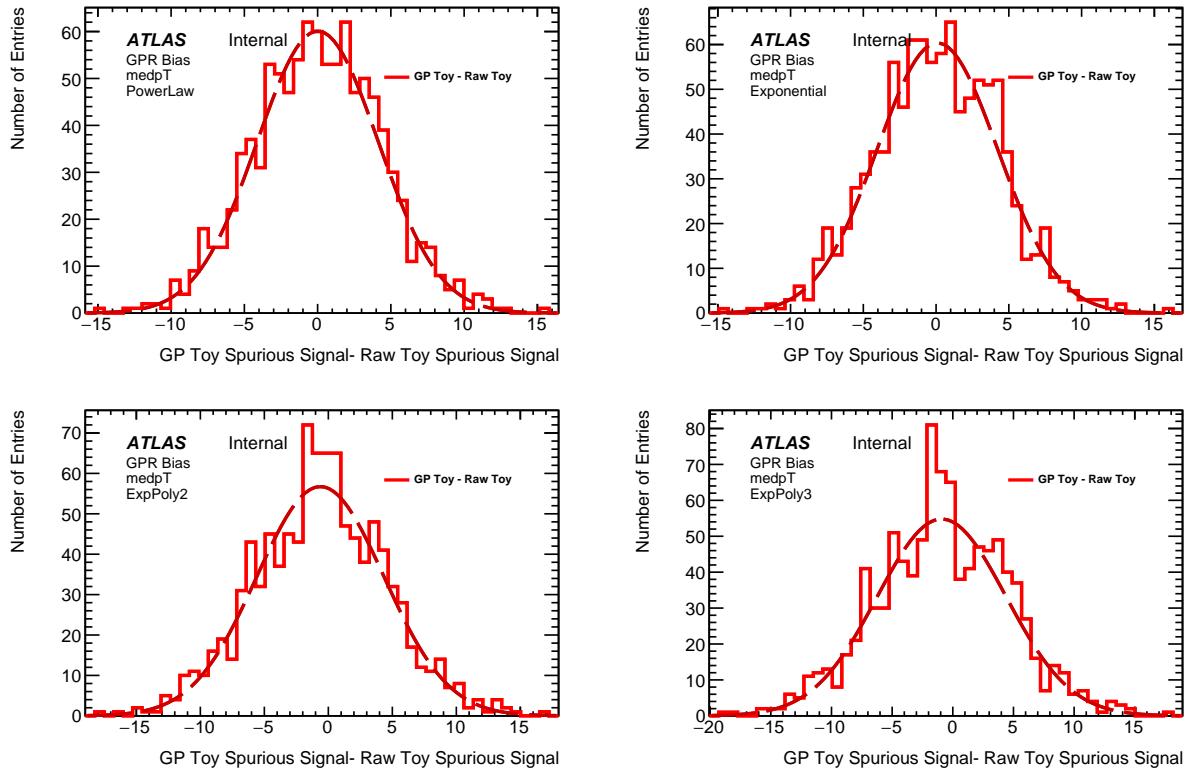


Figure A.85: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template. Each toy in this test has 100k events.

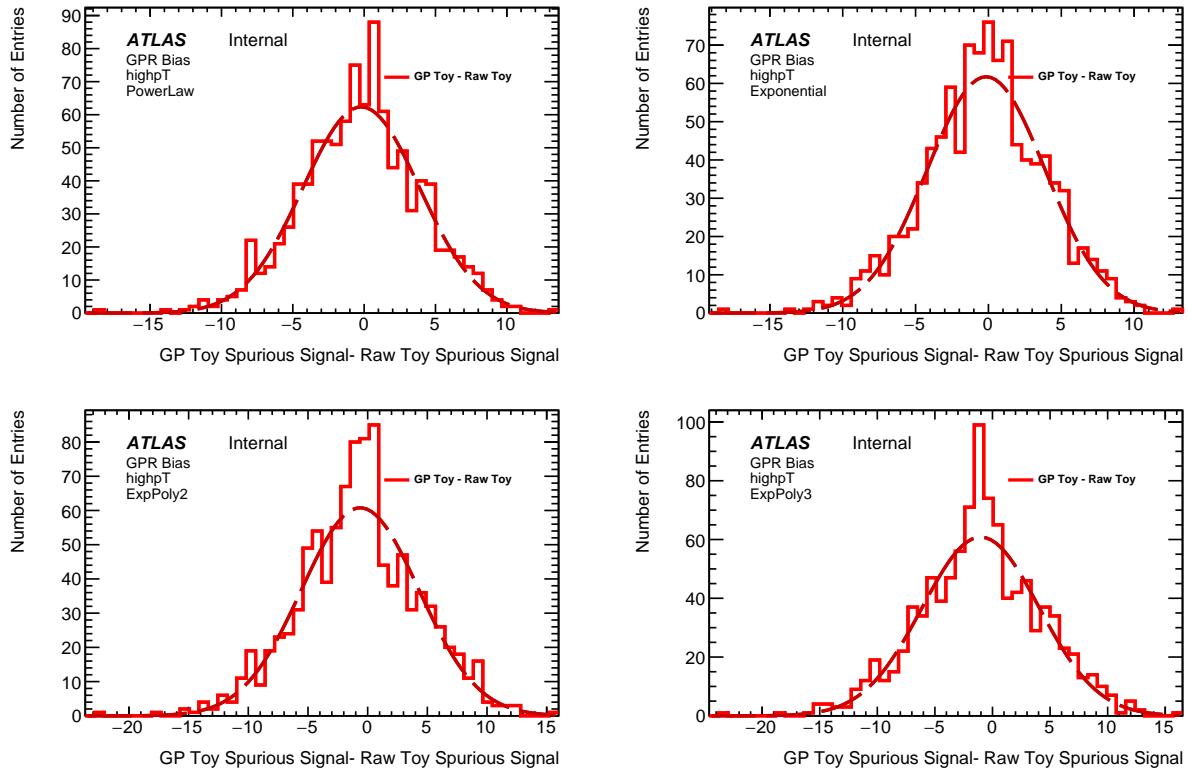


Figure A.86: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template. Each toy in this test has 100k events.

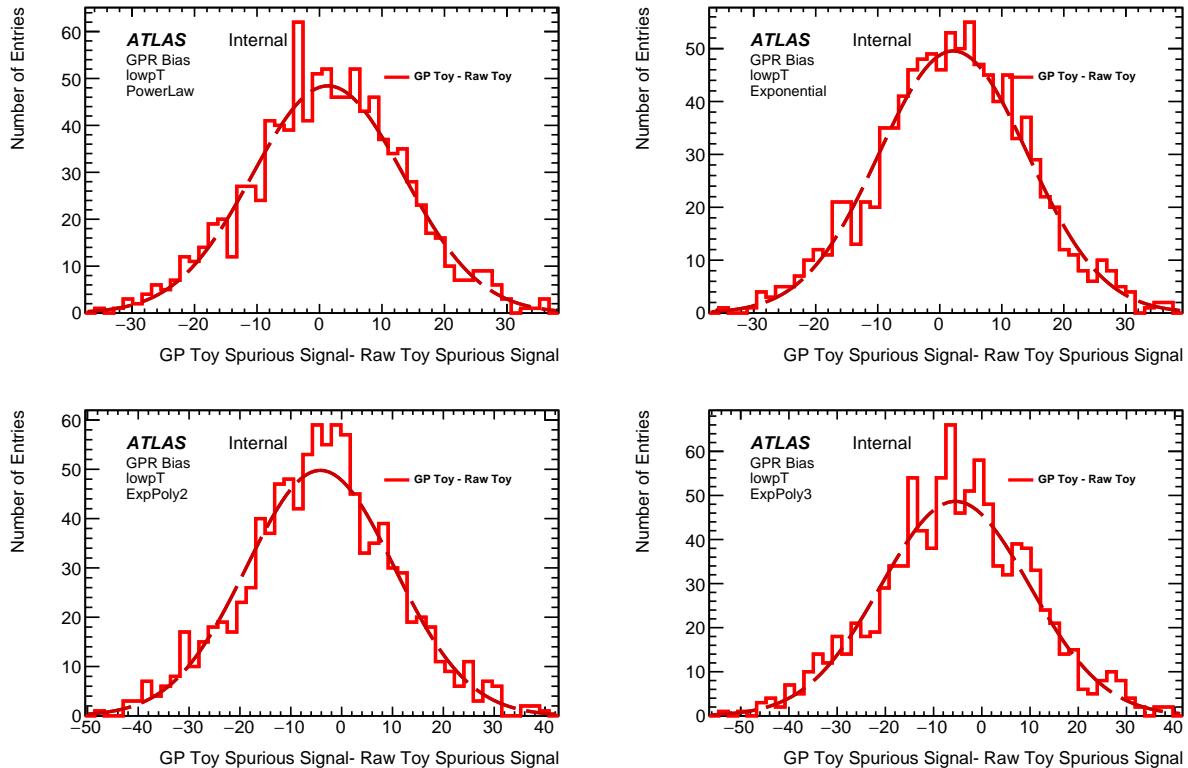


Figure A.87: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template. Each toy in this test has 1,000,000 events.

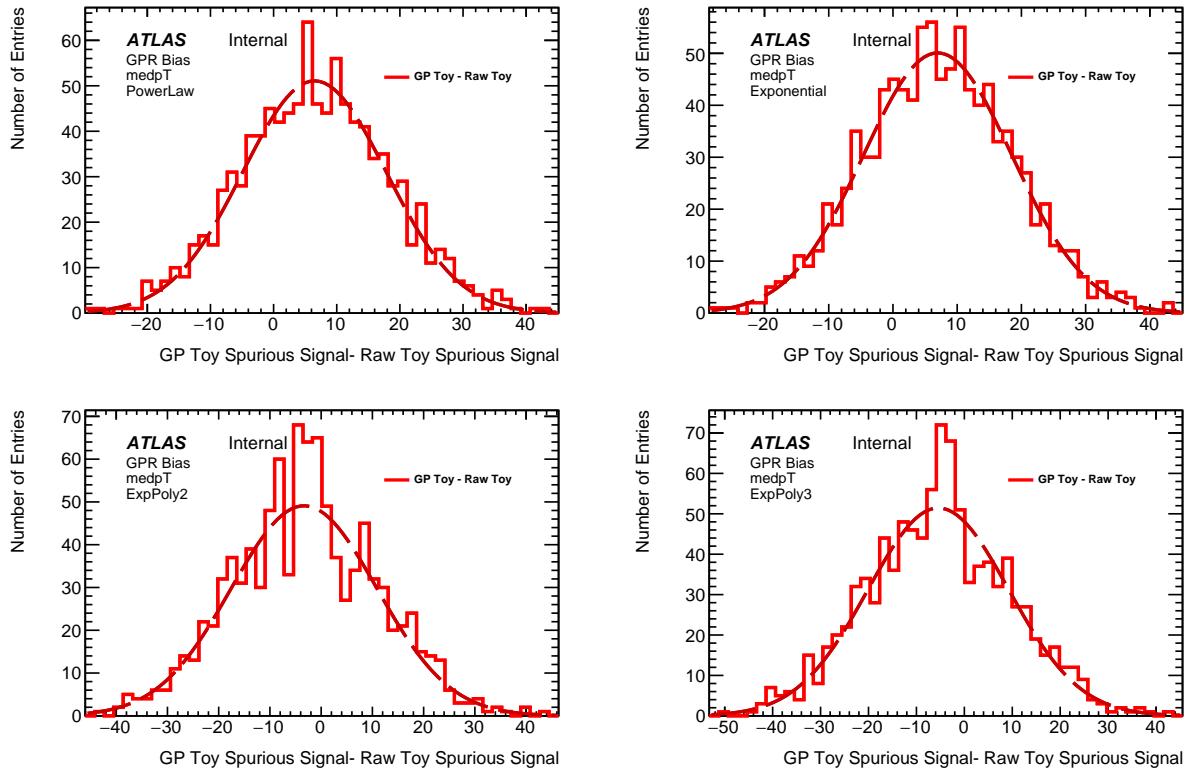


Figure A.88: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template. Each toy in this test has 1,000,000 events.

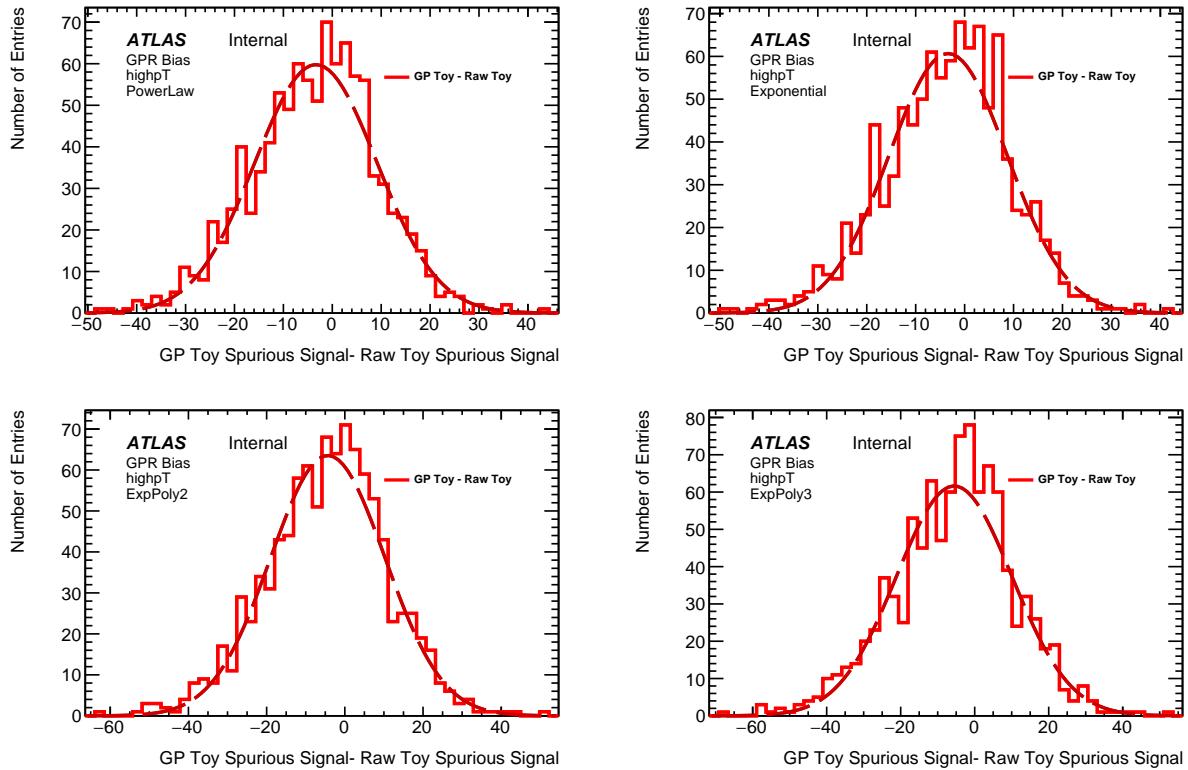


Figure A.89: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template. Each toy in this test has 1,000,000 events.

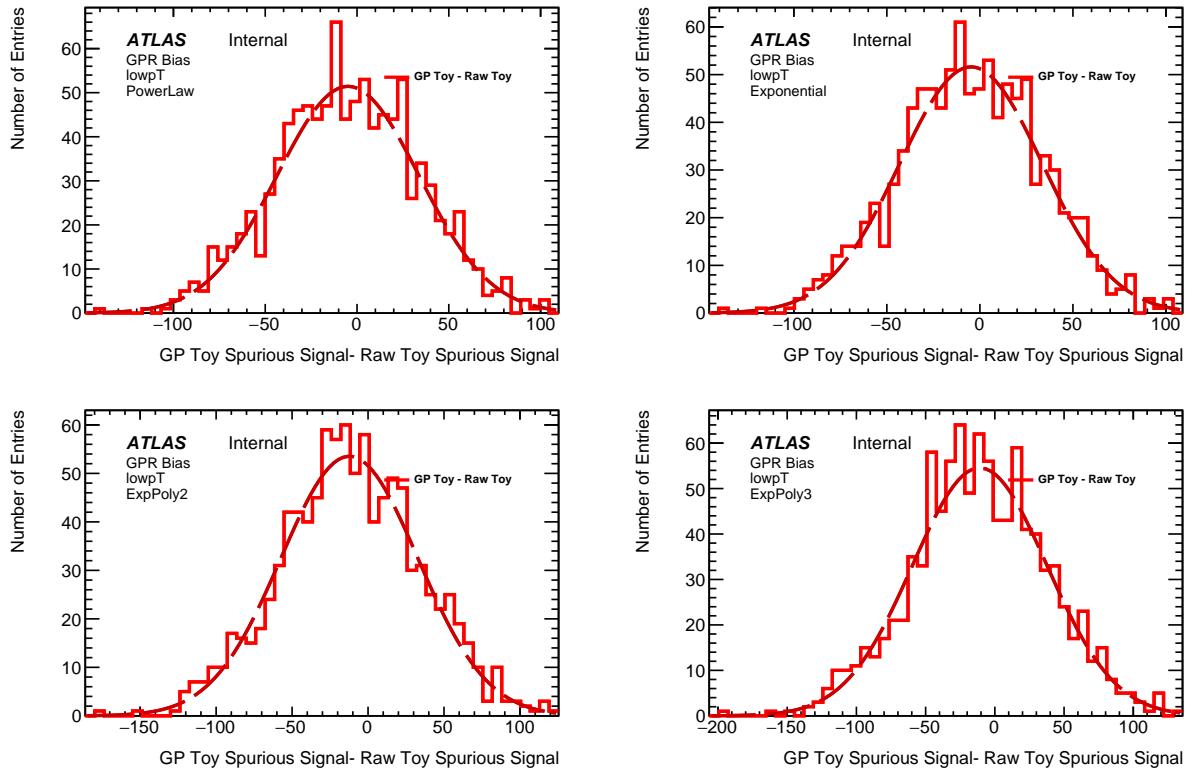


Figure A.90: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template. Each toy in this test has 10,000,000 events.

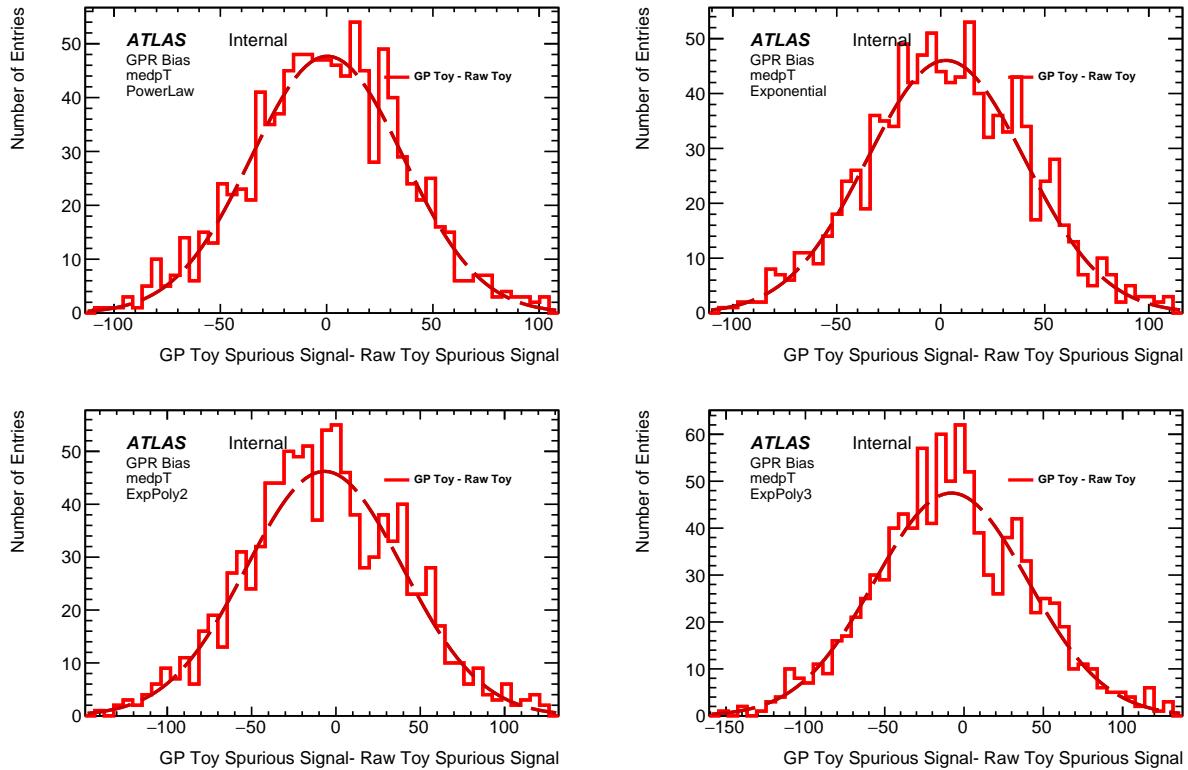


Figure A.91: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template. Each toy in this test has 10,000,000 events.

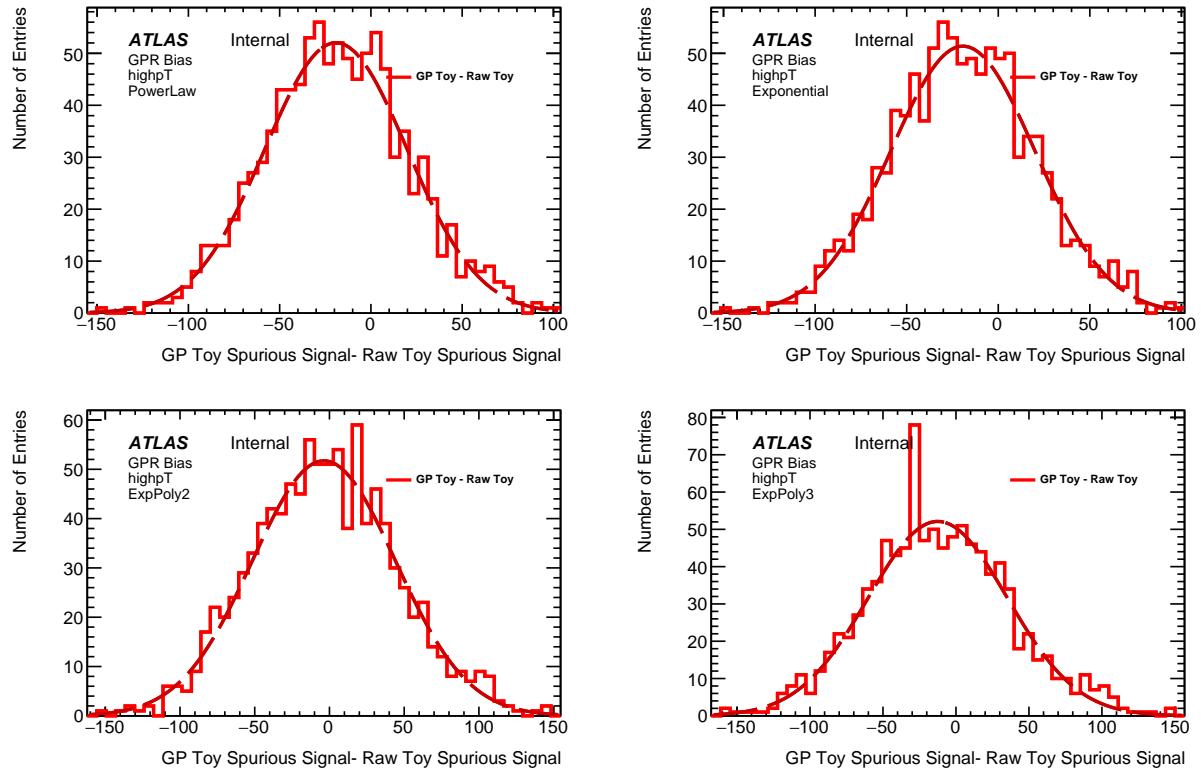


Figure A.92: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template. Each toy in this test has 10,000,000 events.

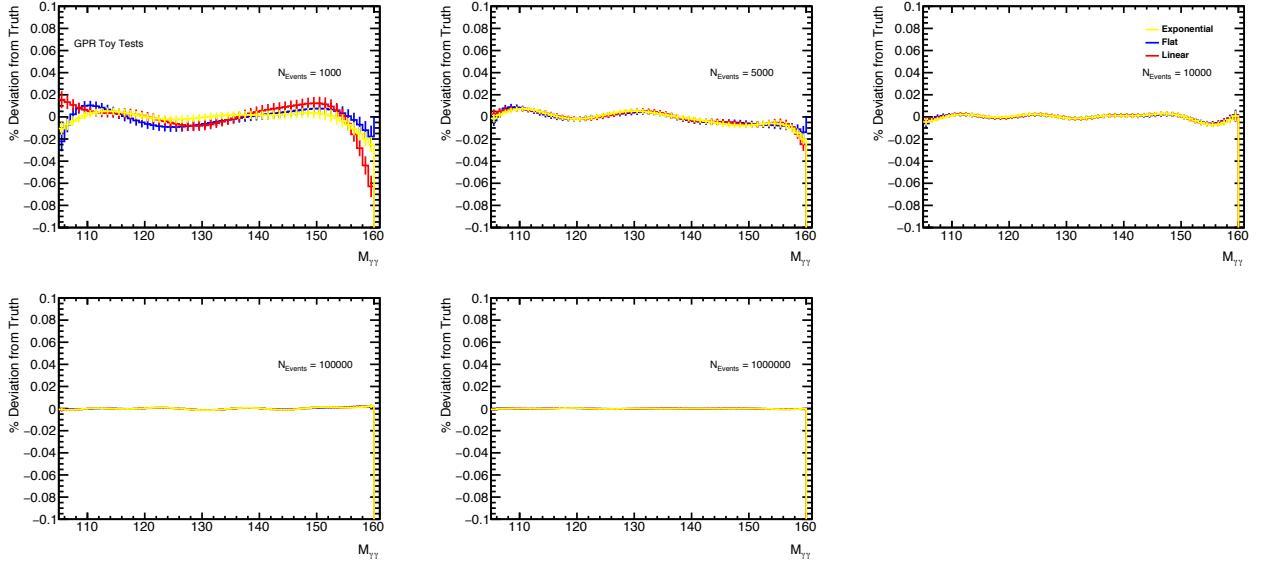


Figure A.93: Comparisons of the average bias induced by the choice of GP mean when fitting toy templates constructed from the analytic Power Law function. The yellow shape shows the results using the default Exponential mean, the blue shape shows the result using a flat line as the mean, and the red shape shows the result using a linear fit as the mean. The top left subplot shows the results using templates containing 1000 events, the top middle shows the results using templates containing 5000 events, and the top right shows the results using templates containing 10,000 events. The bottom left subplot shows the results using templates containing 100,000 events, while the bottom middle shows the results using templates containing one million events.

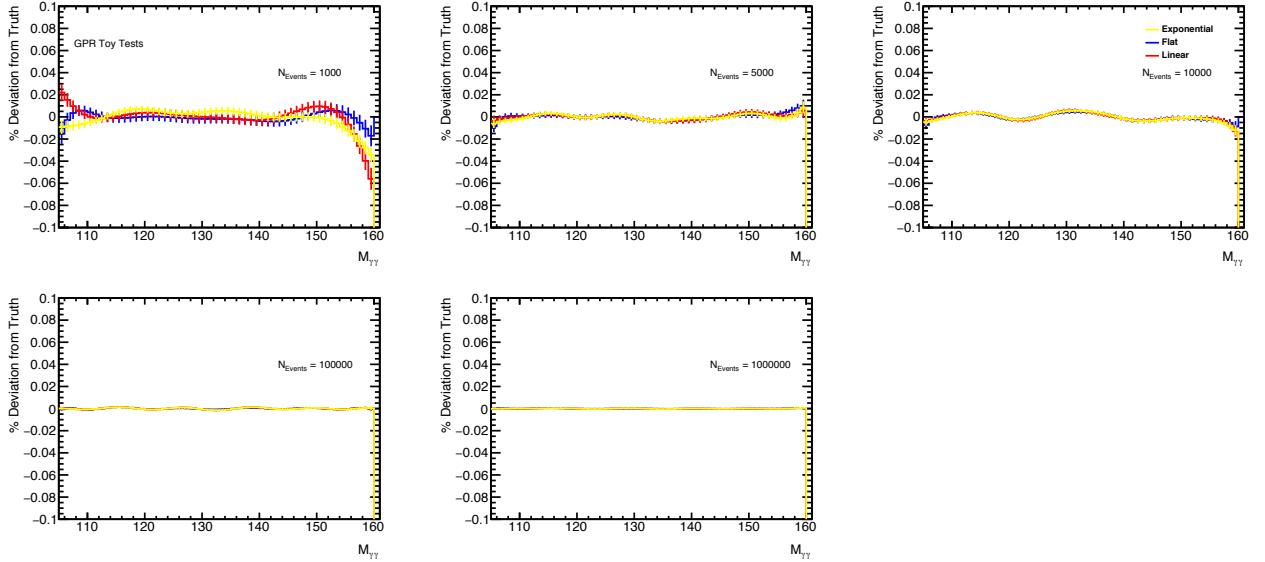


Figure A.94: Comparisons of the average bias induced by the choice of GP mean when fitting toy templates constructed from the analytic ExpPoly2 function. The yellow shape shows the results using the default Exponential mean, the blue shape shows the result using a flat line as the mean, and the red shape shows the result using a linear fit as the mean. The top left subplot shows the results using templates containing 1000 events, the top middle shows the results using templates containing 5000 events, and the top right shows the results using templates containing 10,000 events. The bottom left subplot shows the results using templates containing 100,000 events, while the bottom middle shows the results using templates containing one million events.

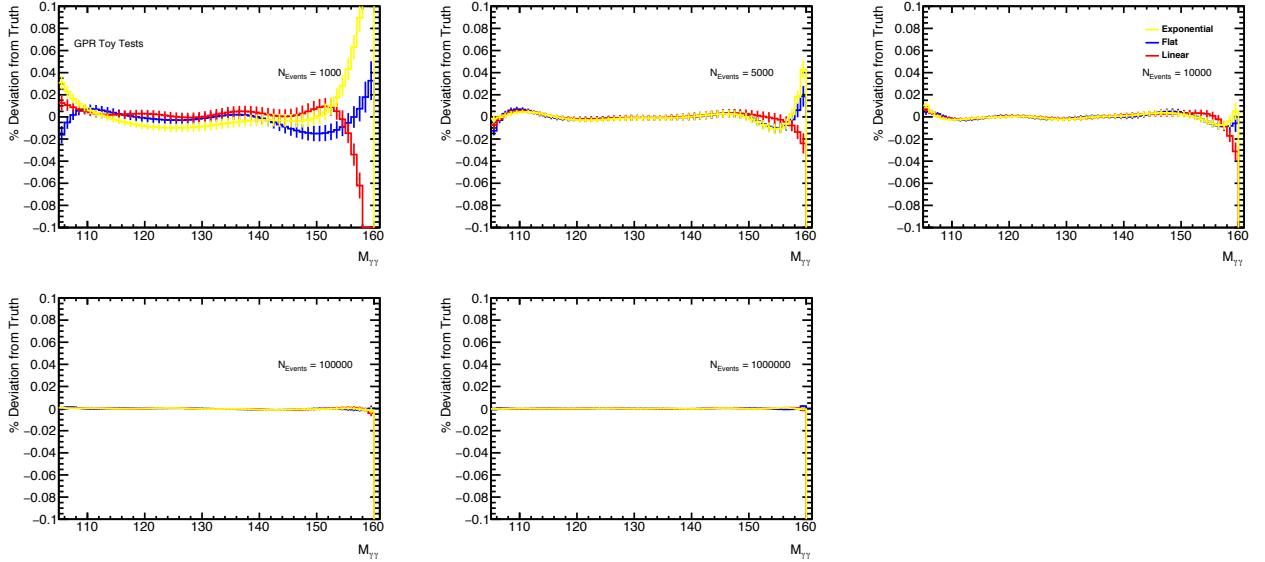


Figure A.95: Comparisons of the average bias induced by the choice of GP mean when fitting toy templates constructed from the analytic Bernstein 5 function. The yellow shape shows the results using the default Exponential mean, the blue shape shows the result using a flat line as the mean, and the red shape shows the result using a linear fit as the mean. The top left subplot shows the results using templates containing 1000 events, the top middle shows the results using templates containing 5000 events, and the top right shows the results using templates containing 10,000 events. The bottom left subplot shows the results using templates containing 100,000 events, while the bottom middle shows the results using templates containing one million events.

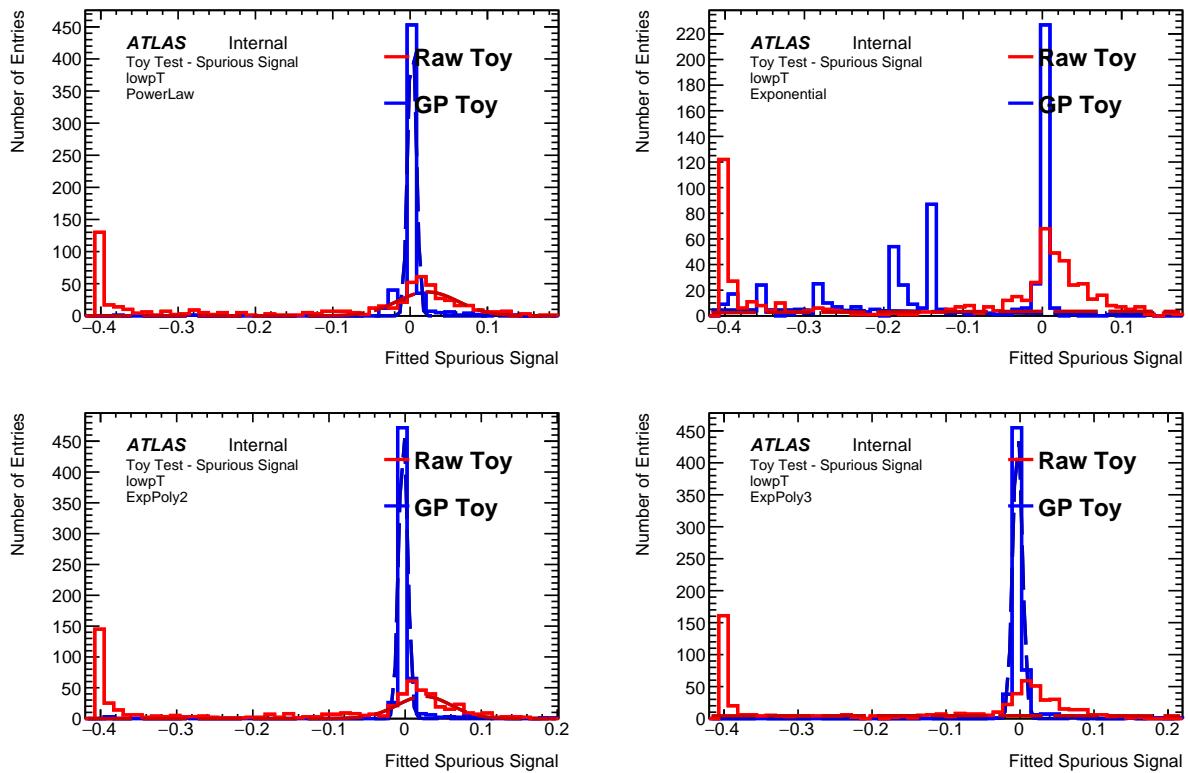


Figure A.96: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 bins on either side. Each toy in this test has 10 events.

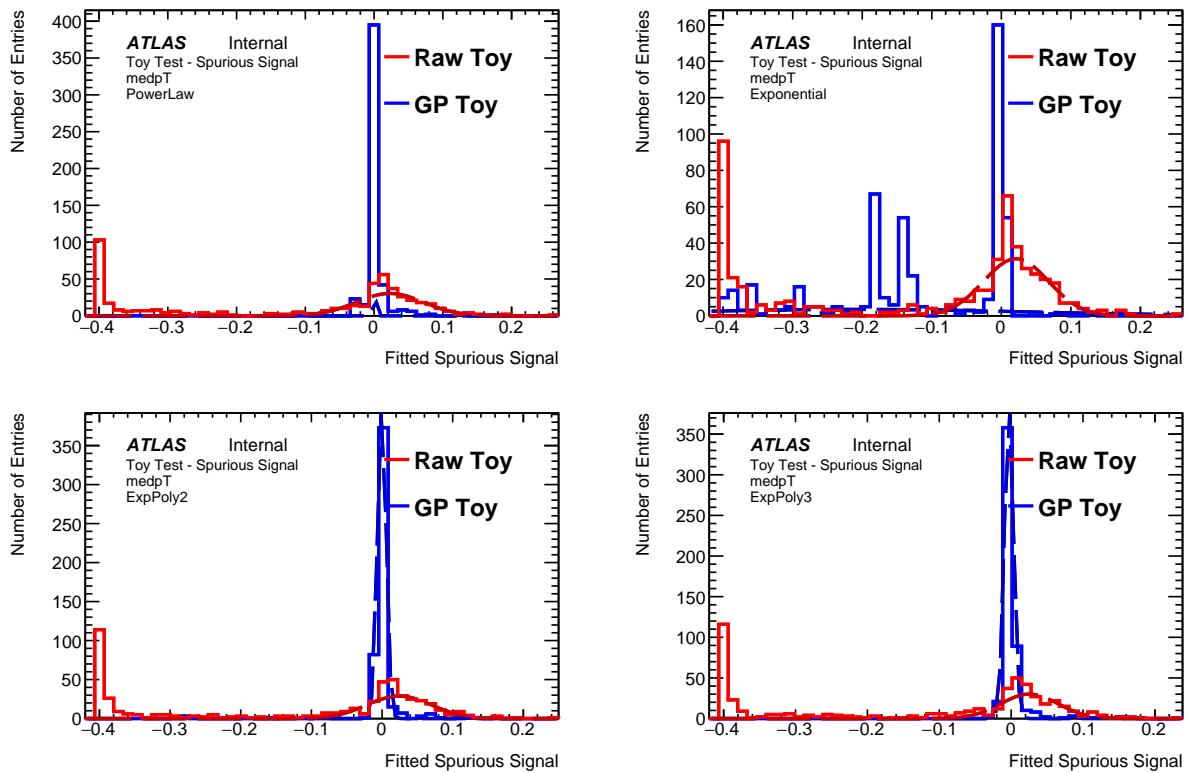


Figure A.97: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 10 events.

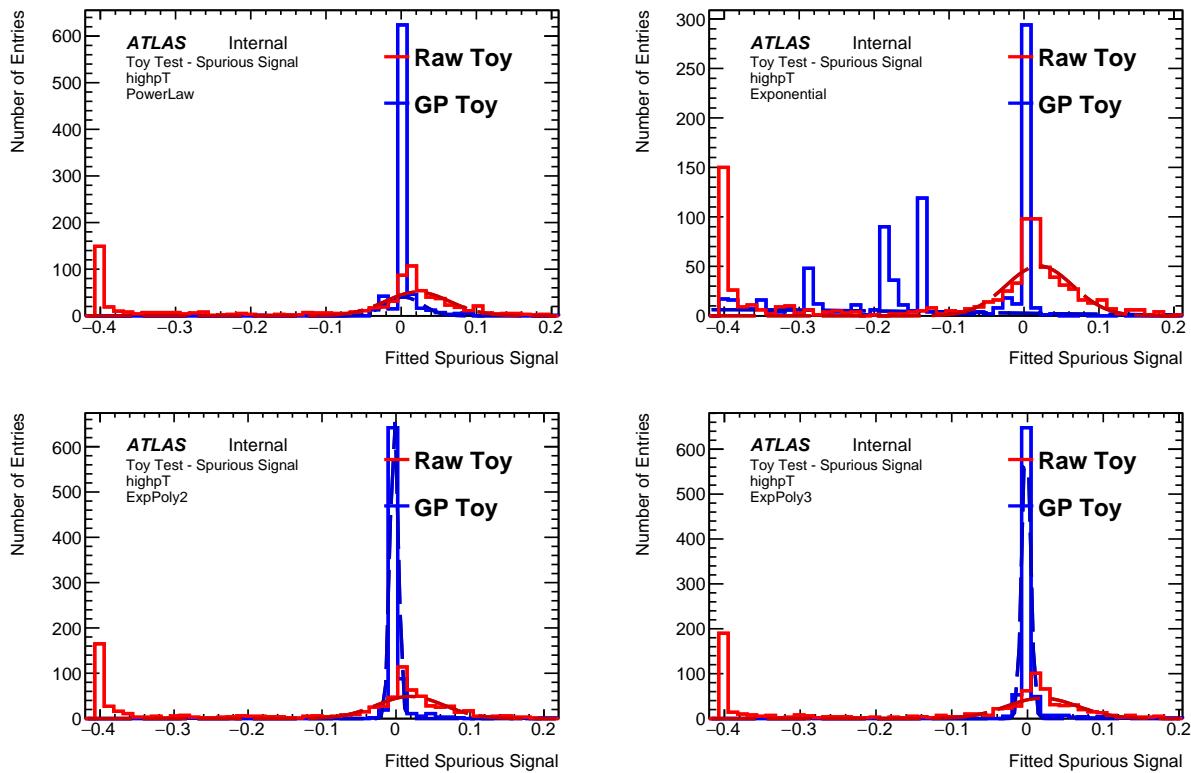


Figure A.98: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 10 events.

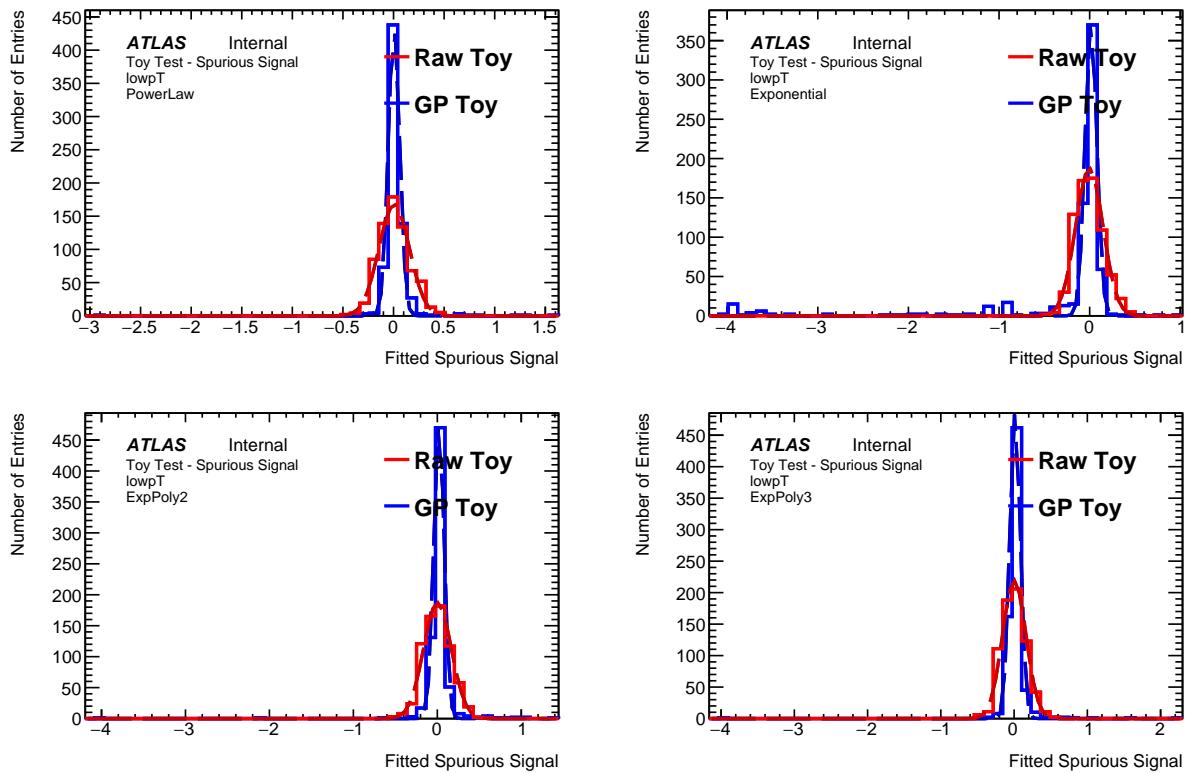


Figure A.99: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side. Each toy in this test has 100 events.

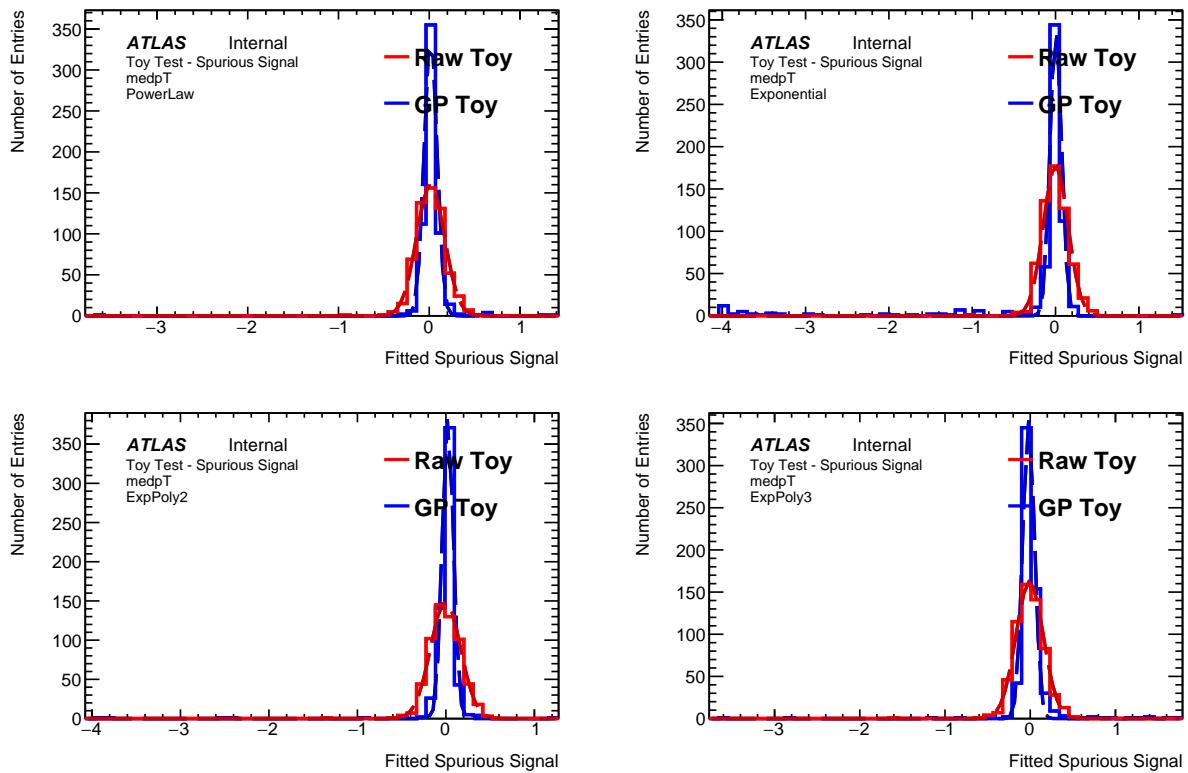


Figure A.100: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 100 events.

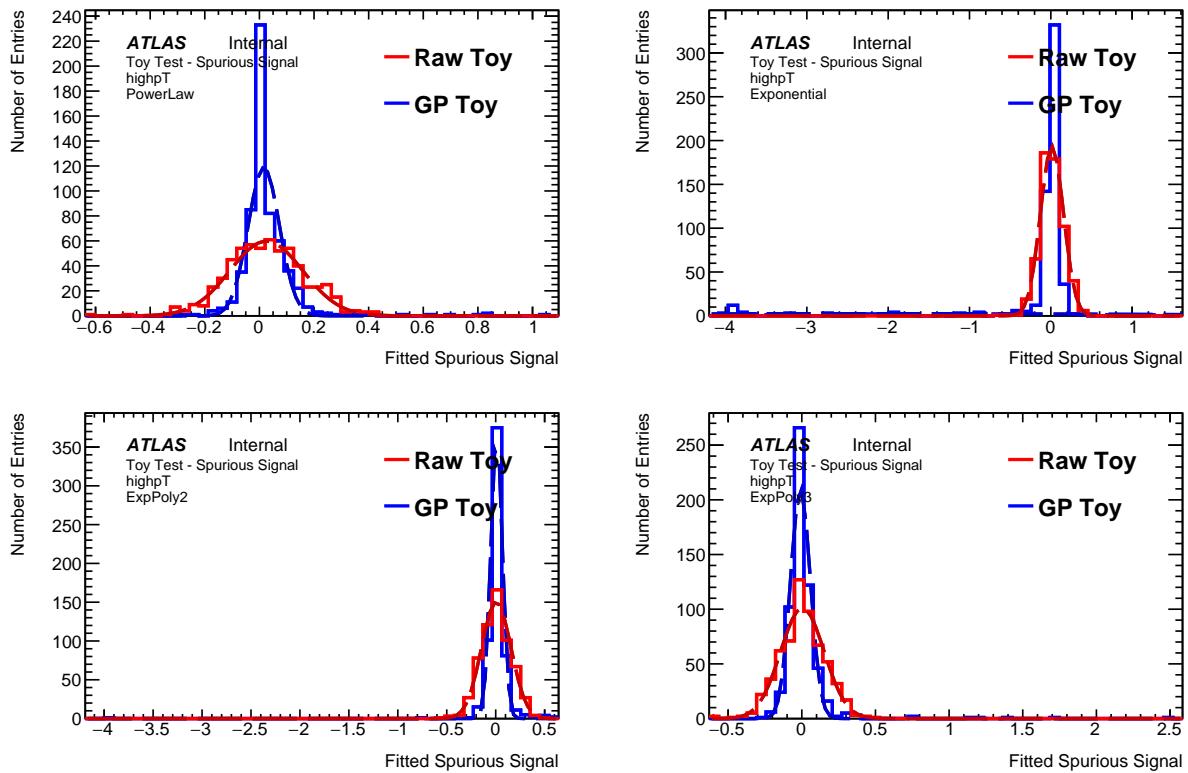


Figure A.101: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 100 events.

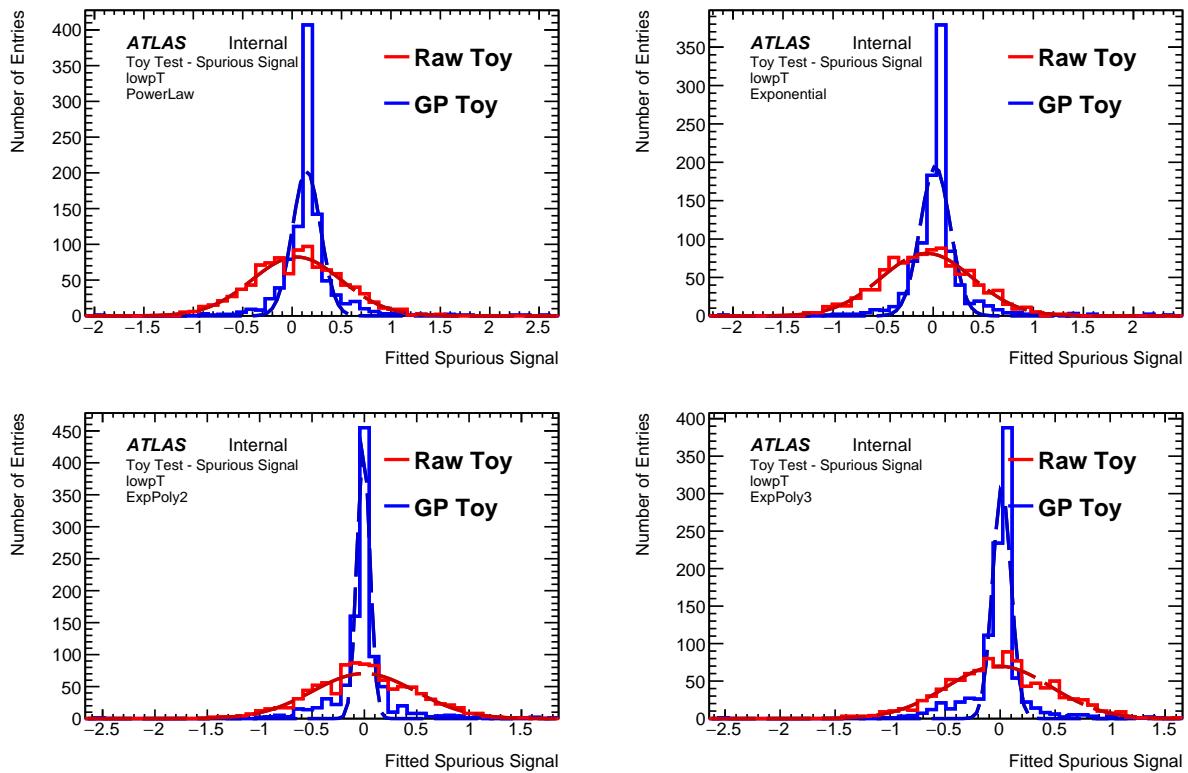


Figure A.102: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side. Each toy in this test has 1000 events.

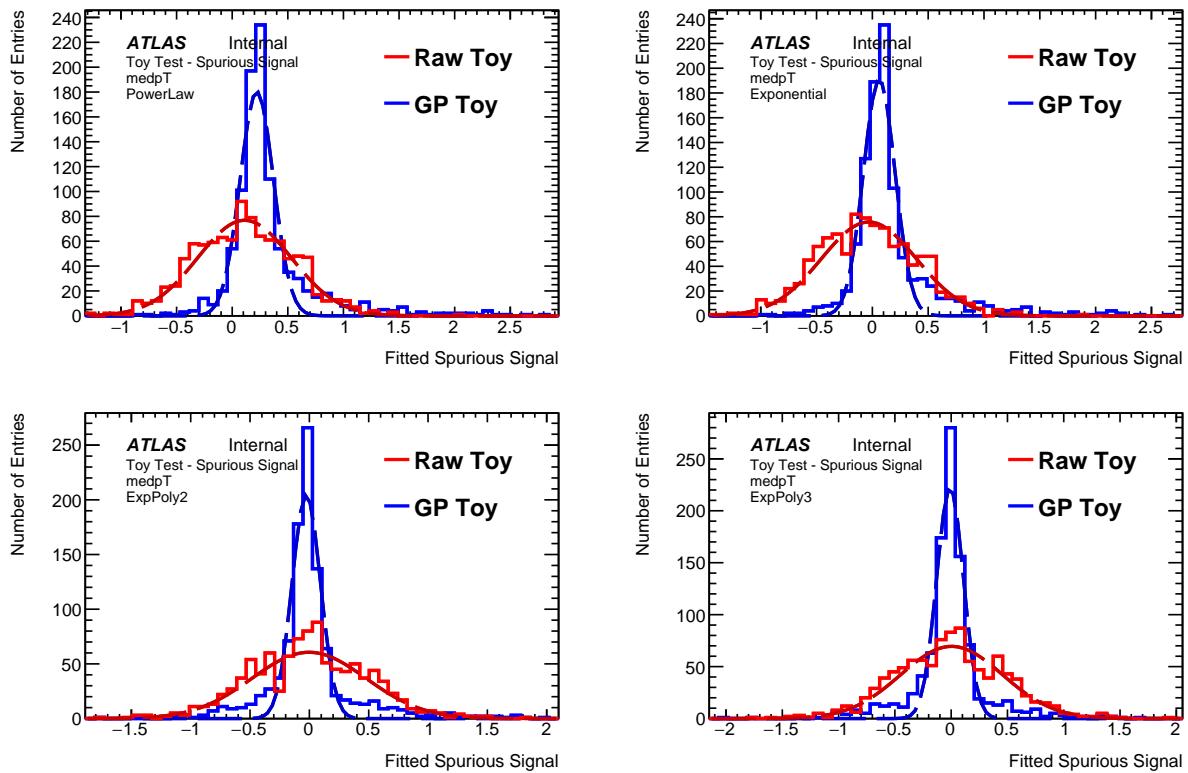


Figure A.103: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 1000 events.

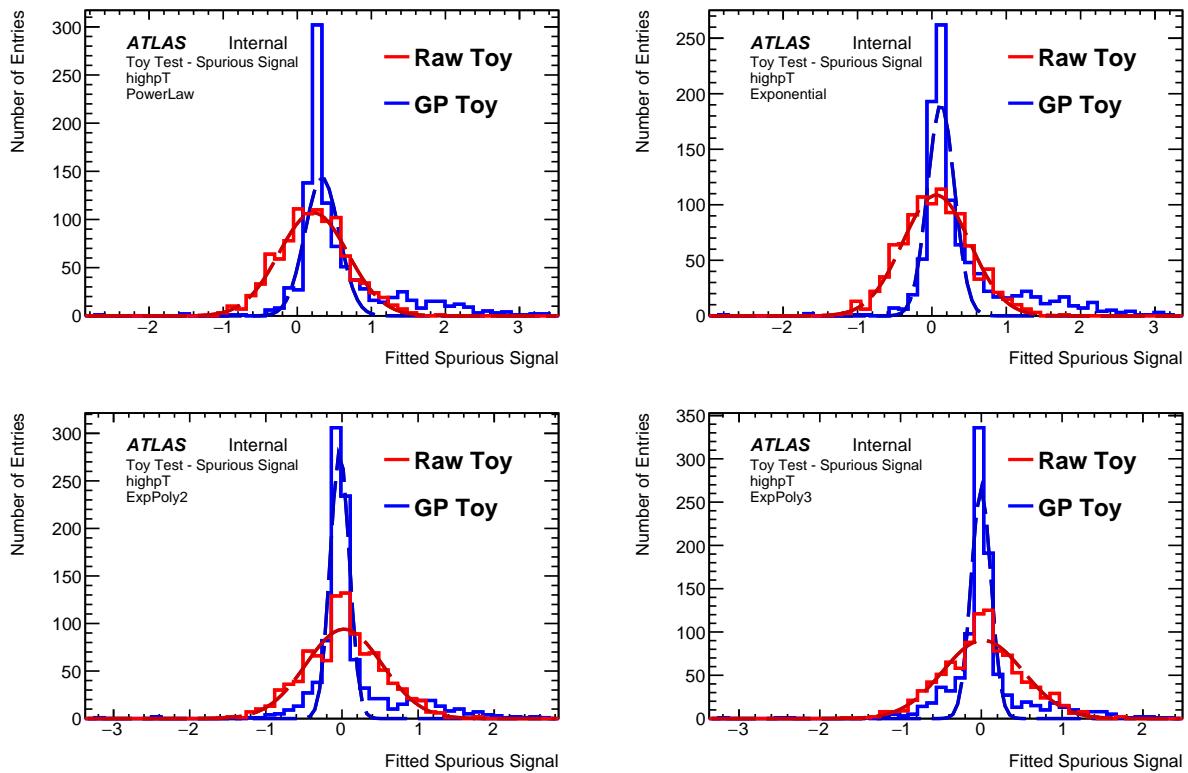


Figure A.104: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 1000 events.

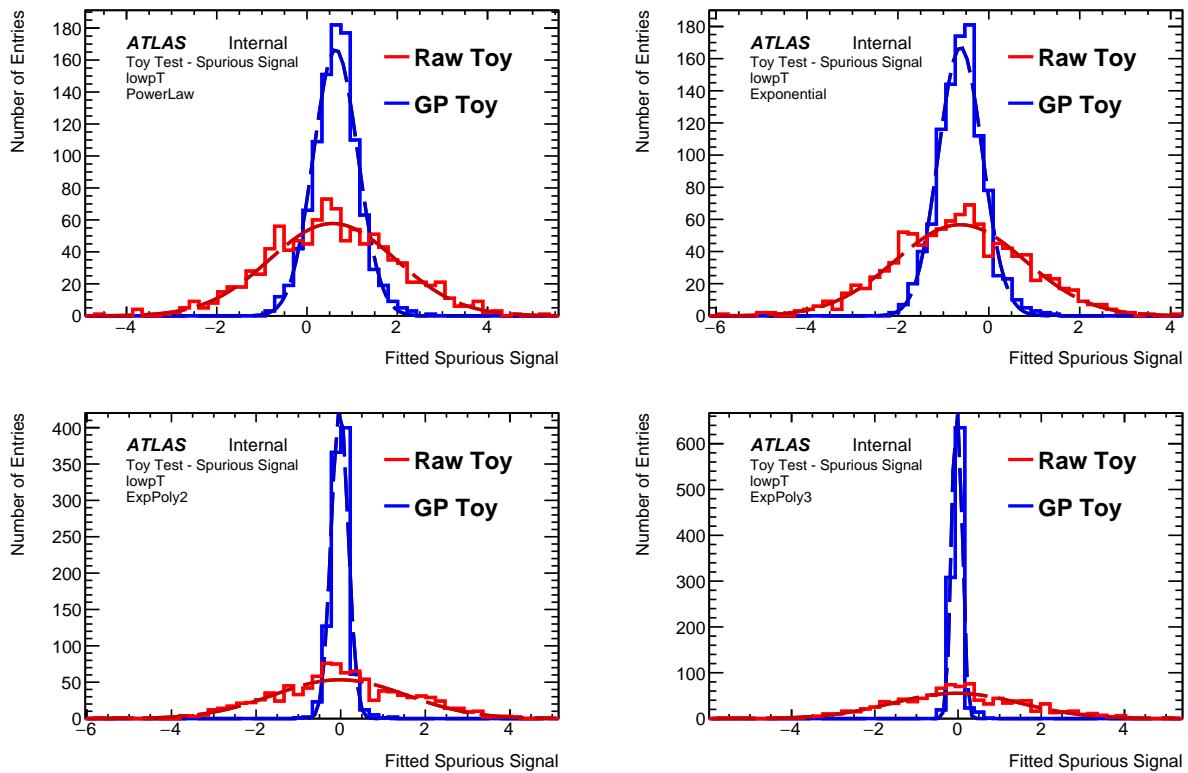


Figure A.105: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side. Each toy in this test has 10k events.

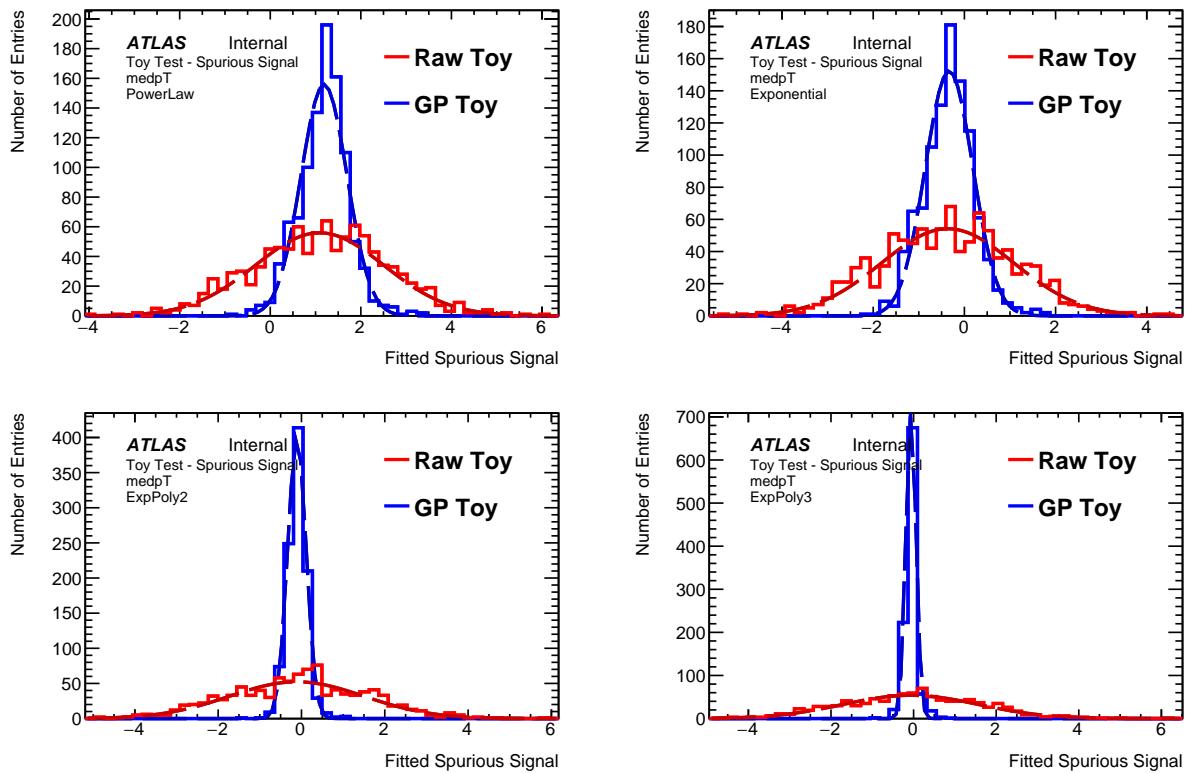


Figure A.106: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 10k events.

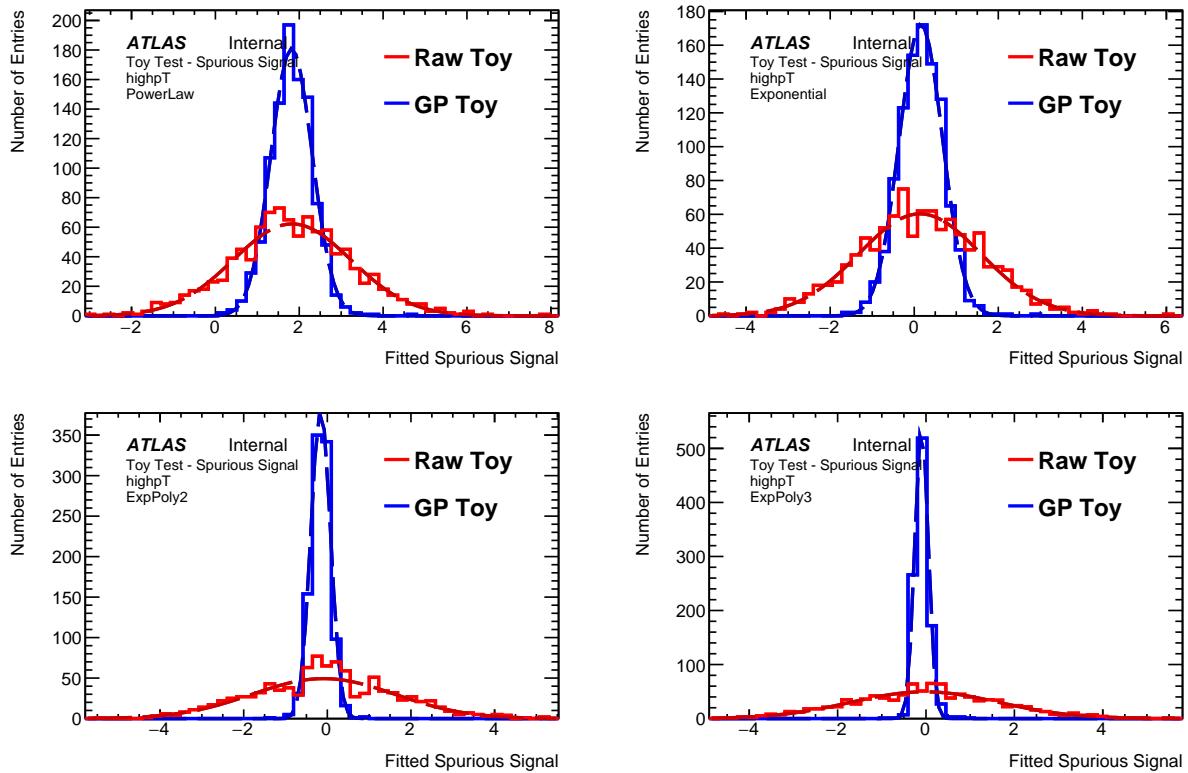


Figure A.107: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 10k events.

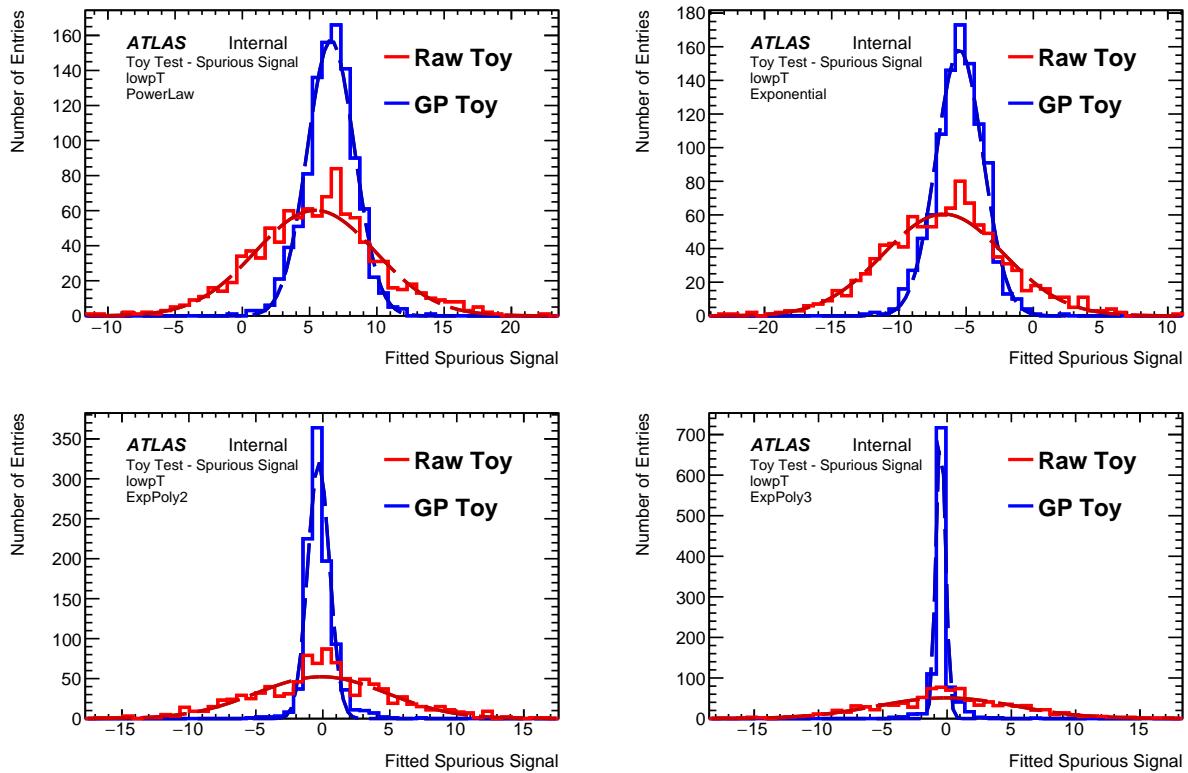


Figure A.108: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side. Each toy in this test has 100k events.

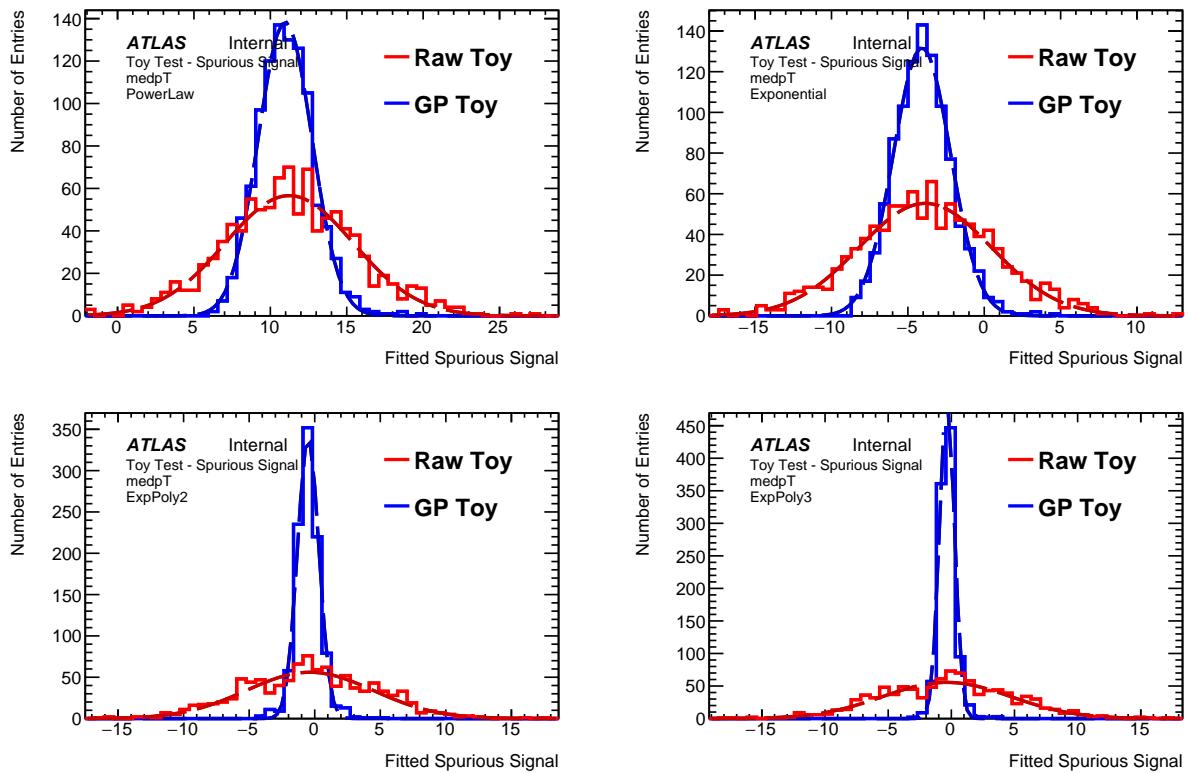


Figure A.109: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 100k events.

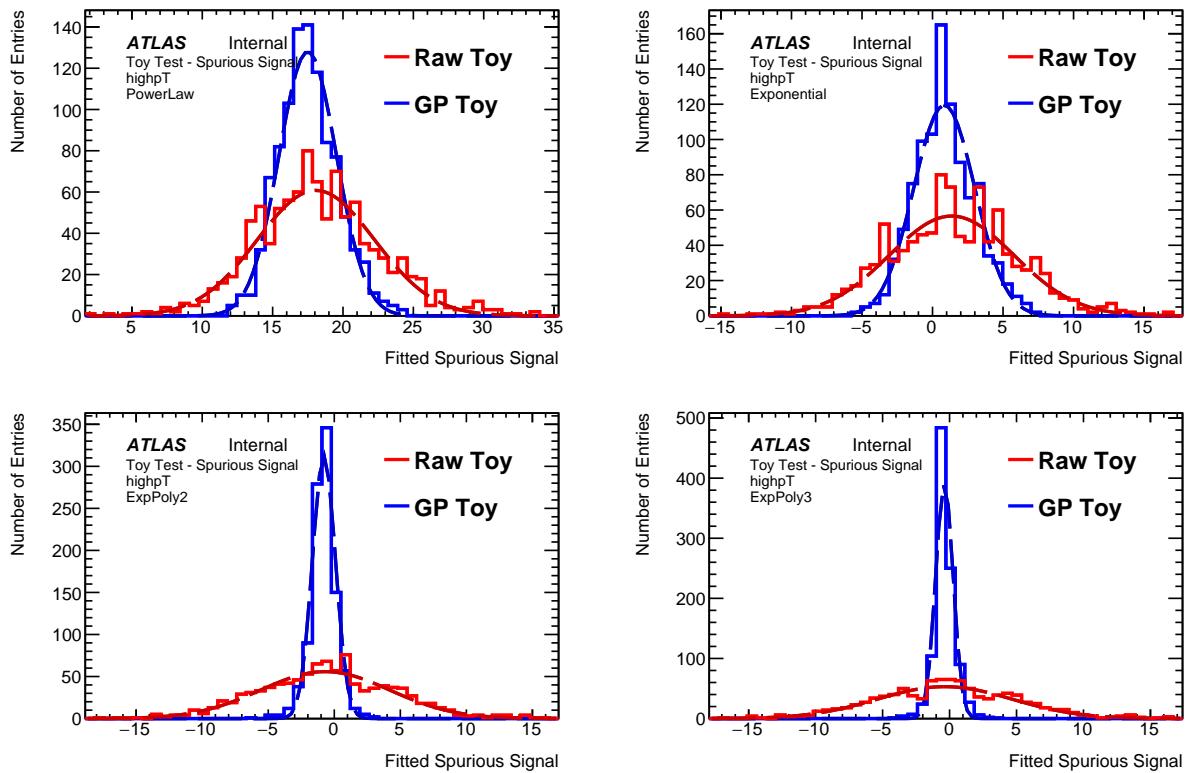


Figure A.110: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 100k events.

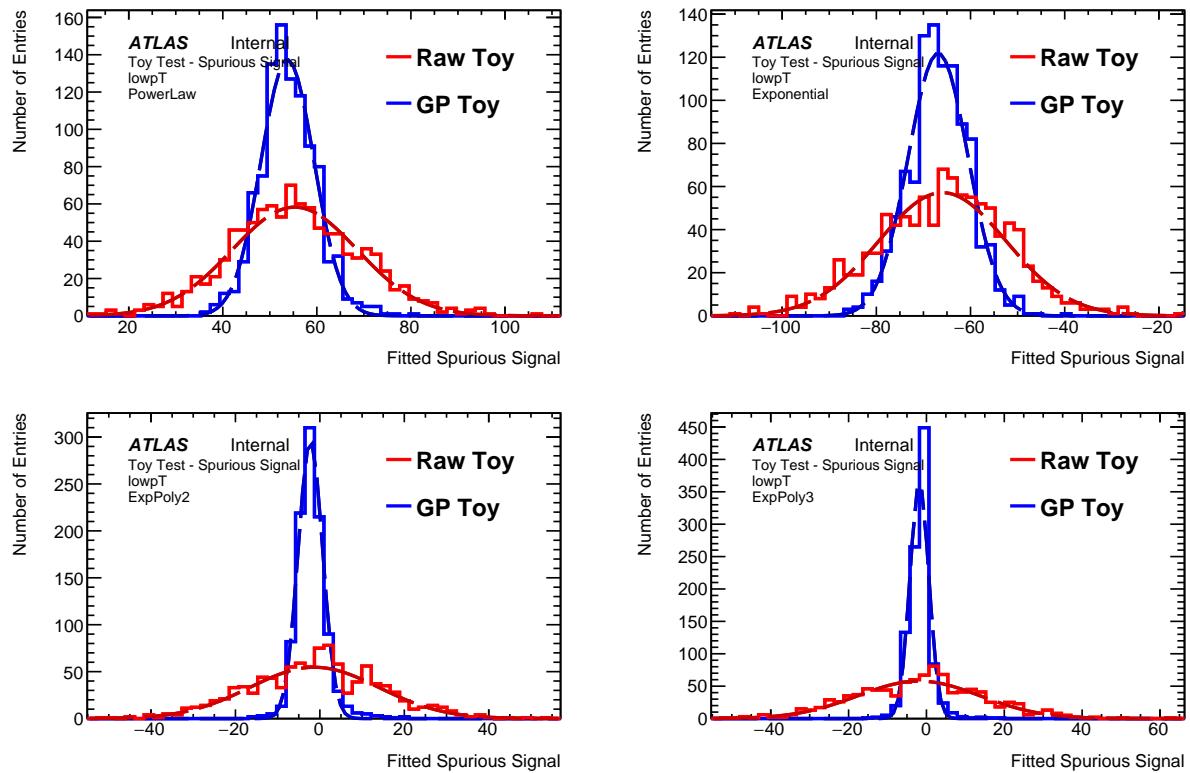


Figure A.111: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side. Each toy in this test has 1,000,000 events.

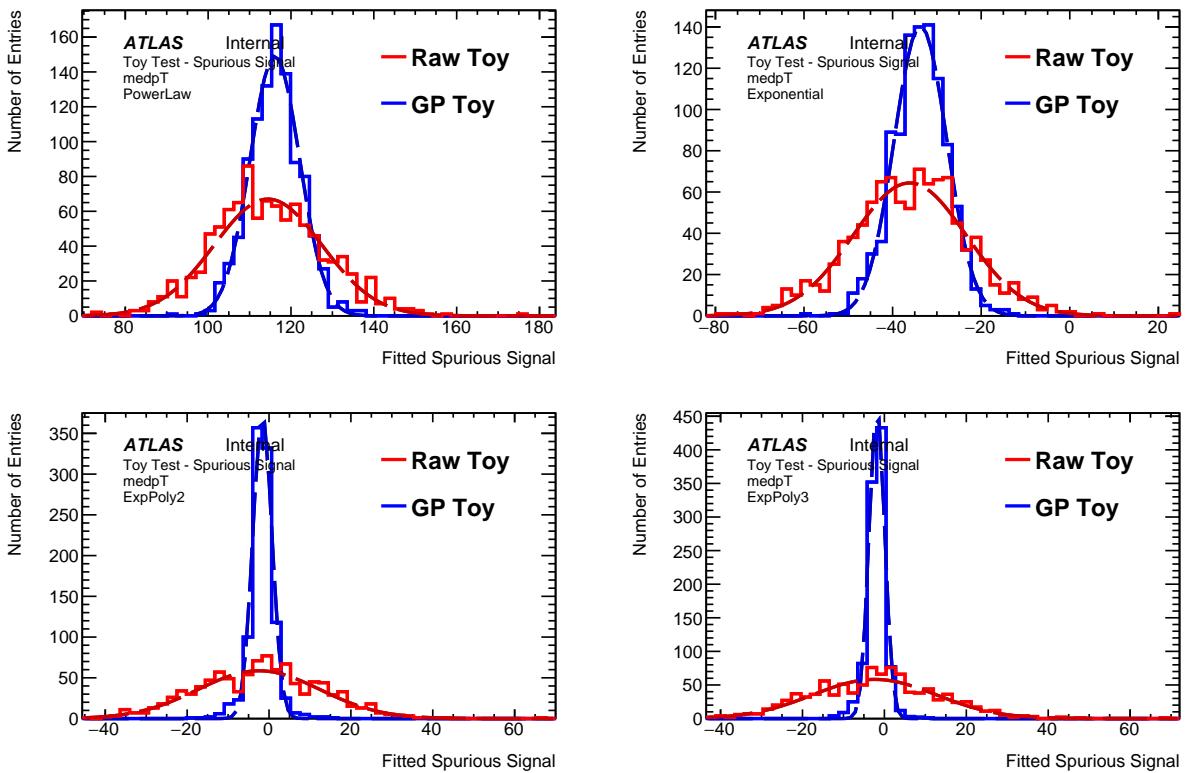


Figure A.112: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 1,000,000 events.

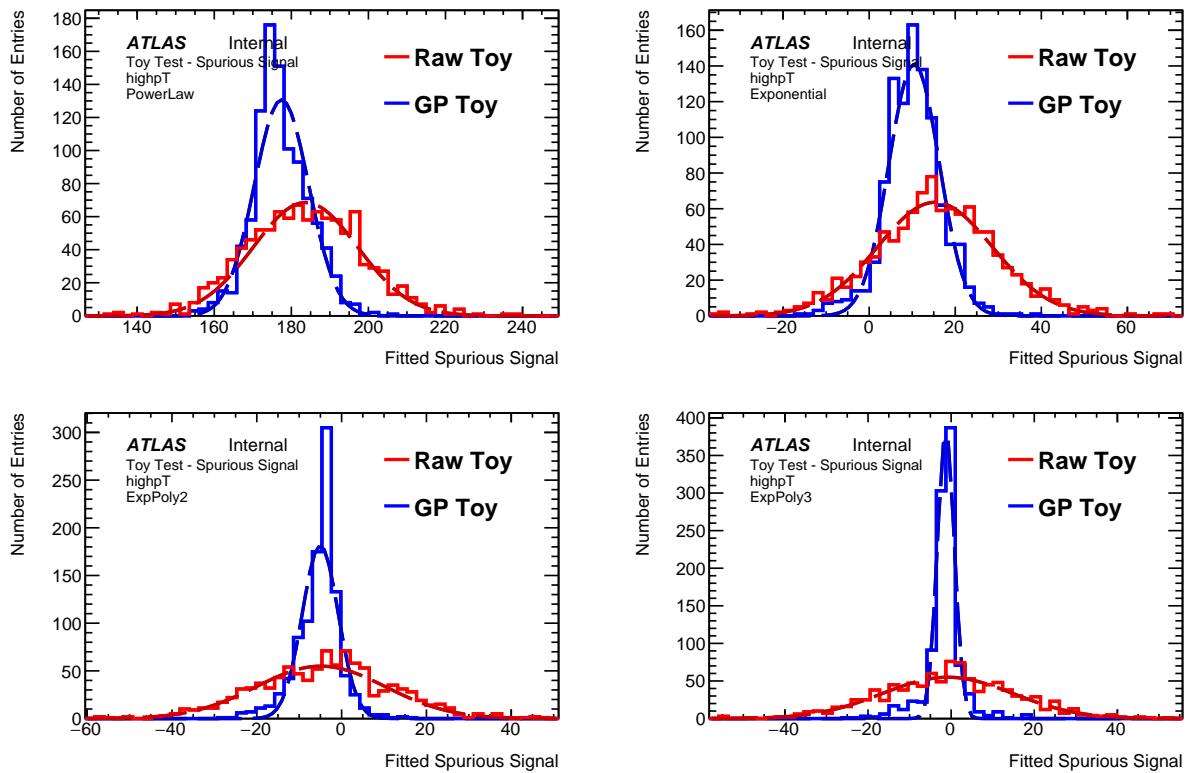


Figure A.113: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 1,000,000 events.

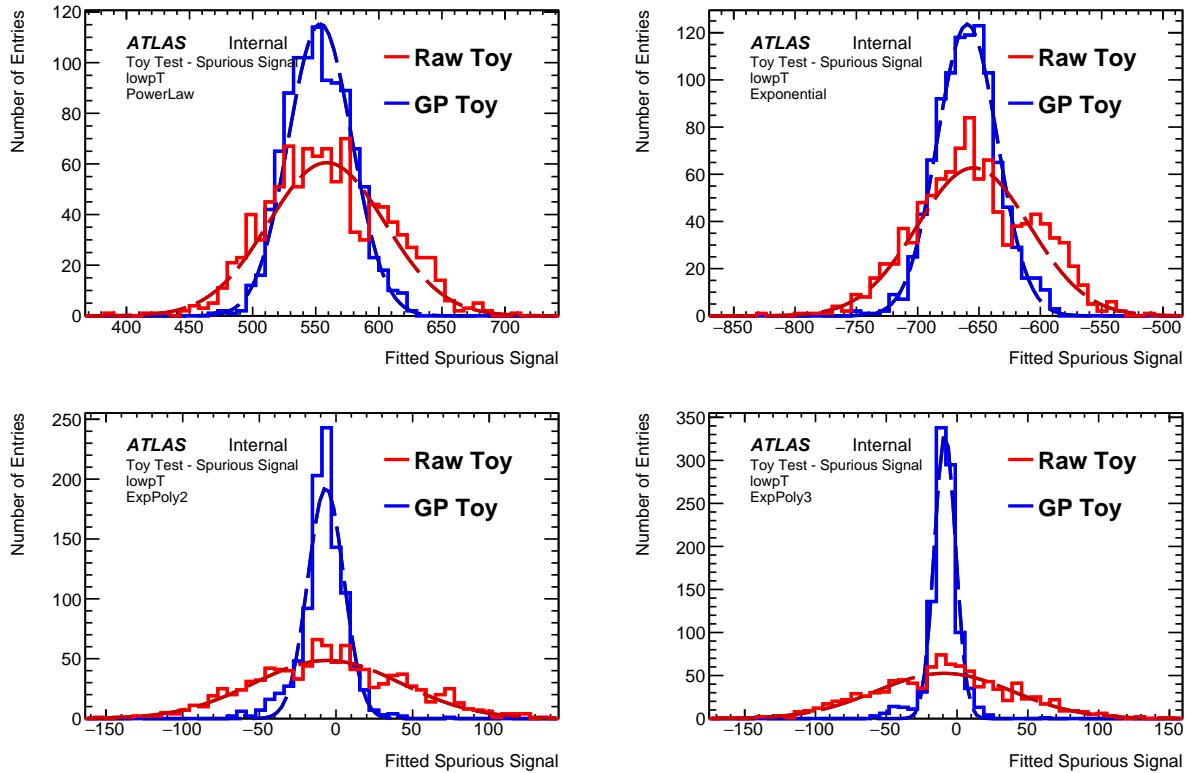


Figure A.114: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side. Each toy in this test has 10,000,000 events.

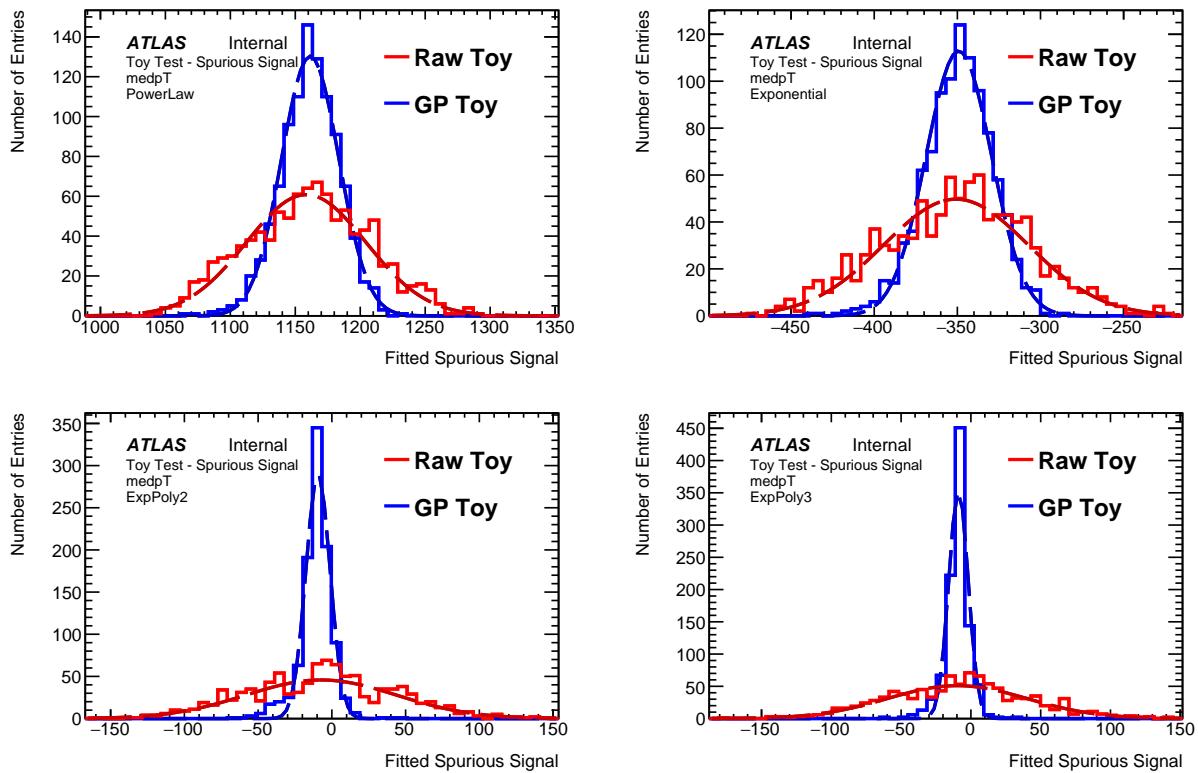


Figure A.115: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 10,000,000 events.

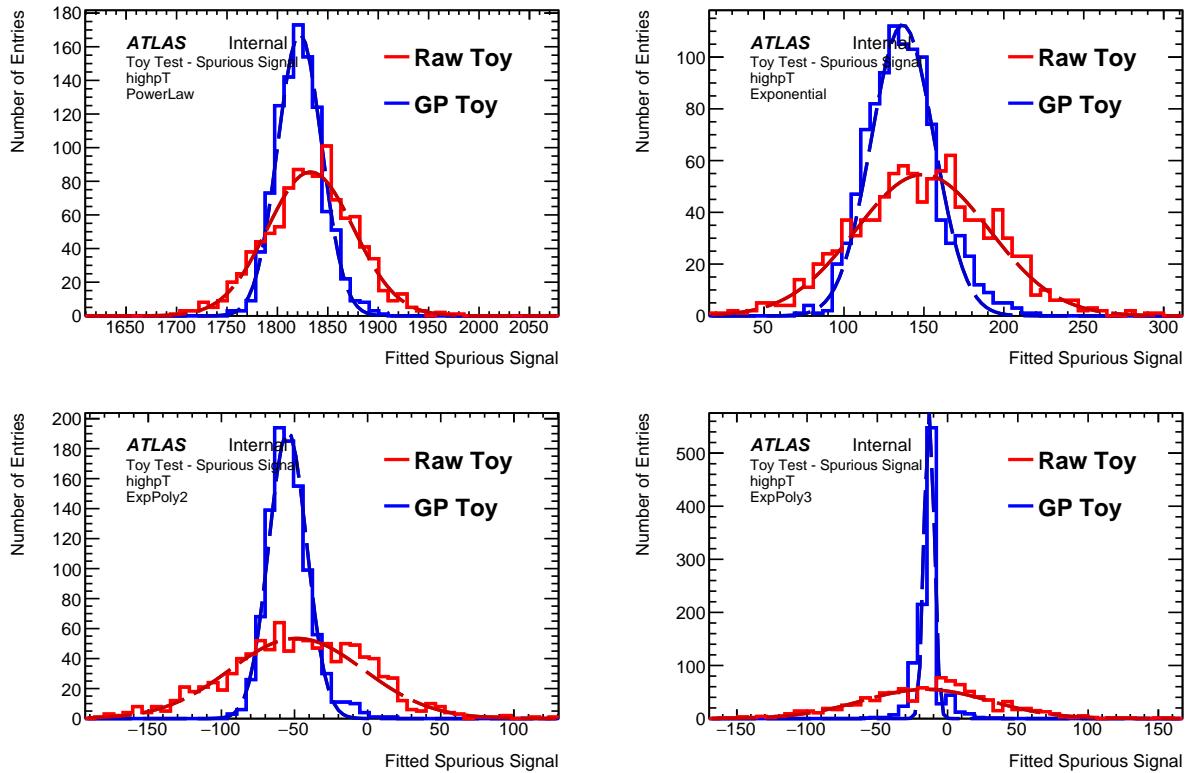


Figure A.116: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 10,000,000 events.

Nominal	N_sig	Unweighted Events	GP-Raw: Bias, Mean	Sigma Toy	Sigma GPR	Bias/sigma(SS_raw)
Fit: Exp						
Generating: Exp2						
low10	0	10^1	-0.01	0.19	0.13	-0.06
low100	0	10^2	0.26	0.15	0.85	1.69
low1k	0	10^2	-0.11	0.45	0.29	-0.24
low10k	0	10^4	-0.04	1.43	0.50	-0.03
low100k	0	10^5	-1.11	4.63	1.79	-0.24
low1M	0	10^6	0.89	13.82	6.48	0.06
low10M	0	10^7	6.19	46.86	25.26	0.13
Fit: Exp2						
Generating: Exp3						
med10	0	10^1	-0.13	0.20	0.03	-0.67
med100	0	10^2	0.00	0.18	0.28	0.03
med1k	0	10^2	-0.01	0.49	0.36	-0.02
med10k	0	10^4	0.04	1.63	0.24	0.02
med100k	0	10^5	0.07	4.95	0.97	0.01
med1M	0	10^6	-0.49	14.85	3.08	-0.03
med10M	0	10^7	3.32	49.83	10.93	0.07
Fit: Exp2						
Generating: Exp3						
high10	0	10^1	-0.12	0.19	0.03	-0.60
high100	0	10^2	0.00	0.15	0.19	0.00
high1k	0	10^2	-0.10	0.51	0.55	-0.20
high10k	0	10^4	0.10	1.66	0.23	0.06
high100k	0	10^5	0.29	5.09	0.99	0.06
high1M	0	10^6	1.23	15.91	5.18	0.08
high10M	0	10^7	5.86	47.57	14.70	0.12

Table A.12: Spurious signal means and widths for the three test functional-form distributions for a range of different template statistics.

A.5.3 Extended Templates, Linear Error Kernel

For the stage in which we condition the Gaussian Process on the template, we can model the expected noise in a number of ways. The nominal approach performed in the previous study treats the noise level as being defined by the error bands of the original input template. However, the known physics of the template (smoothly-falling) suggests it is reasonable to implement a linear error kernel to model the noise. In this step, rather than fixing the variance at each training point to the value given by the errors on the templates, we can treat the variance as a hyperparameter that is optimized in the GP fit- a "white" kernel uses one hyperparameter (the constant noise value) and a linear error kernel uses two (the initial value and rate-of-decrease of the variance on the training points).

Using both the extended templates and the linear noise kernel, we are able to reduce the bias to less than 20% of the stat uncertainty on the spurious signal for templates with greater than an average 20 effective MC events/ bin (i.e, templates with more than 2600 events) which is enough to claim that GPR is unbiased in this context. From a statistics perspective, 20 effective events per bin is also the known threshold above which the Poission-as-Gaussian approximation that allows us to employ GPR holds.

We note that the seeming lack of bias far below the 20 effective Monte-Carlo event per bin threshold is likely due to the prevalence of a large number of flat-line fits due to empty bins, while a larger bias is apparent just below the threshold (i.e., the 1000 event templates, which have 9 events per bin, are dominated by Poisson statistics).

We report the results of the padded-template, linear-error kernel bias study for all categories in the Table ???. Additionally, we plot the per-toy bias (that is, $\text{GP_SS} - \text{Raw_SS}$ for each toy) for each shape and stat regime.

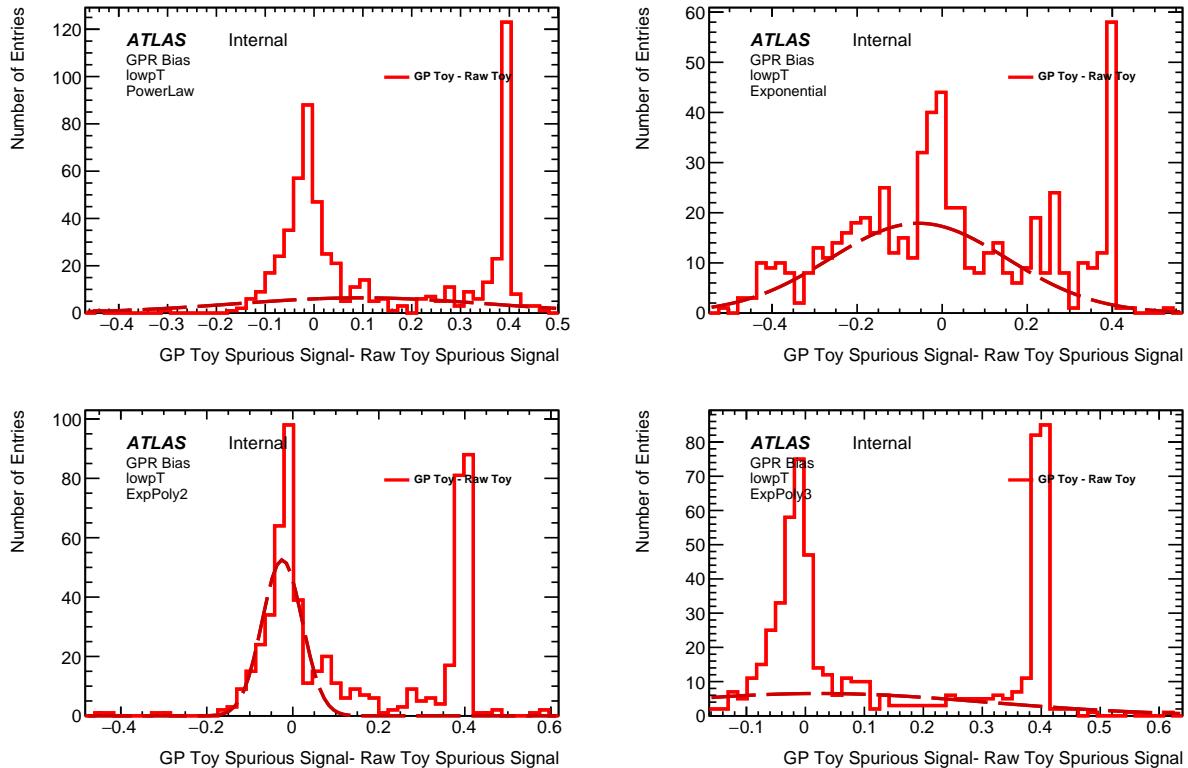


Figure A.117: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template extended by 5 GeV on either side. Each toy in this test has 10 events.

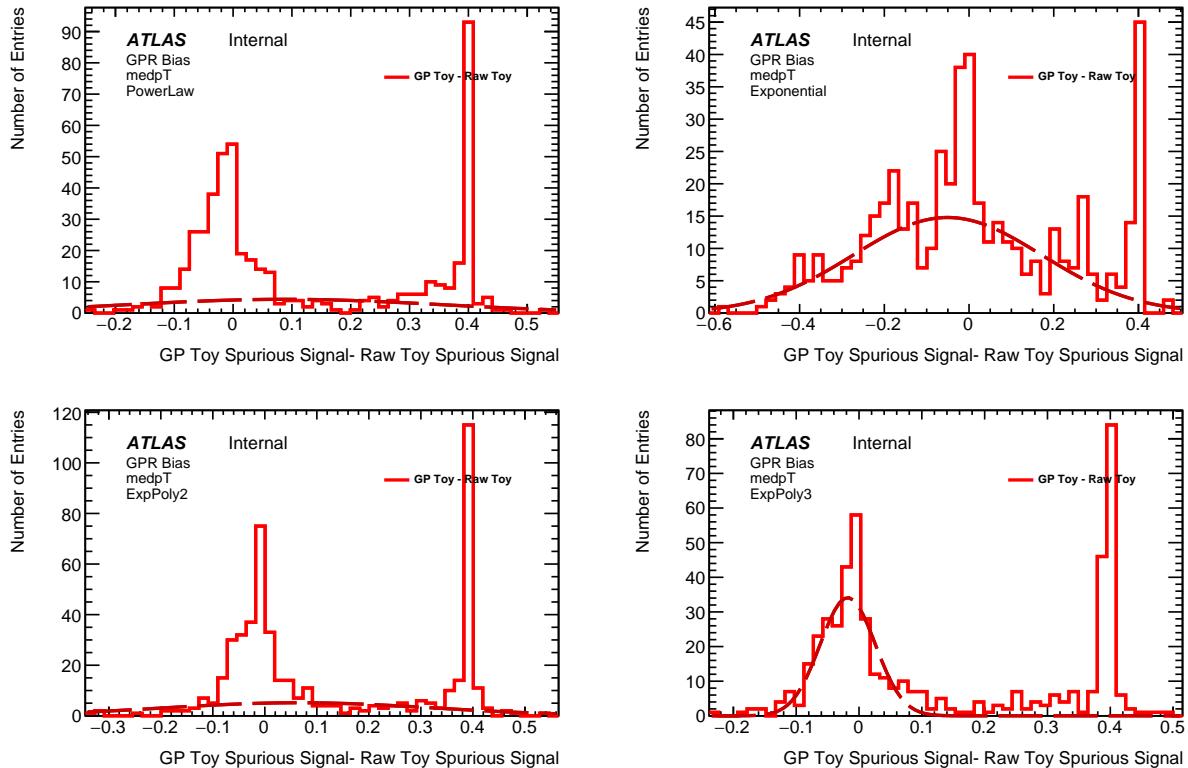


Figure A.118: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 10 events.

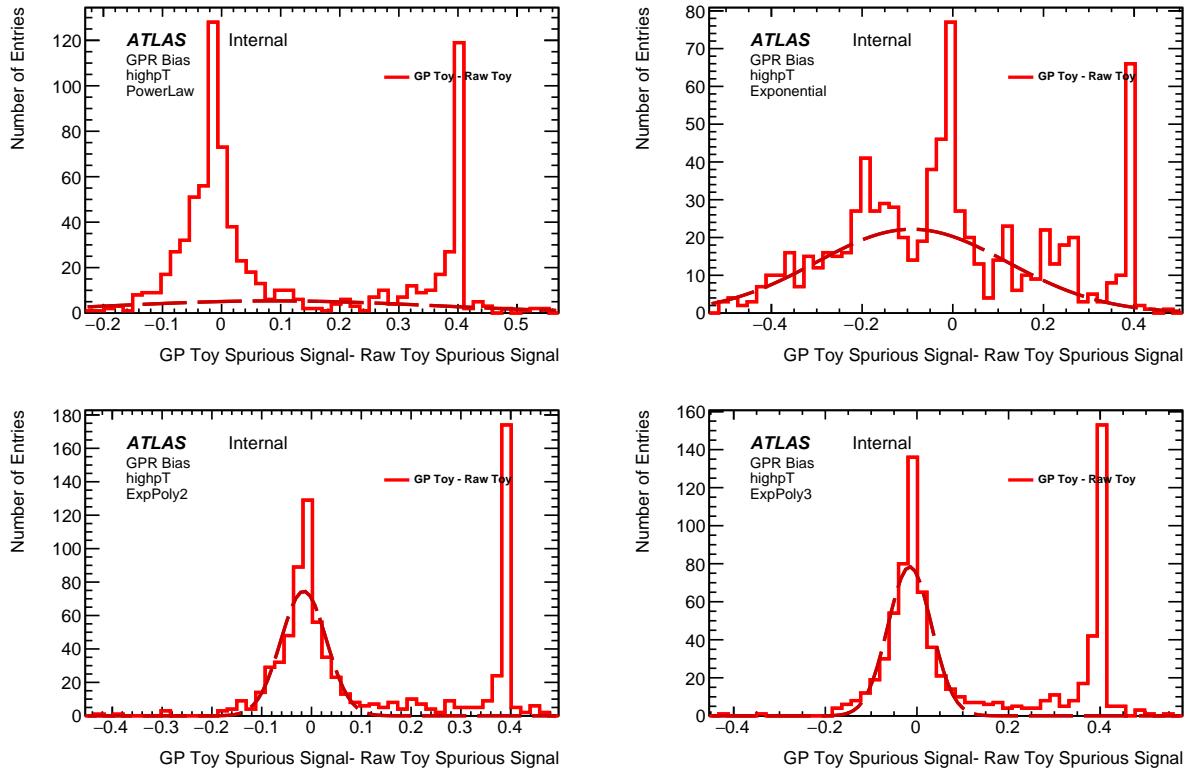


Figure A.119: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 10 events.

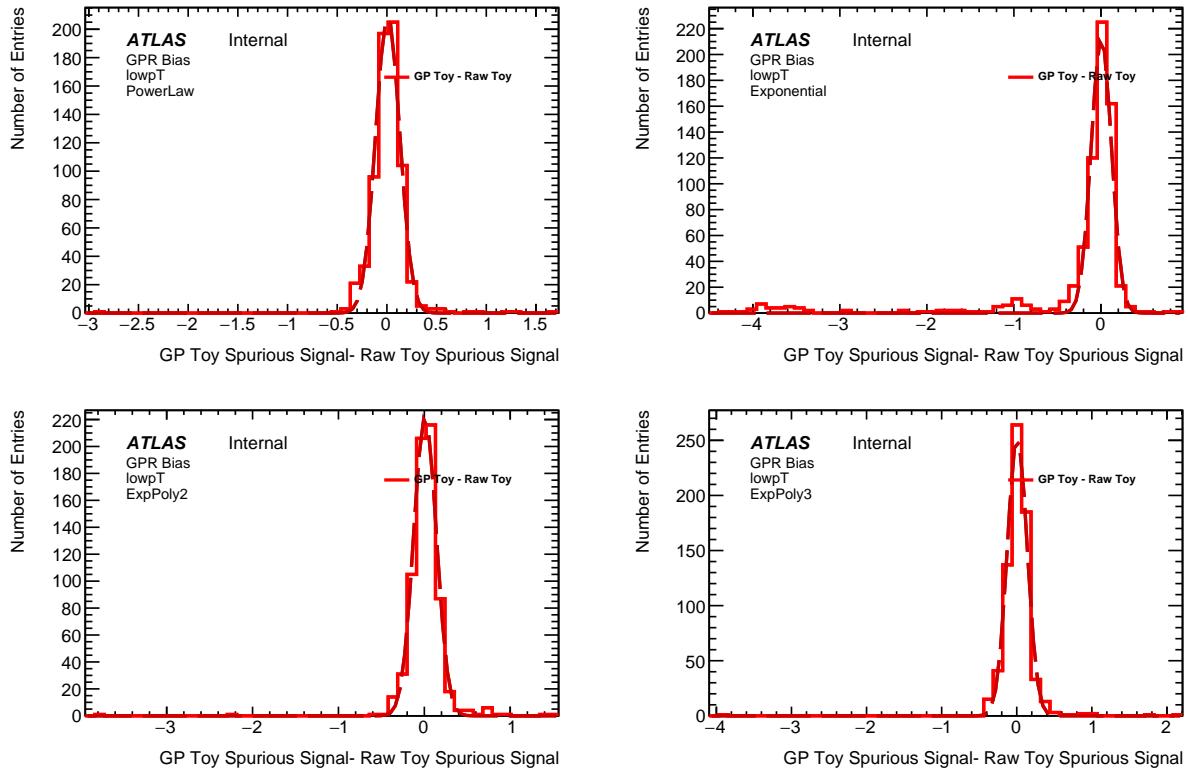


Figure A.120: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template extended by 5 GeV on either side. Each toy in this test has 100 events.

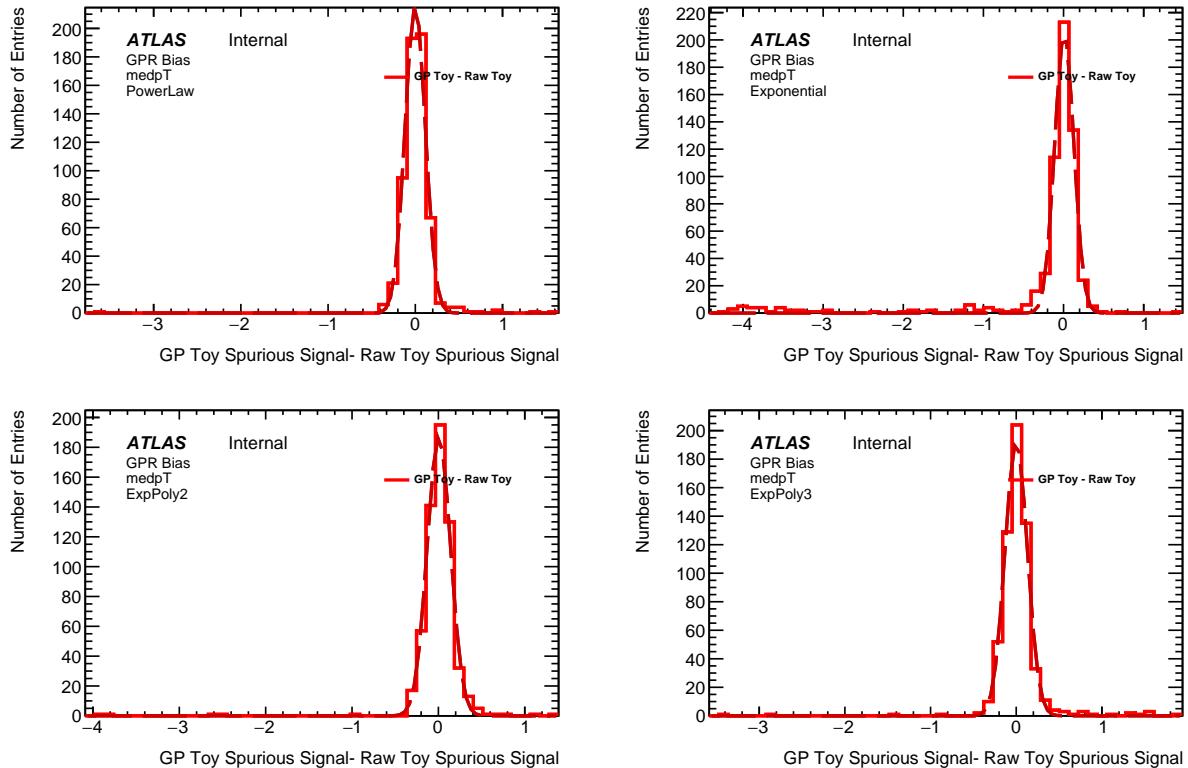


Figure A.121: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 100 events.

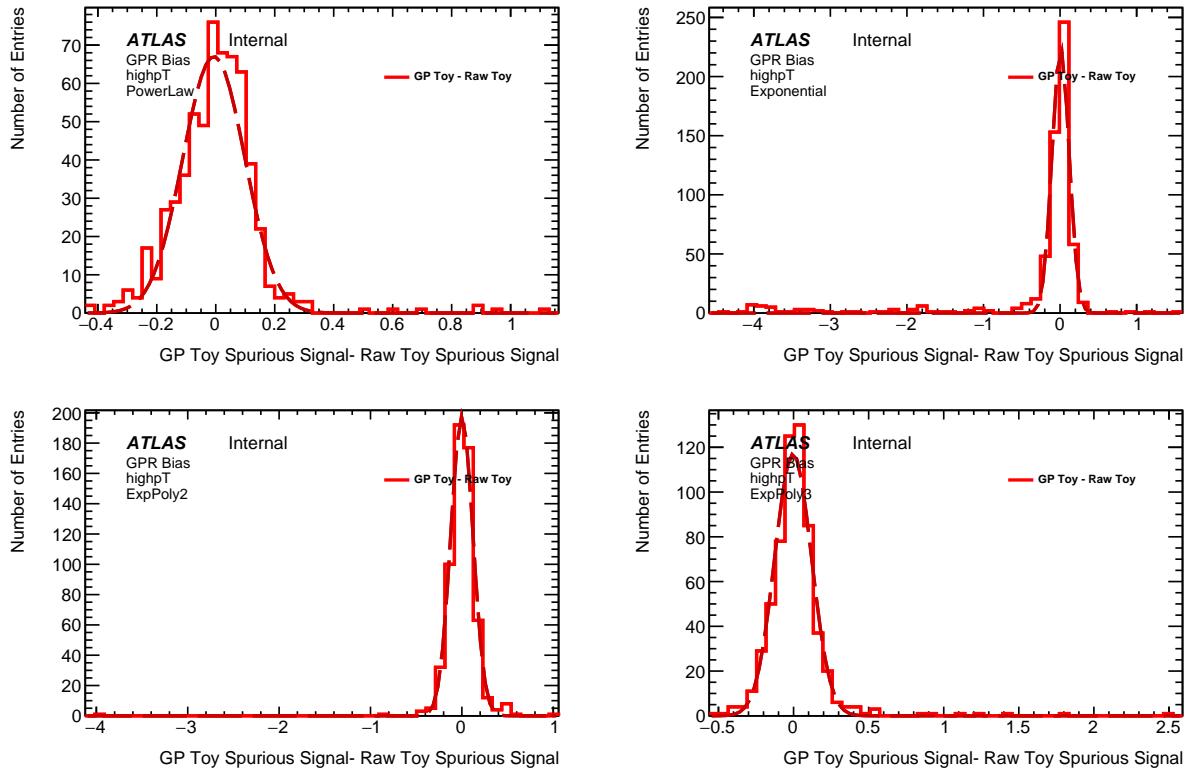


Figure A.122: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 100 events.

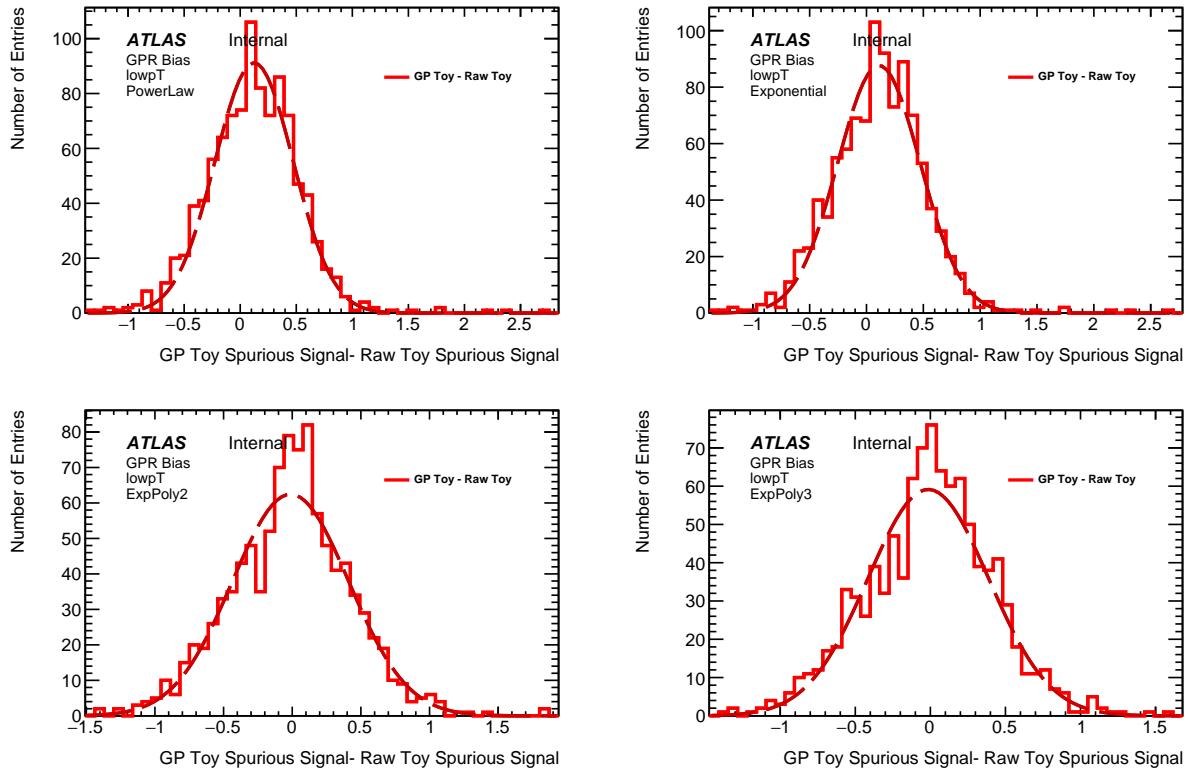


Figure A.123: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template extended by 5 GeV on either side. Each toy in this test has 1000 events.

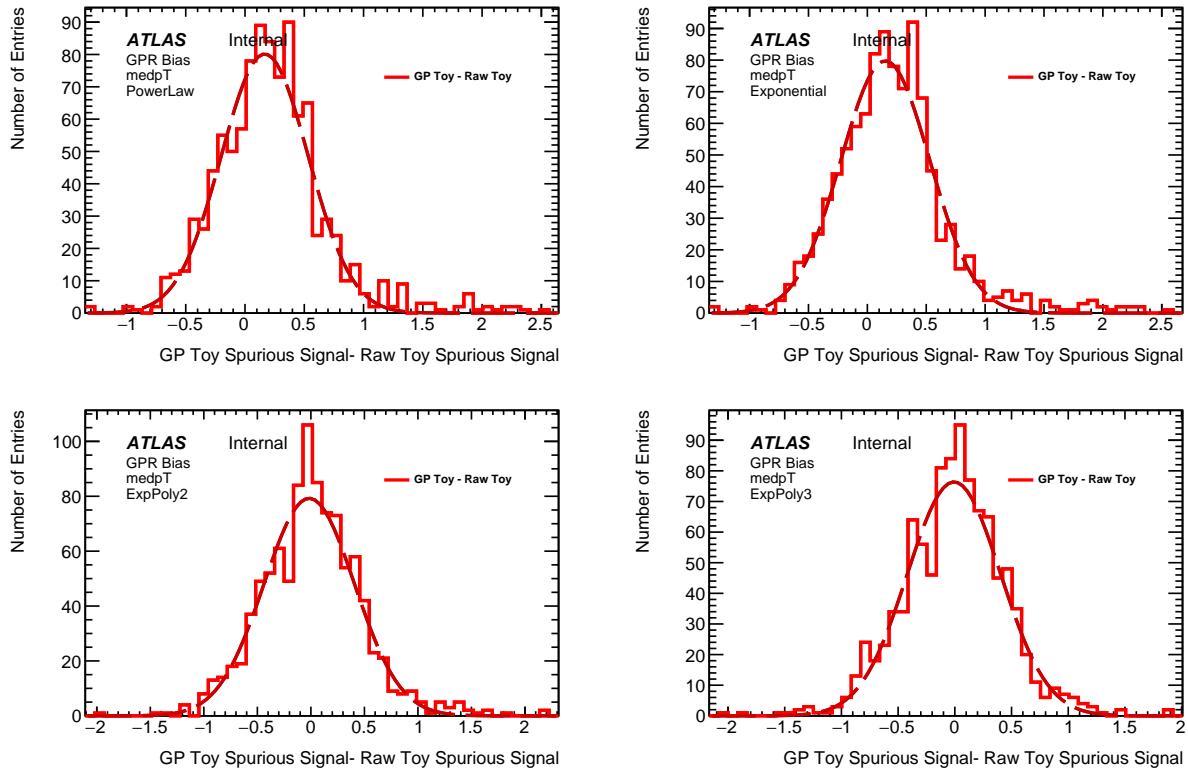


Figure A.124: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 1000 events.

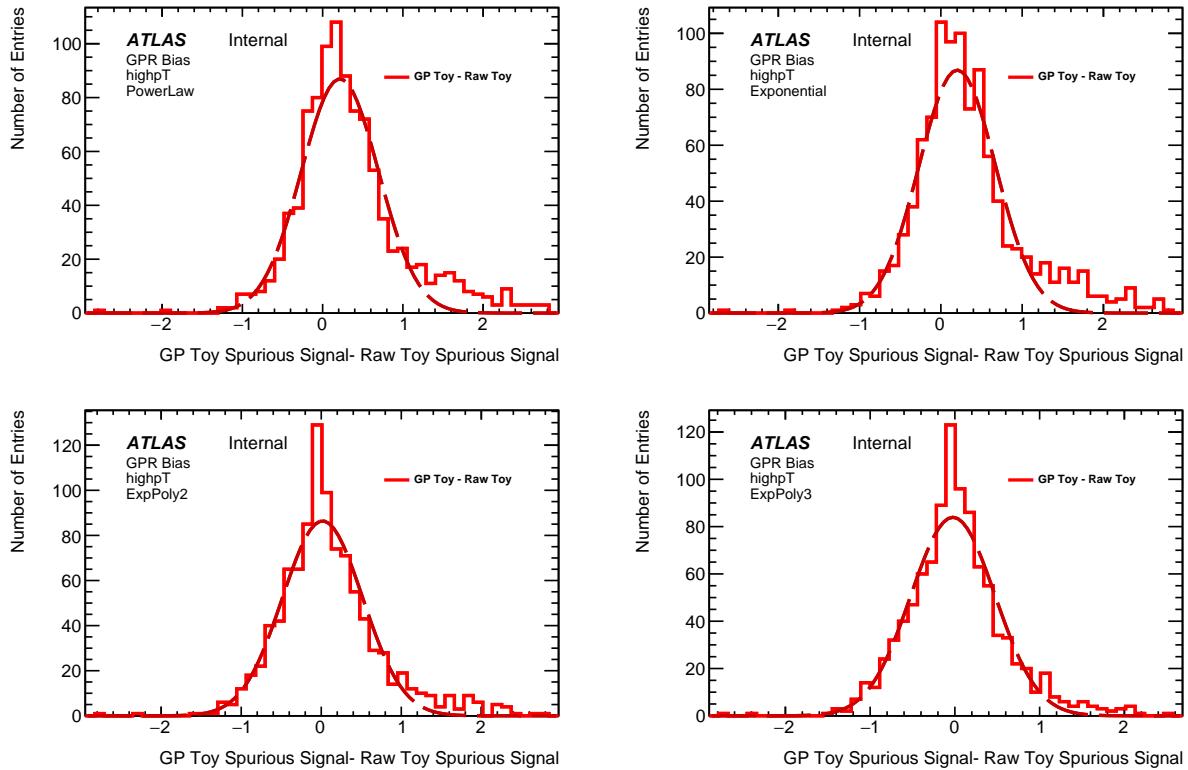


Figure A.125: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 1000 events.

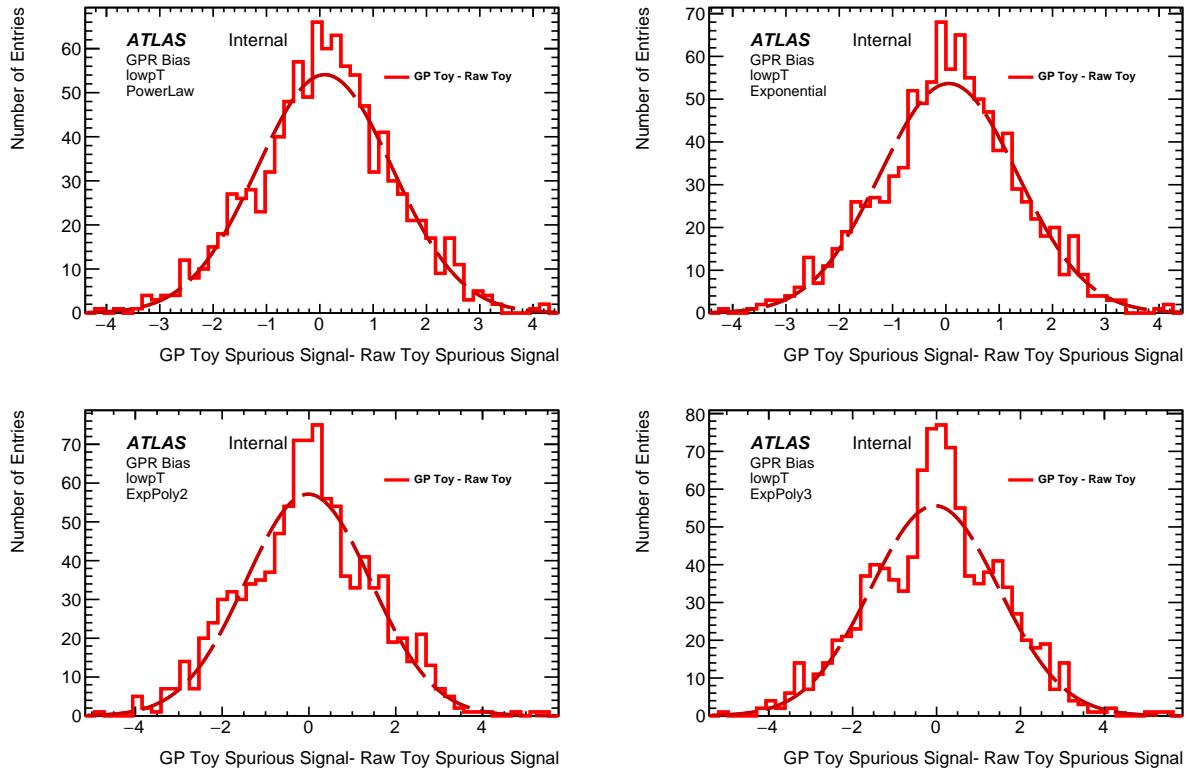


Figure A.126: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template extended by 5 GeV on either side. Each toy in this test has 10k events.

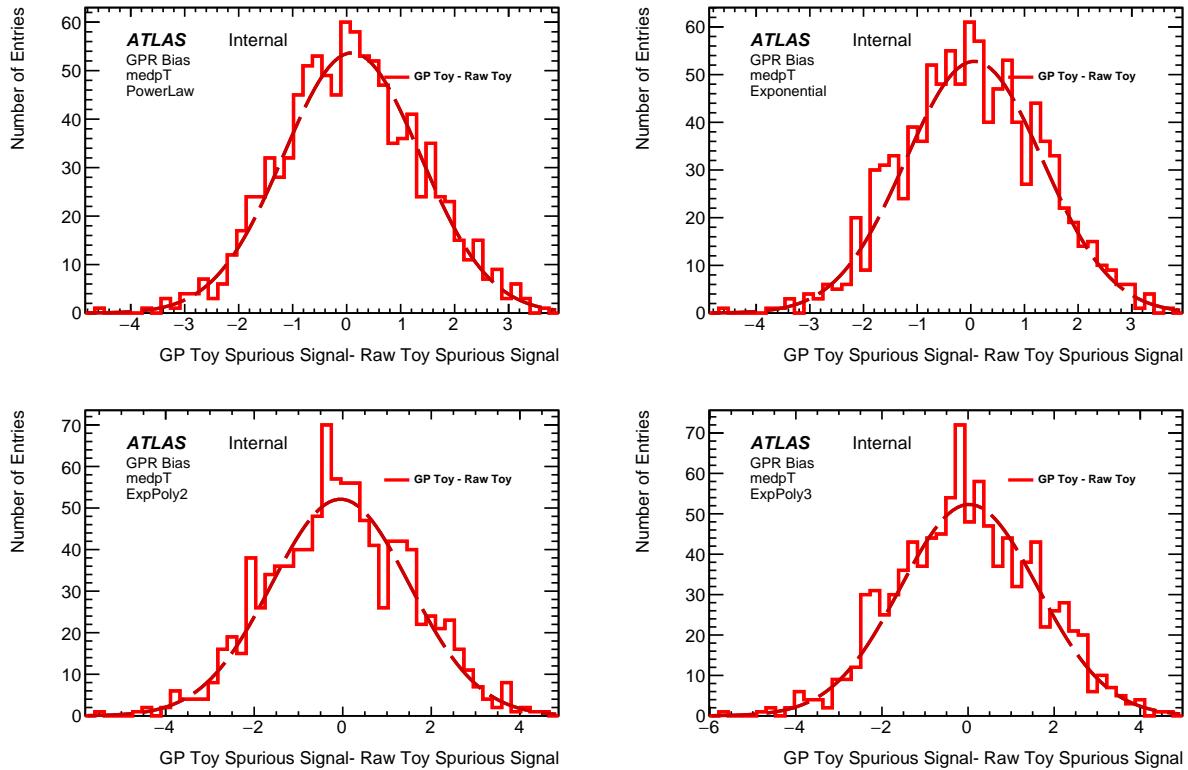


Figure A.127: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 10k events.

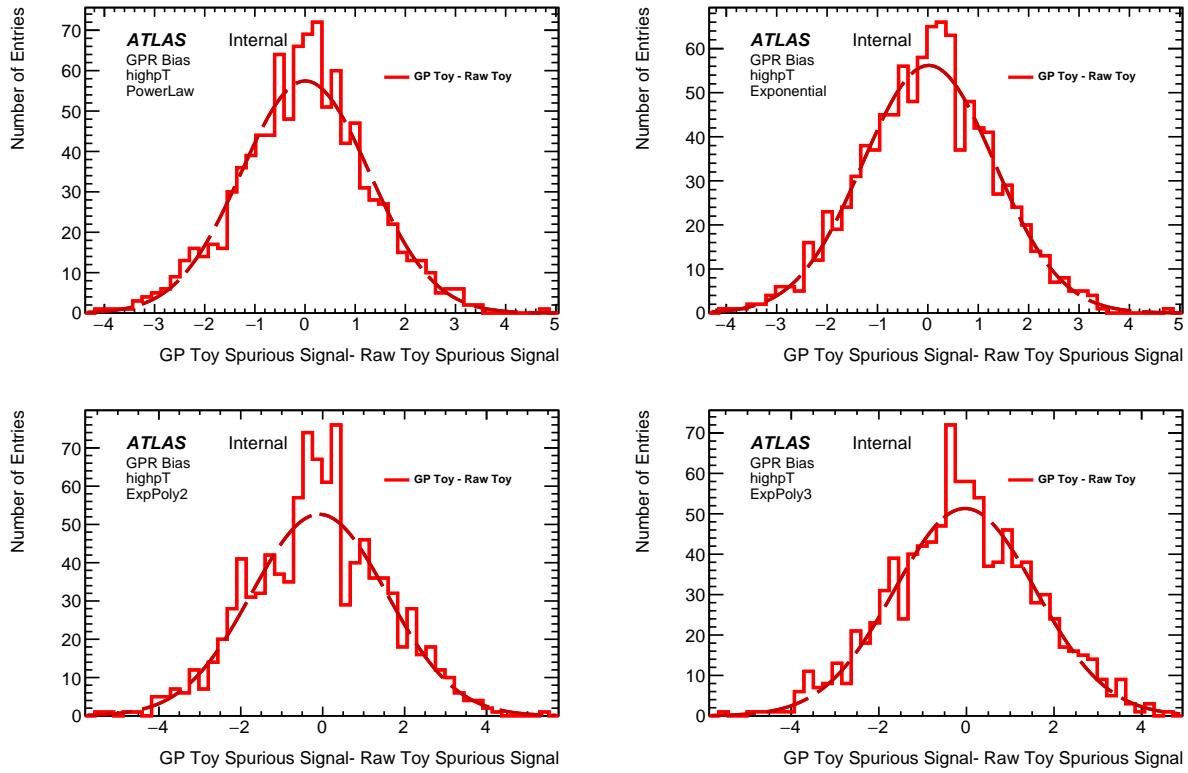


Figure A.128: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 10k events.

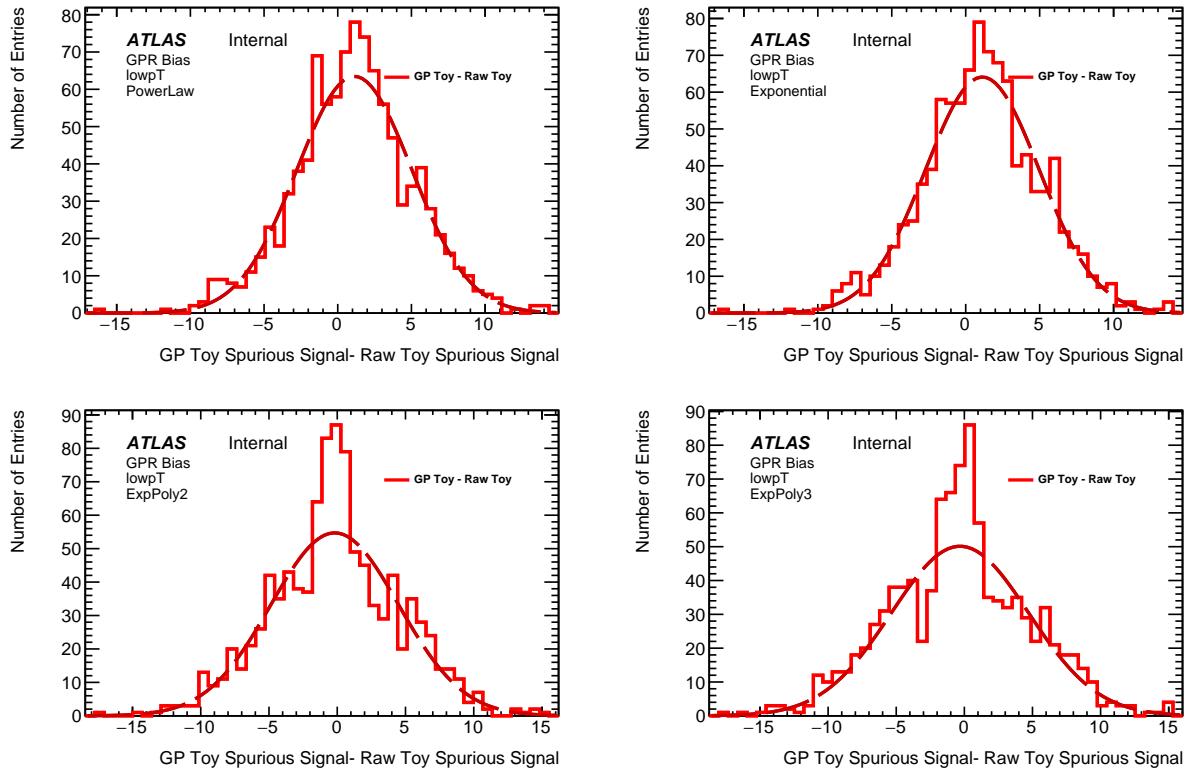


Figure A.129: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template extended by 5 GeV on either side. Each toy in this test has 100k events.

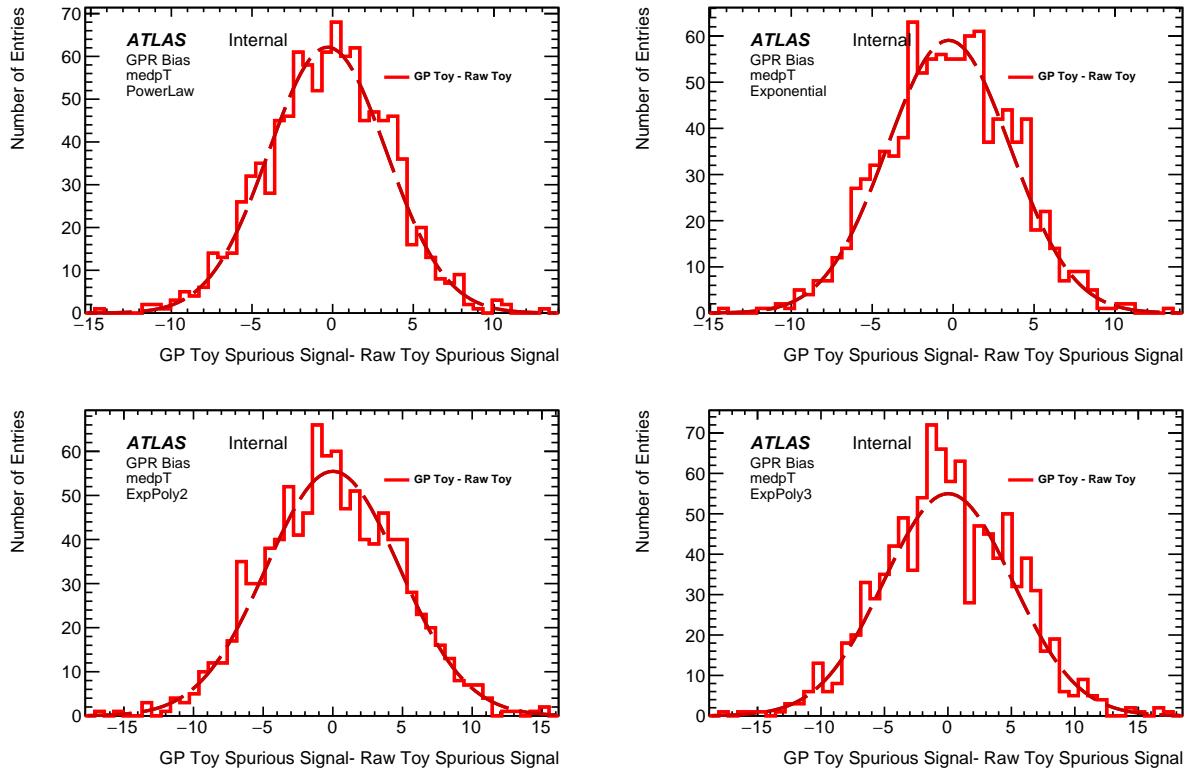


Figure A.130: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 100k events.

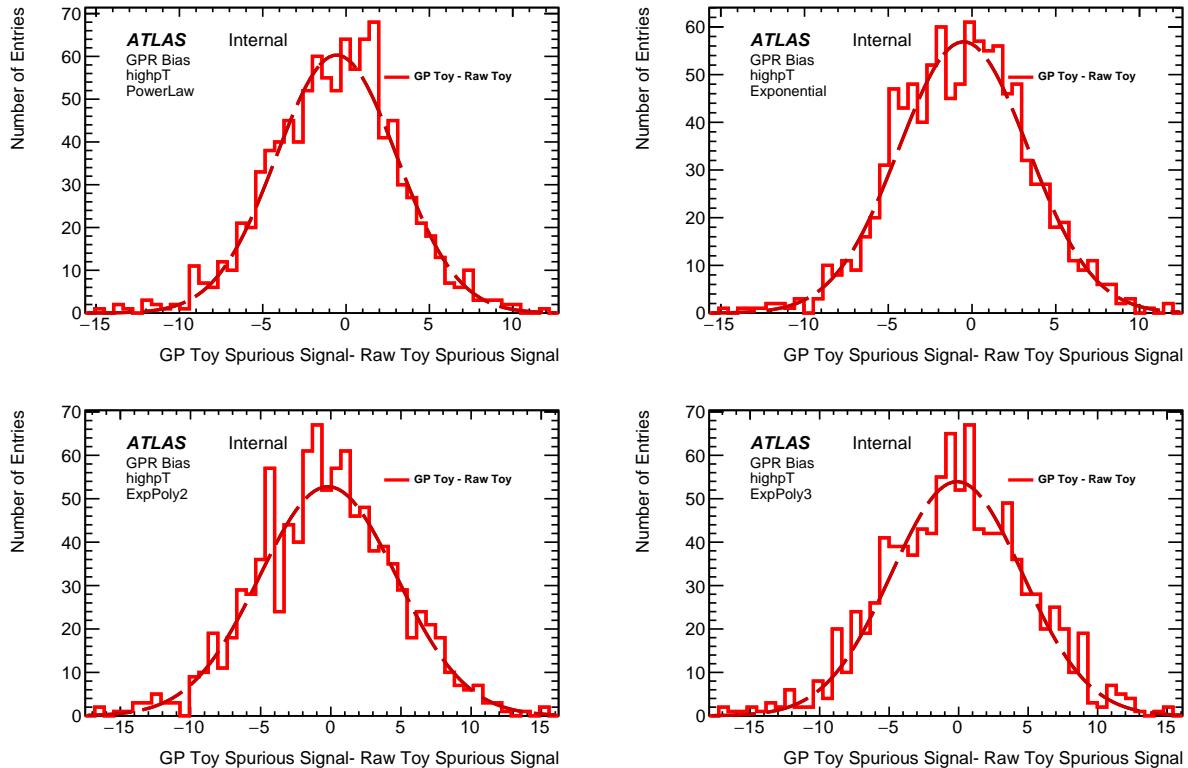


Figure A.131: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 100k events.

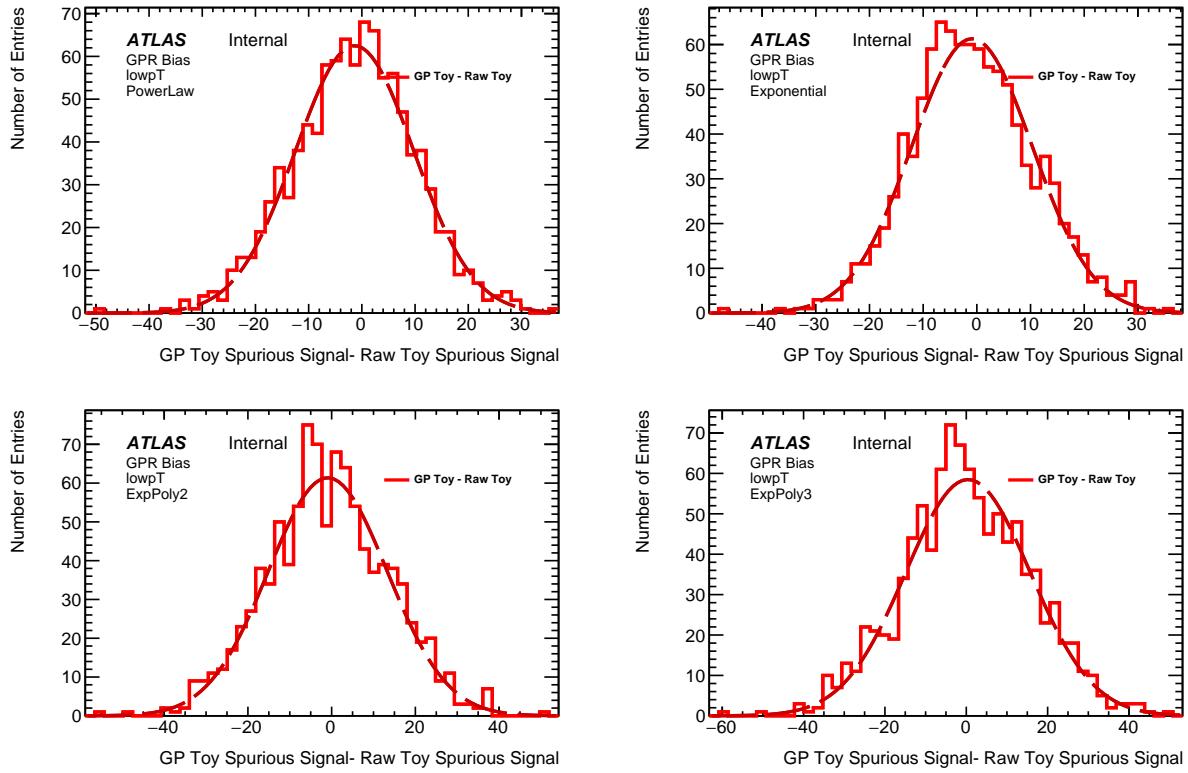


Figure A.132: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template extended by 5 GeV on either side. Each toy in this test has 1,000,000 events.

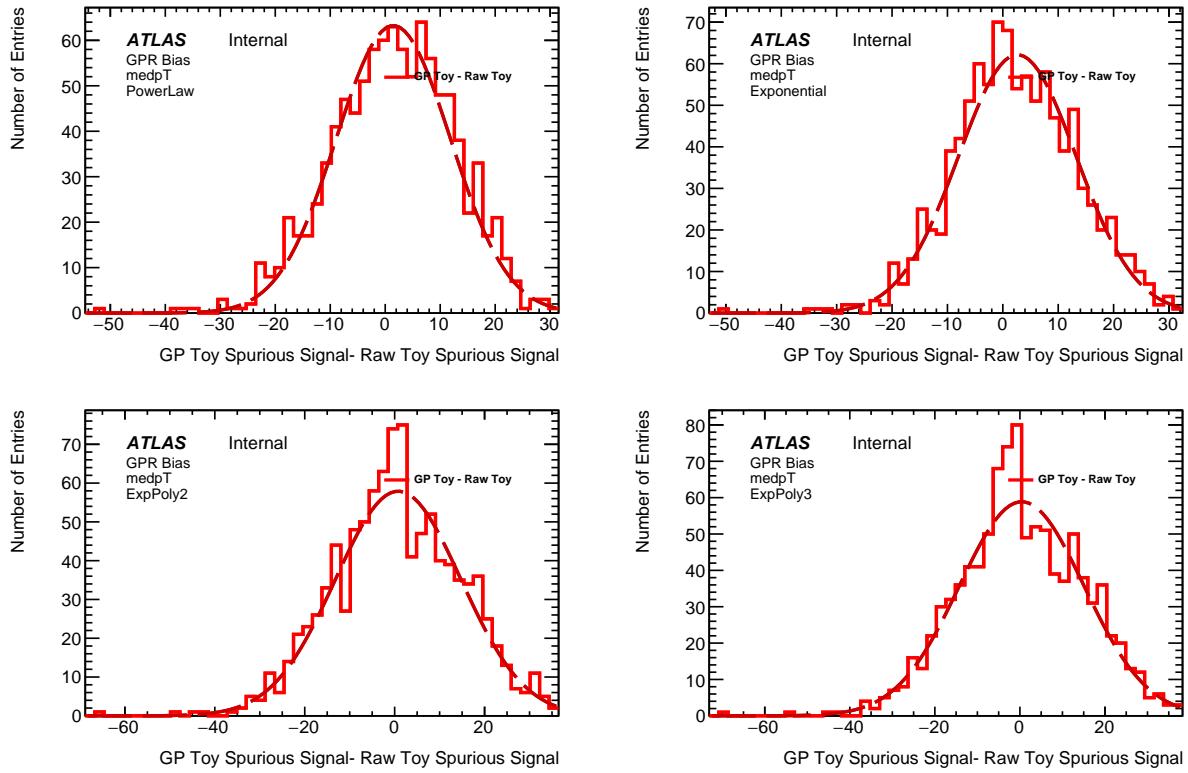


Figure A.133: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 1,000,000 events.

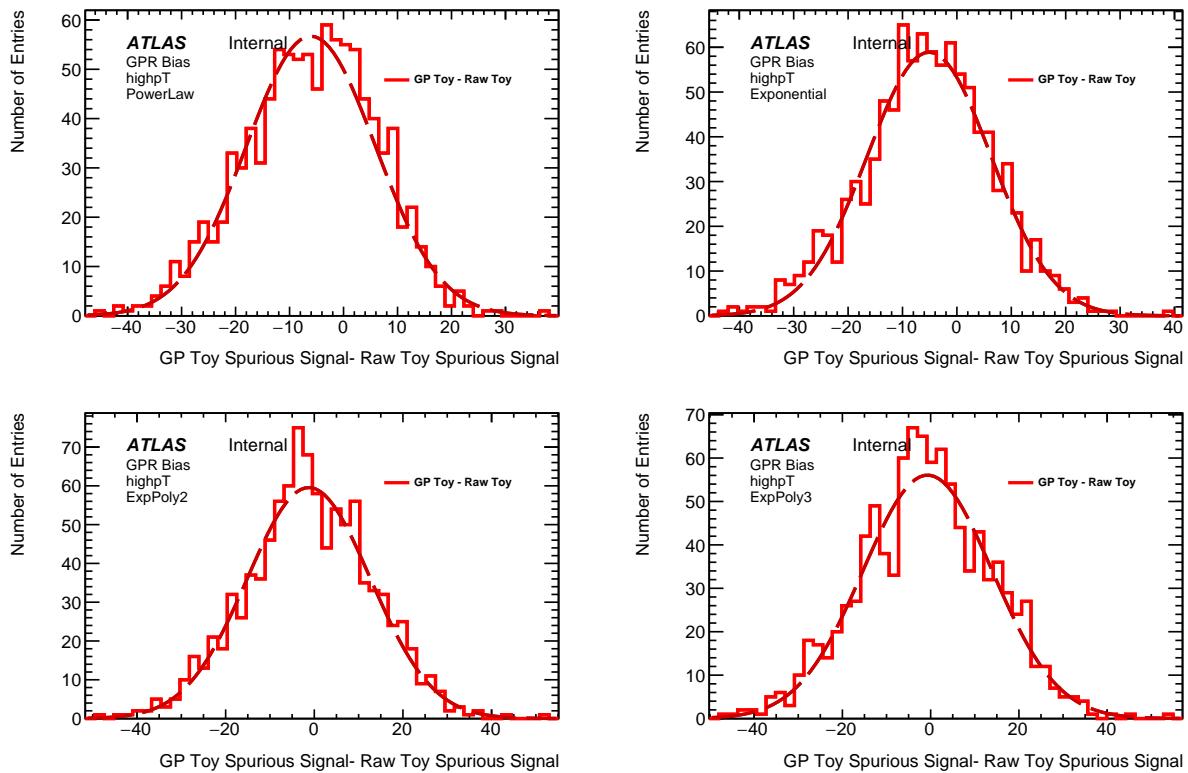


Figure A.134: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 1,000,000 events.

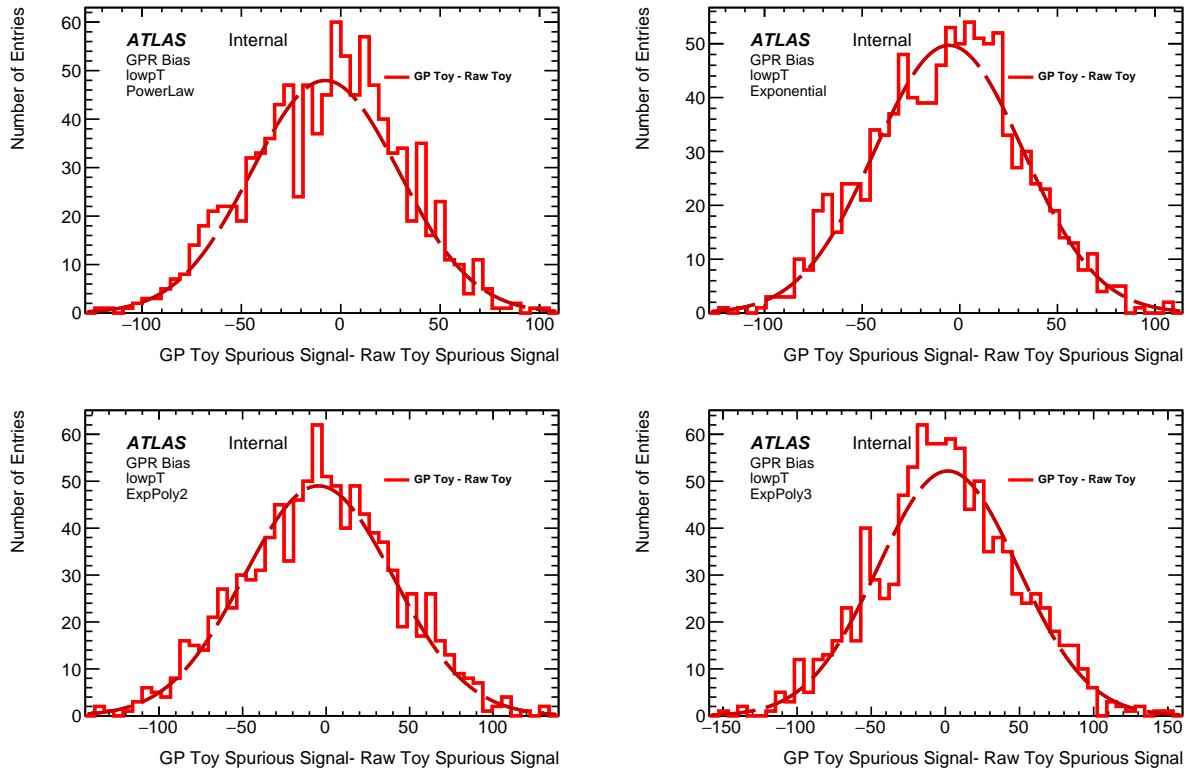


Figure A.135: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template extended by 5 GeV on either side. Each toy in this test has 10,000,000 events.

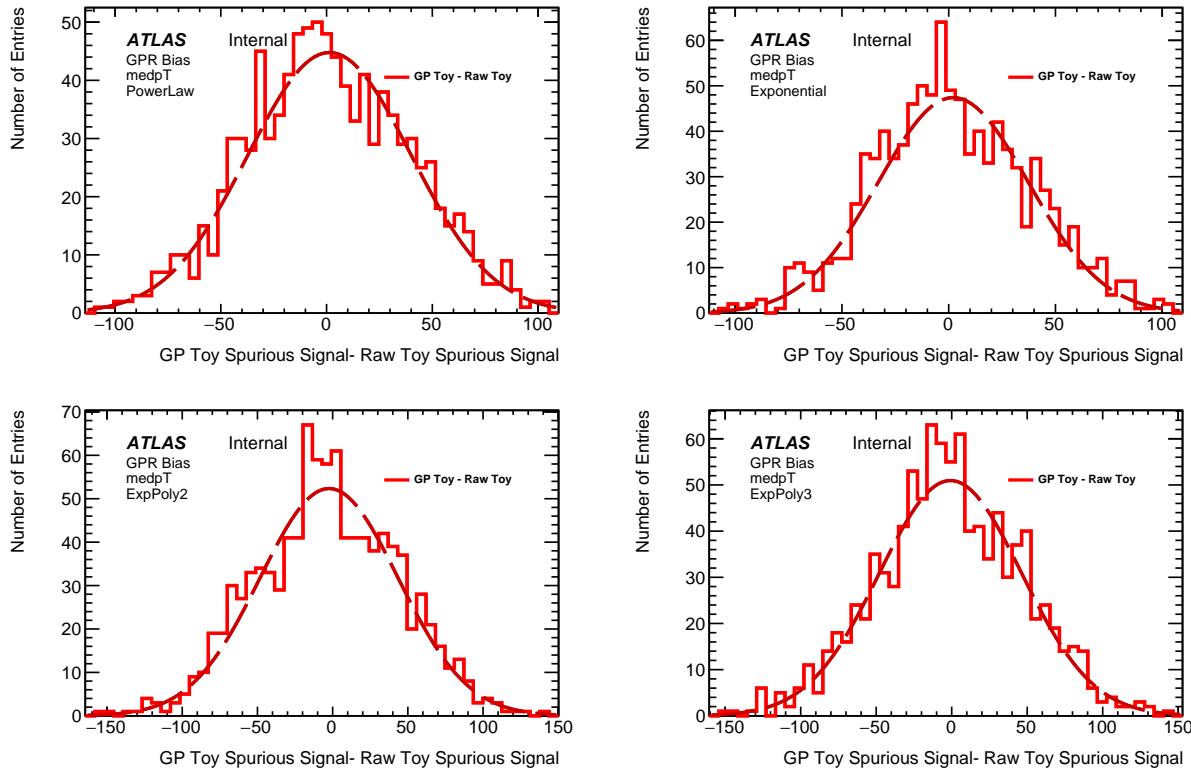


Figure A.136: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 10,000,000 events.

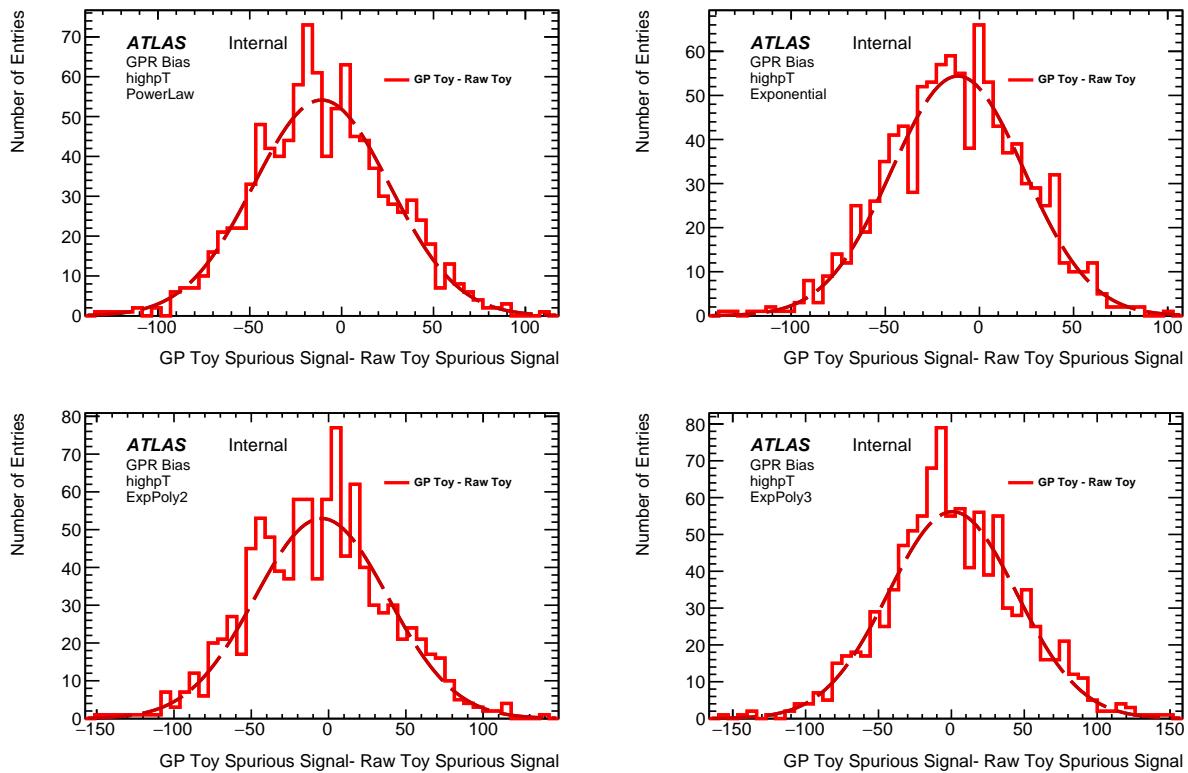


Figure A.137: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template extended by 5 GeV on either side. Each toy in this test has 10,000,000 events.

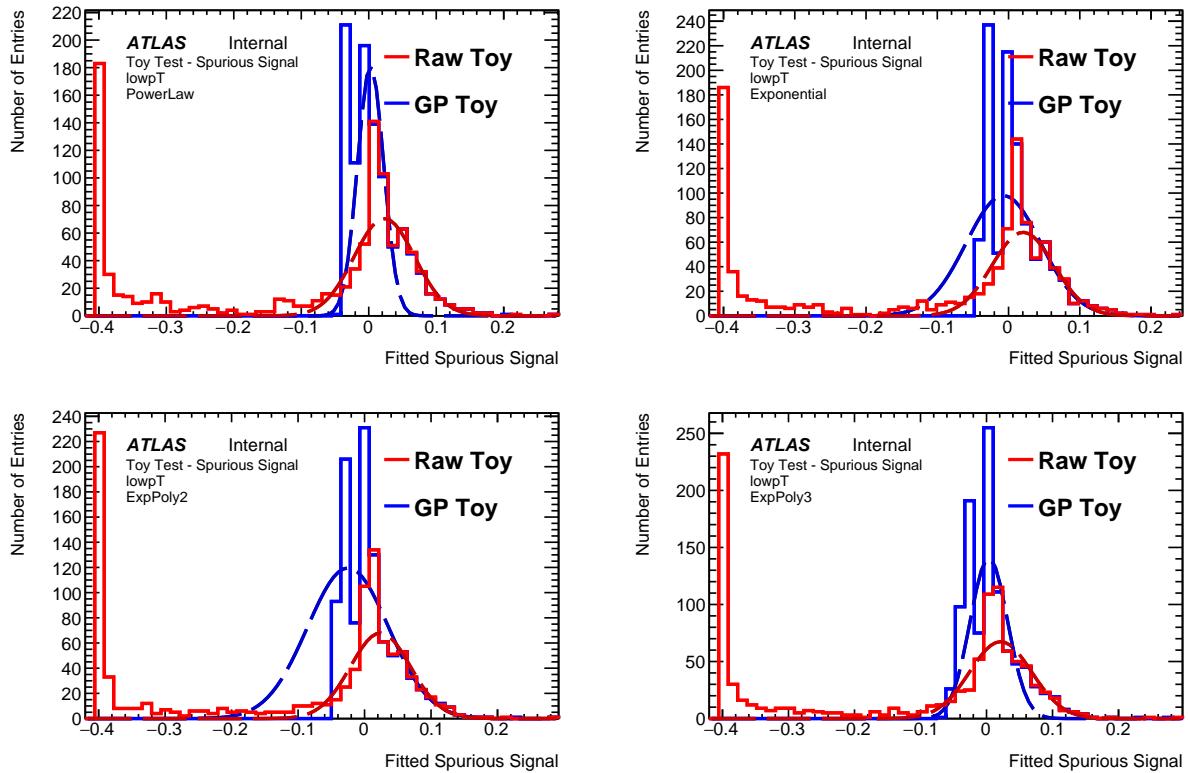


Figure A.138: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10 events.

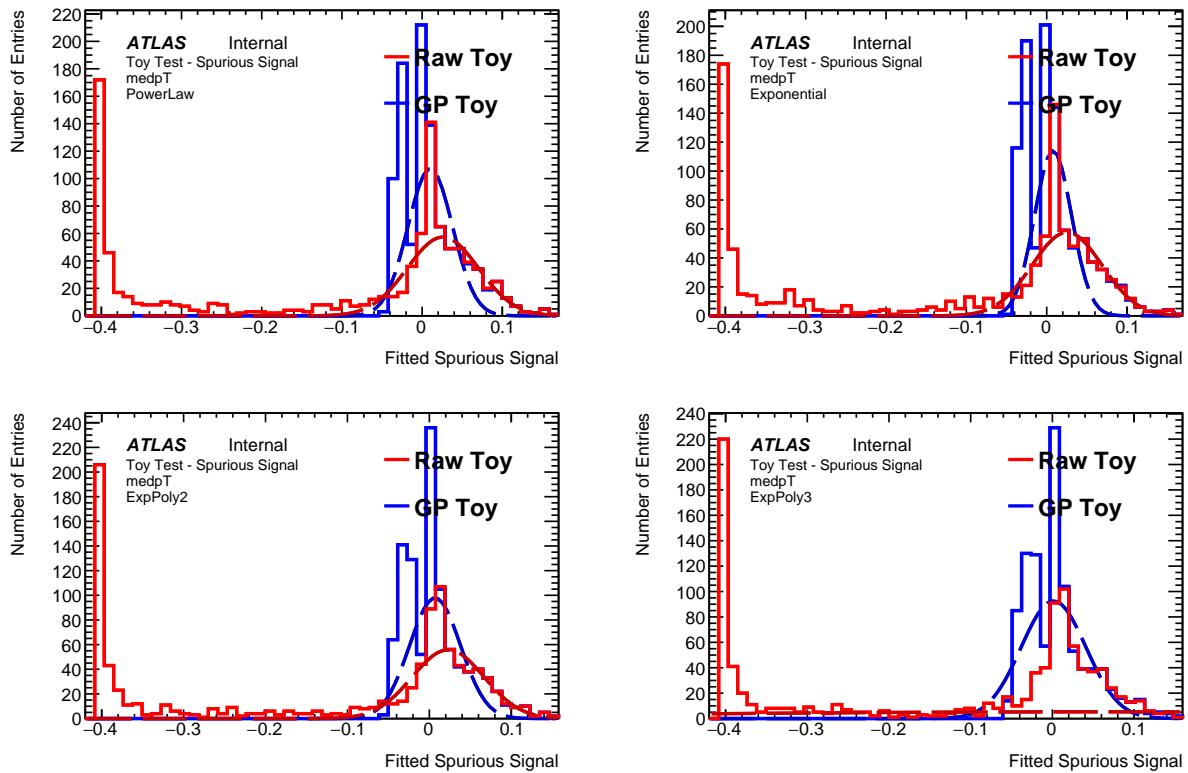


Figure A.139: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10 events.

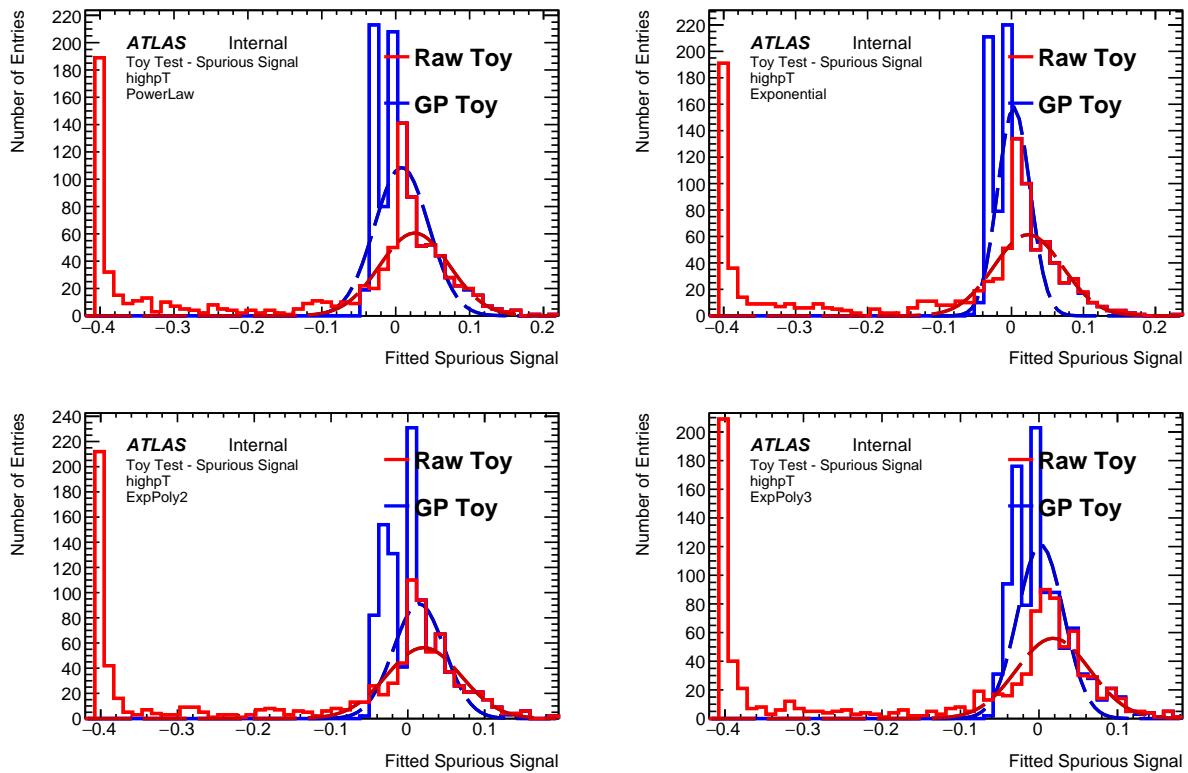


Figure A.140: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10 events.

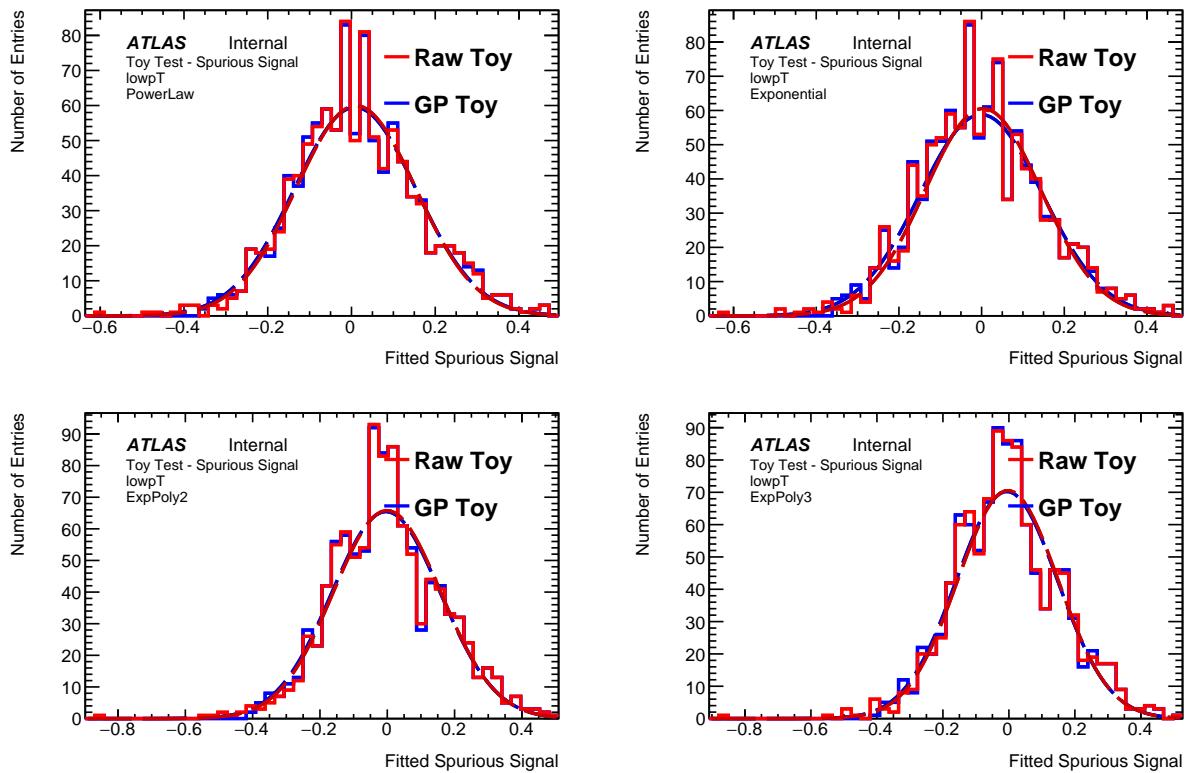


Figure A.141: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100 events.

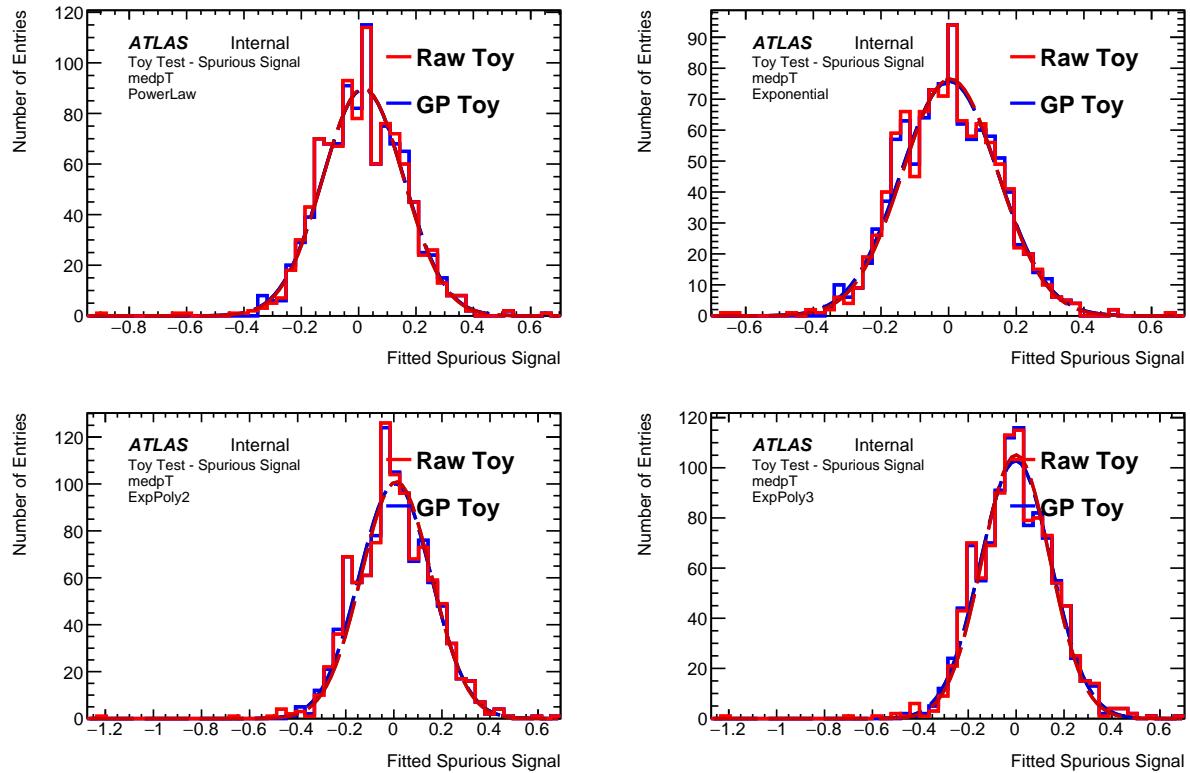


Figure A.142: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100 events.

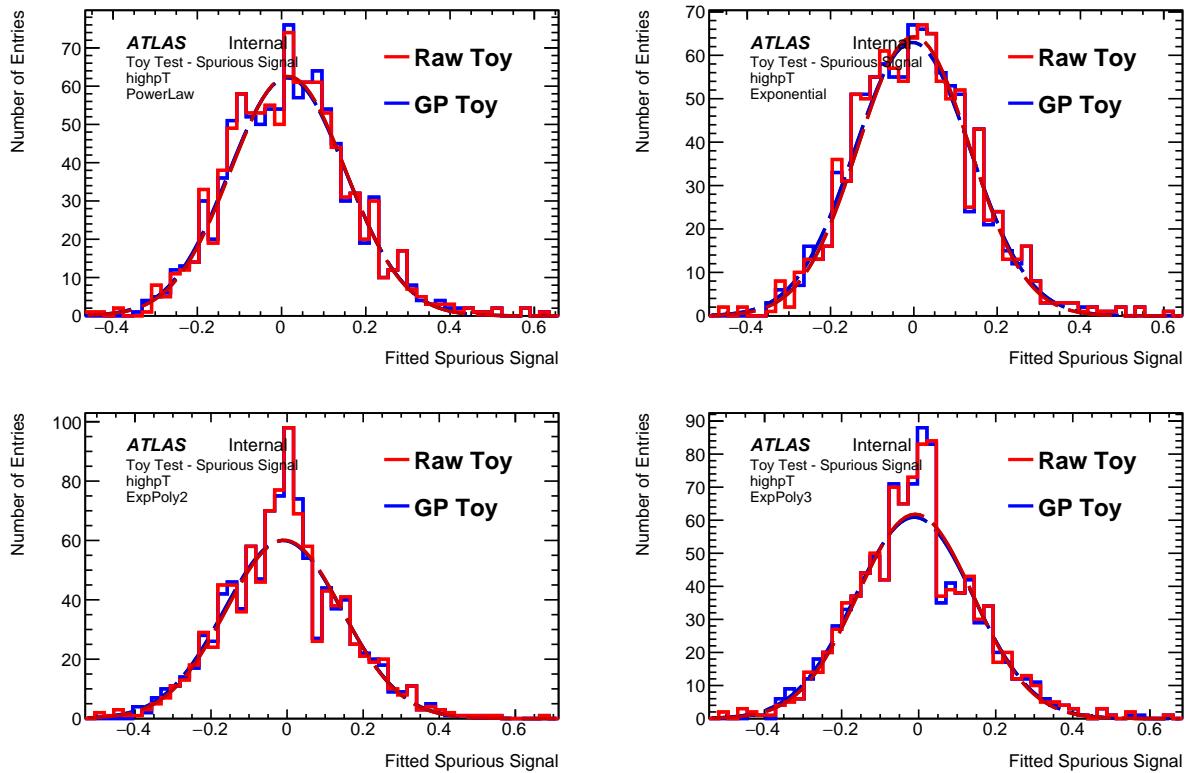


Figure A.143: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100 events.

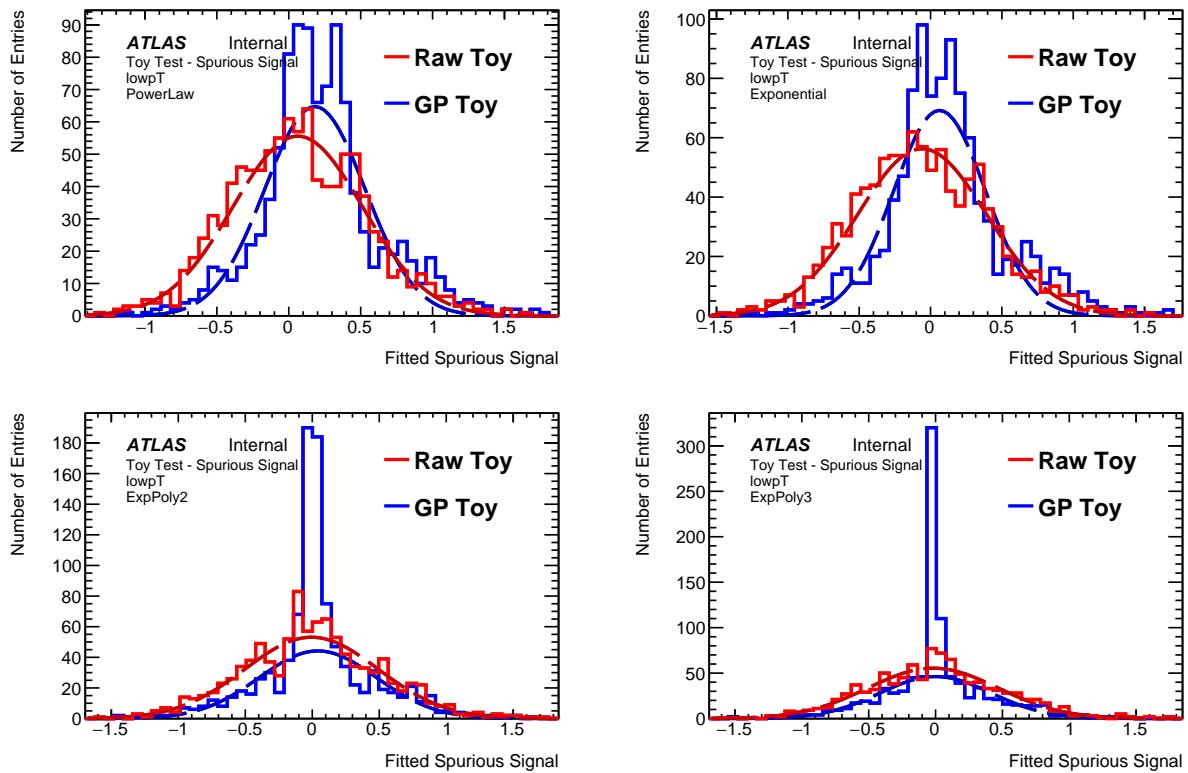


Figure A.144: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1000 events.

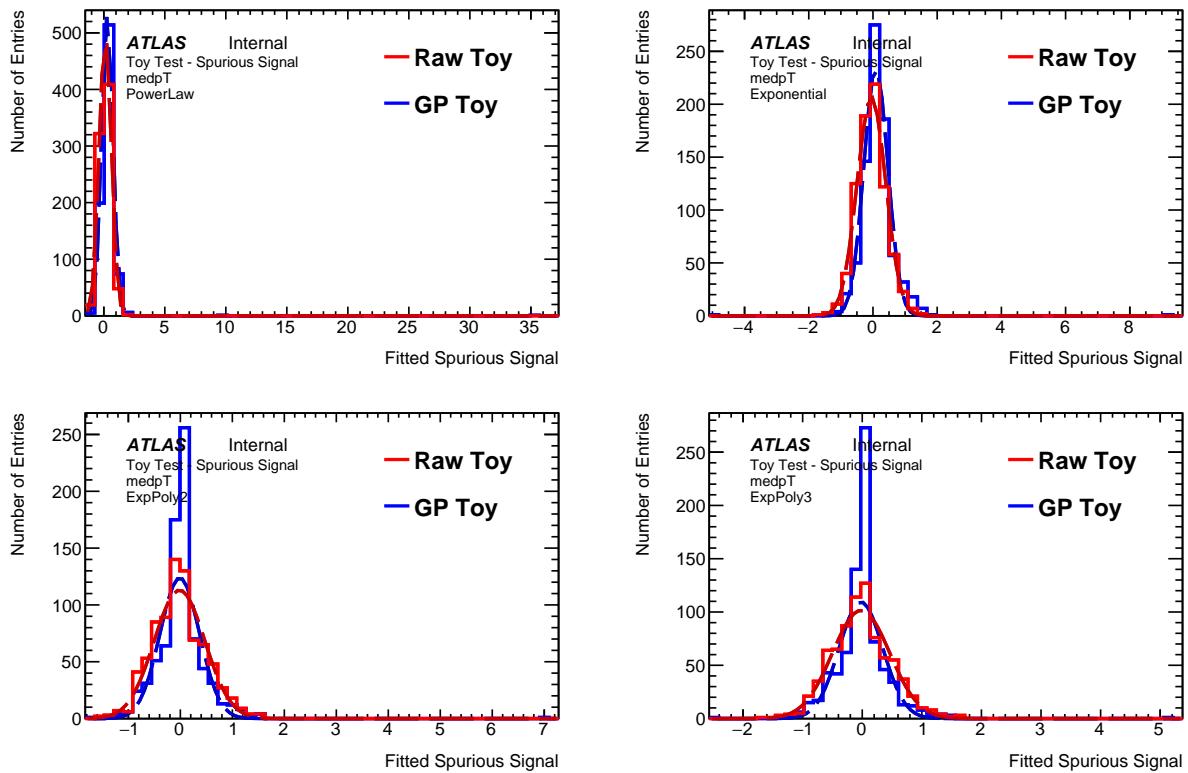


Figure A.145: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1000 events.

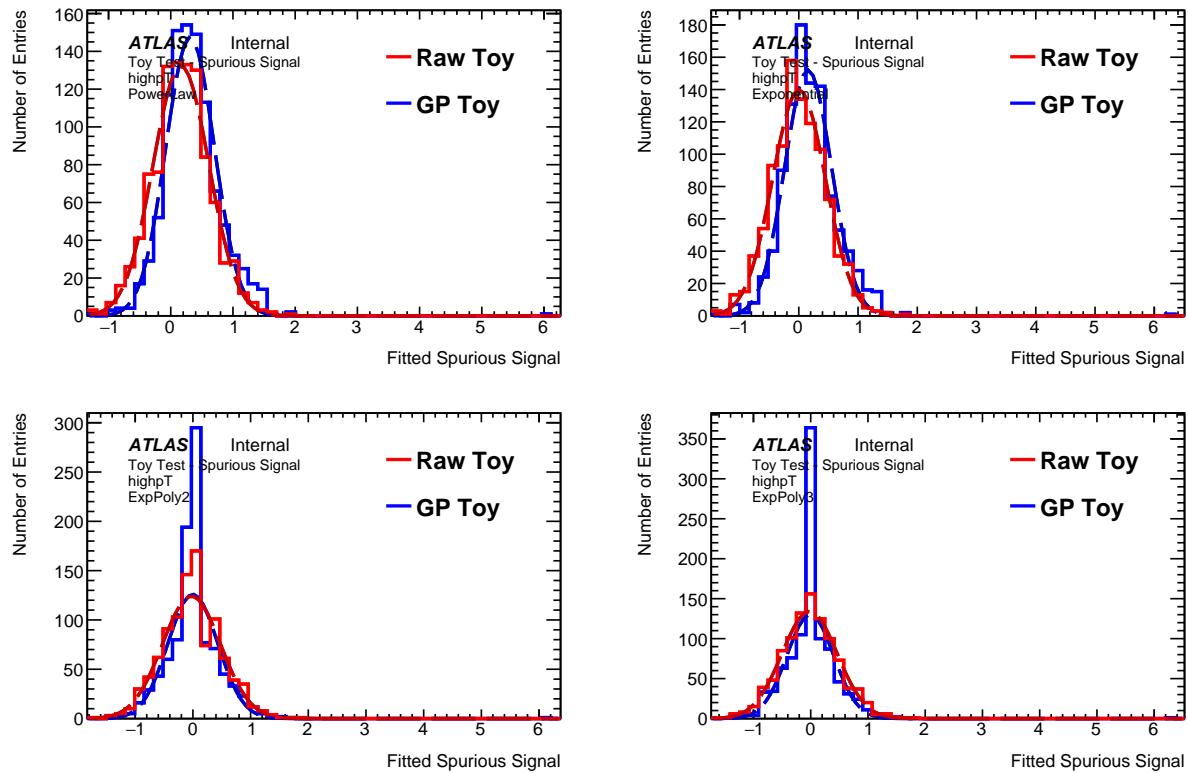


Figure A.146: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1000 events.

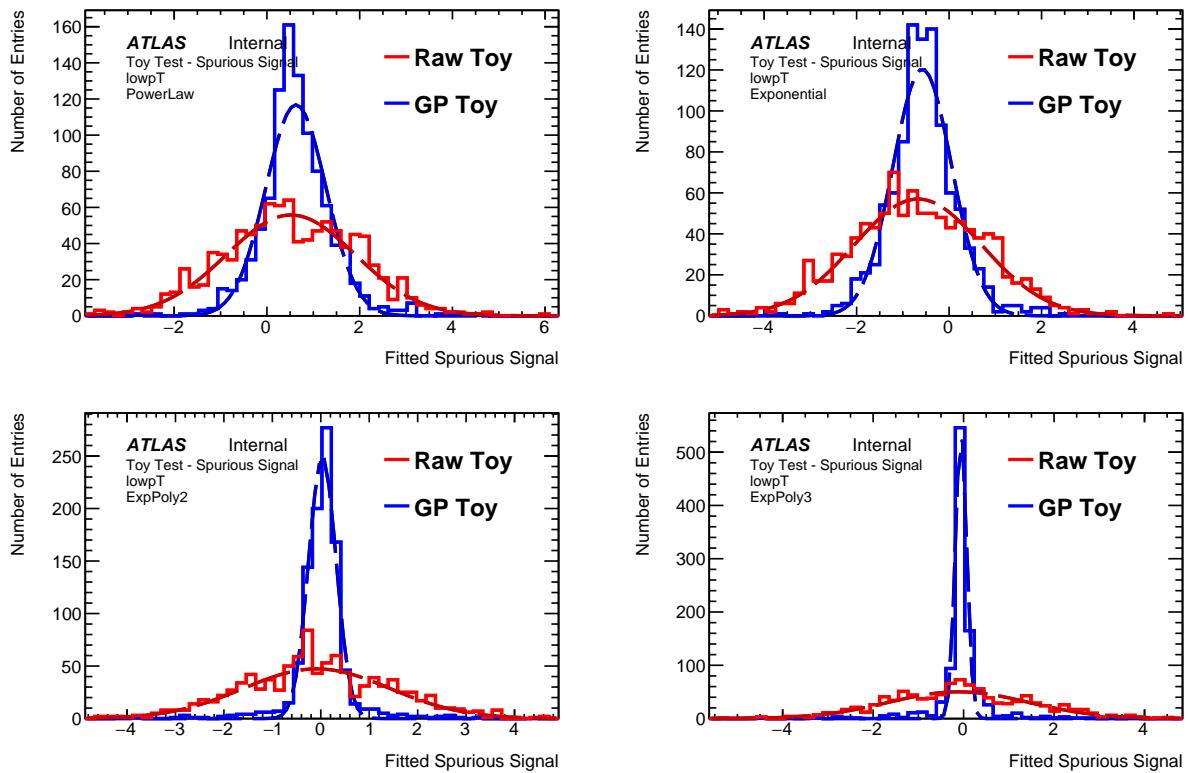


Figure A.147: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10k events.

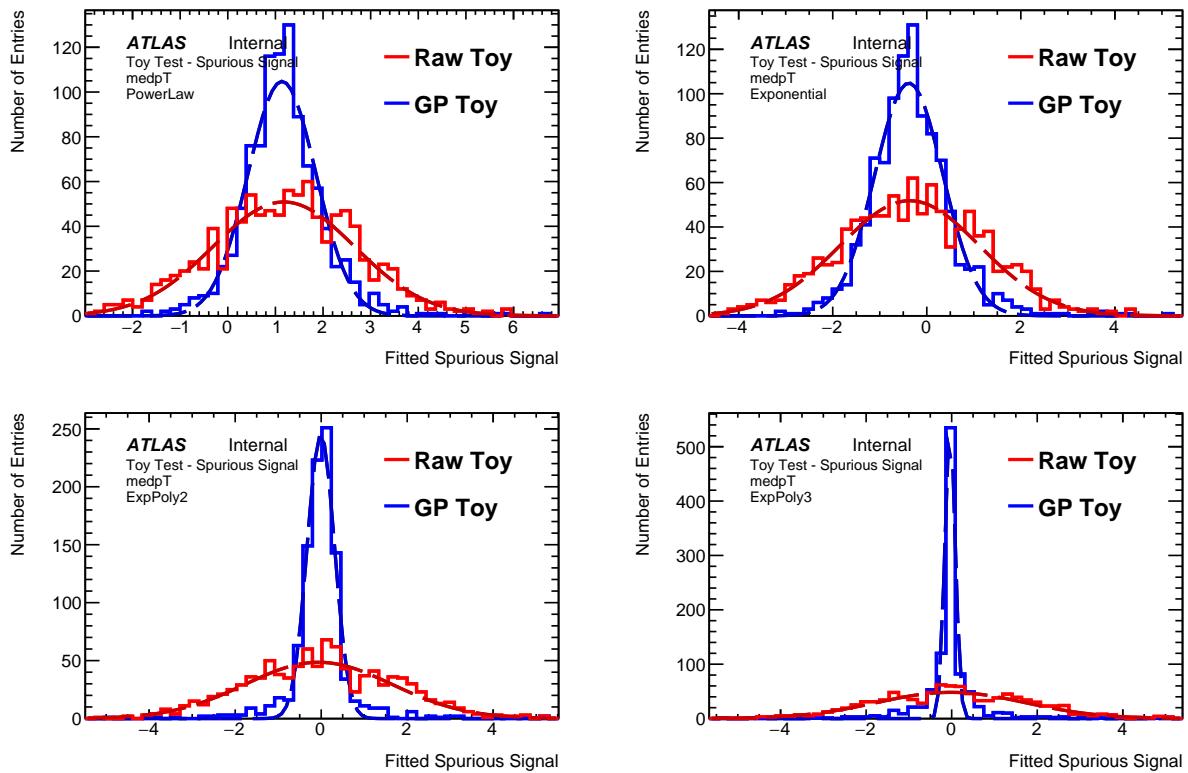


Figure A.148: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10k events.

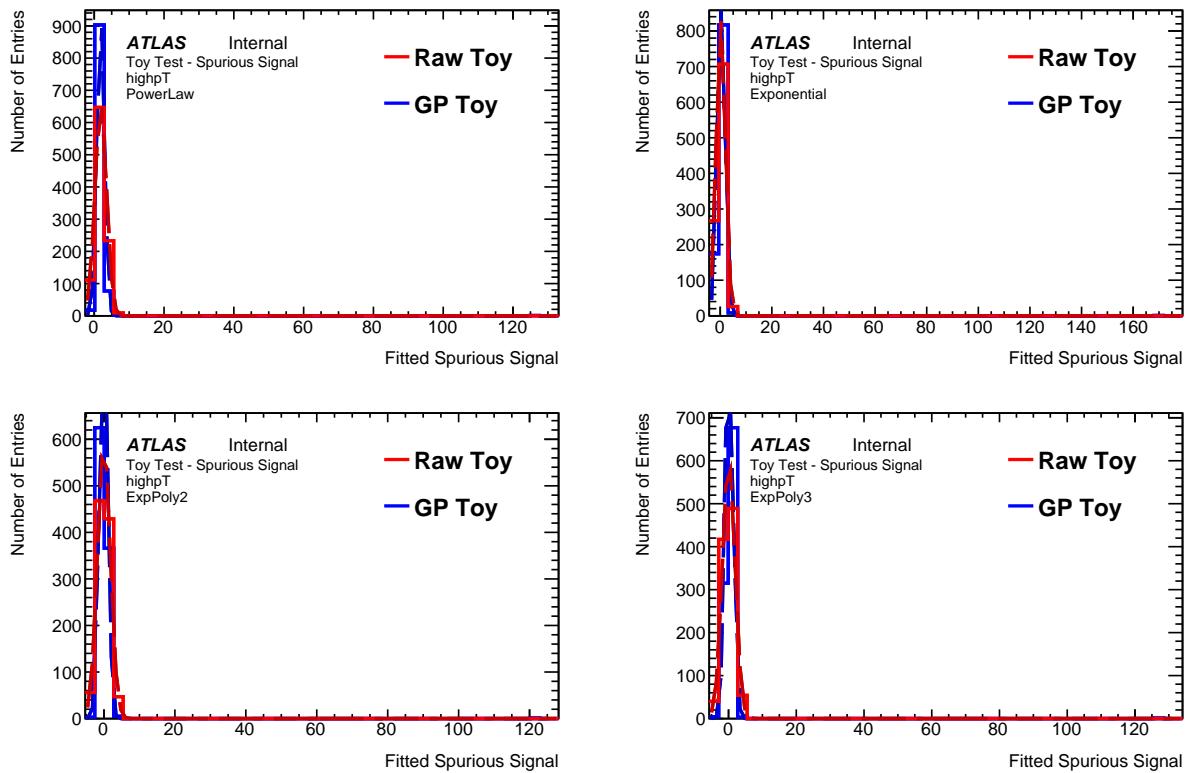


Figure A.149: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10k events.

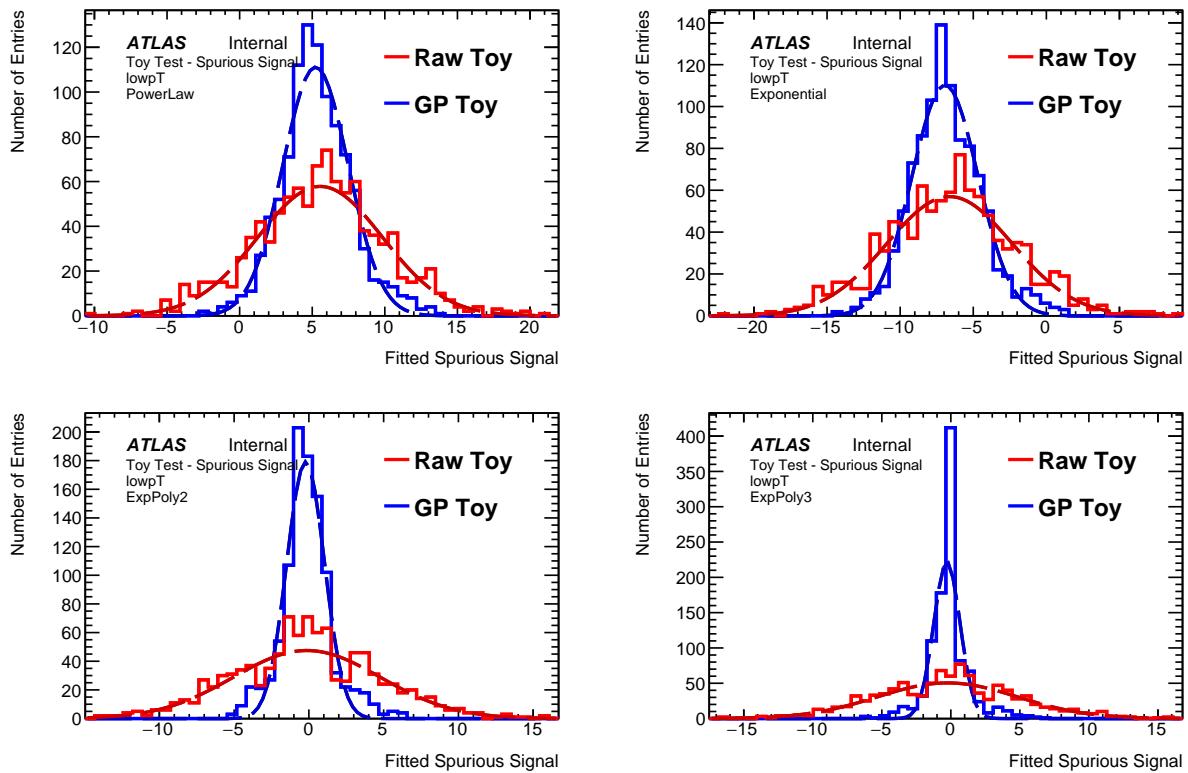


Figure A.150: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100k events.

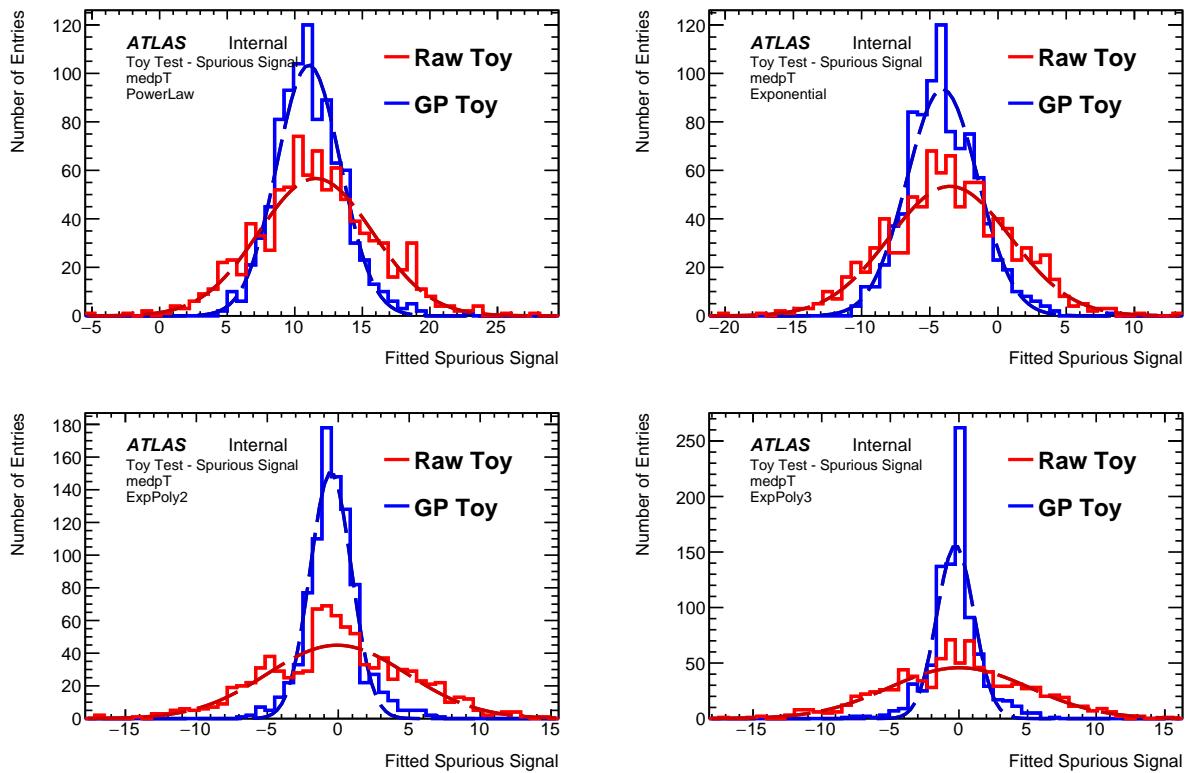


Figure A.151: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100k events.

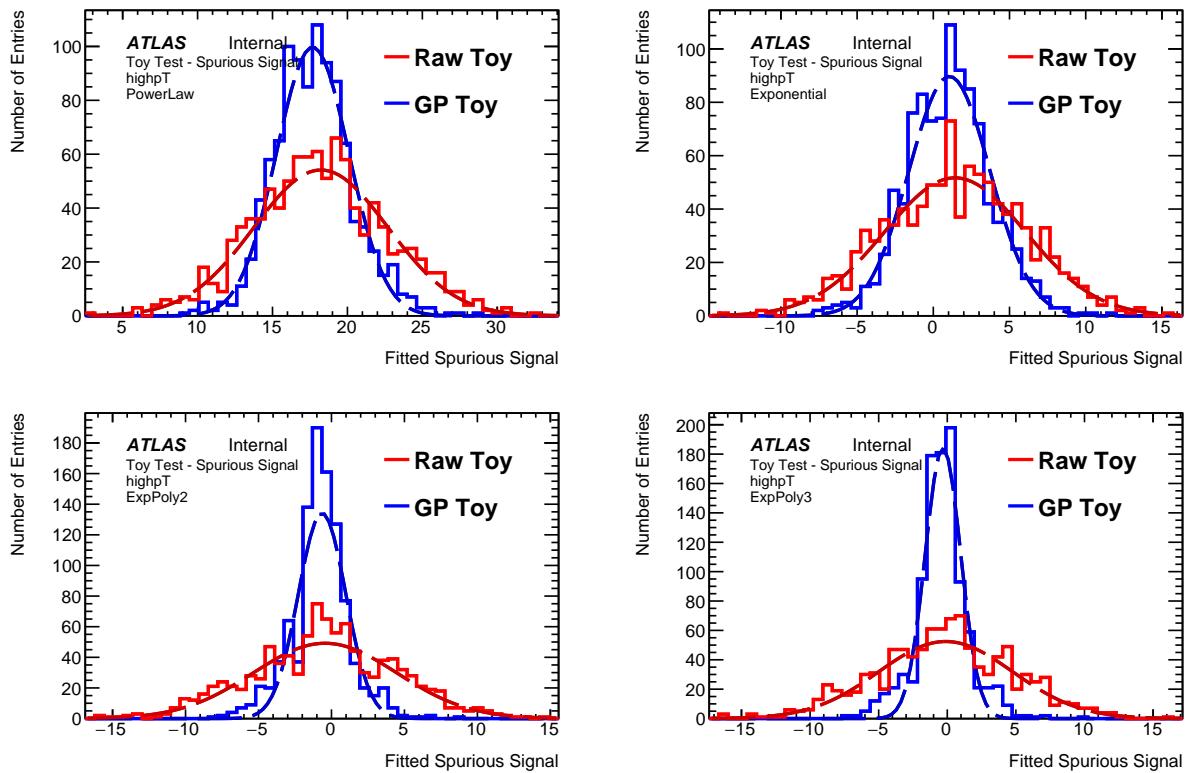


Figure A.152: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100k events.

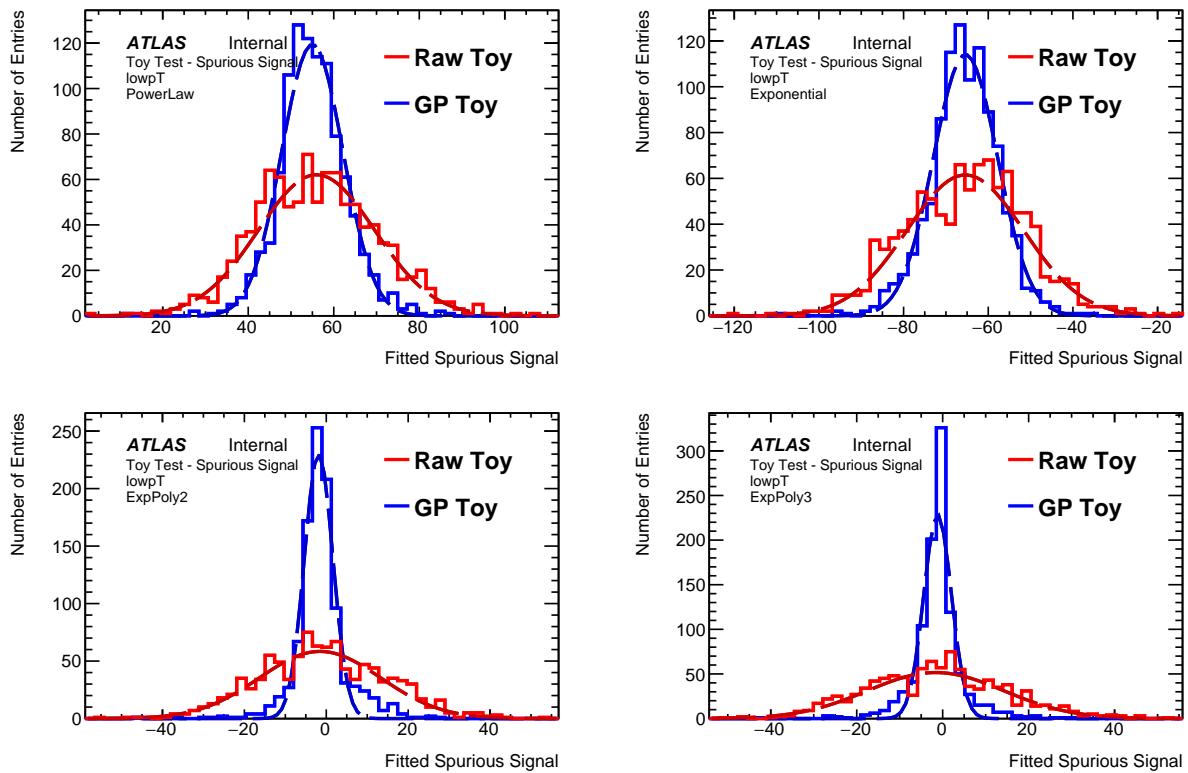


Figure A.153: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

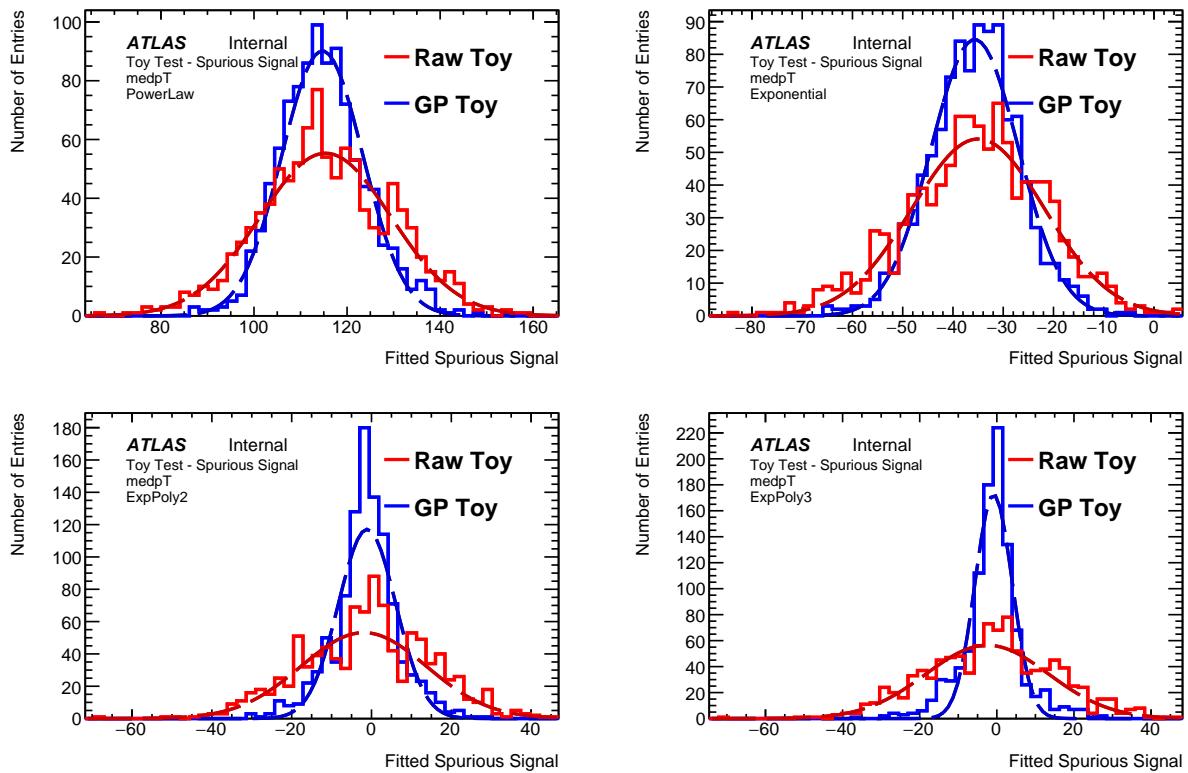


Figure A.154: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

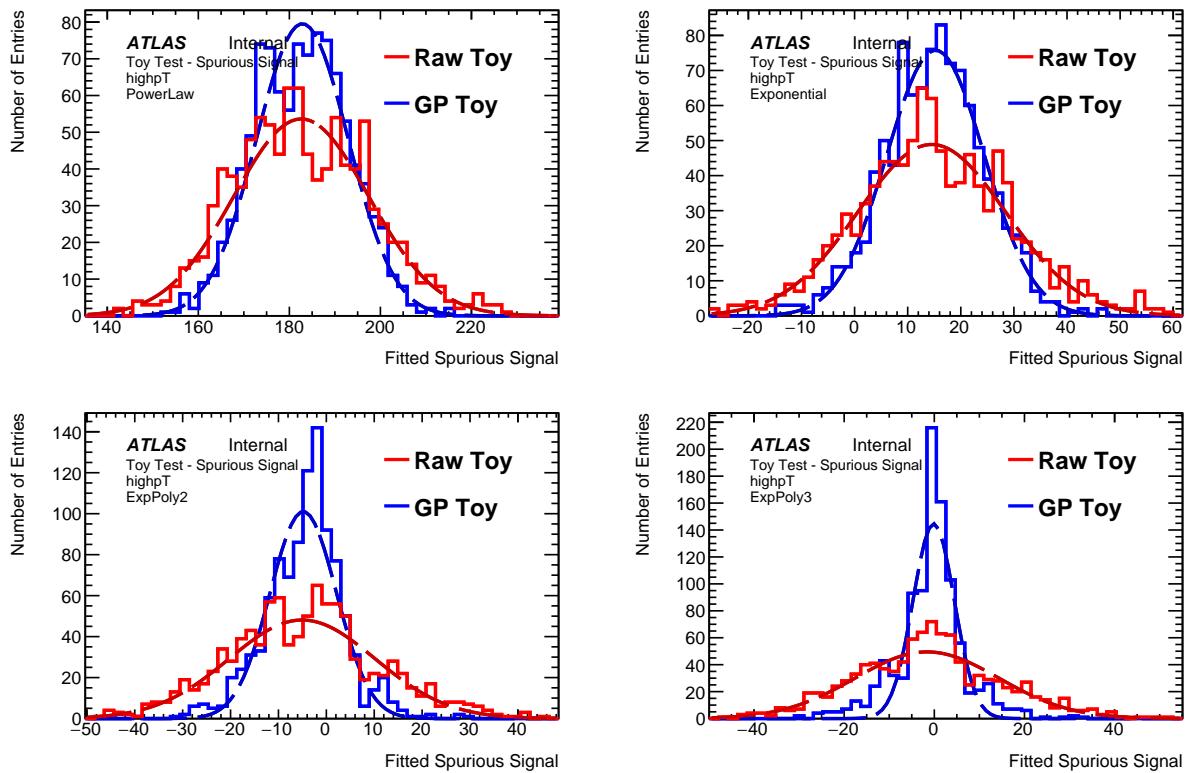


Figure A.155: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

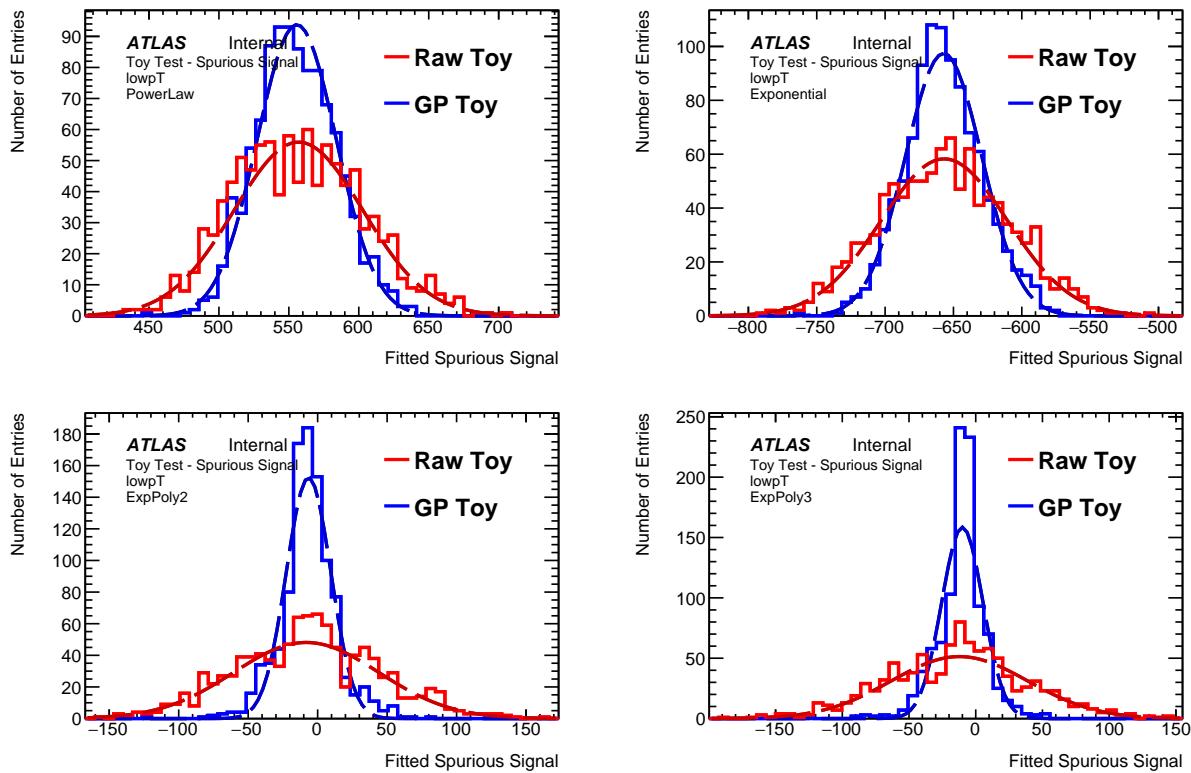


Figure A.156: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10,000,000 events.

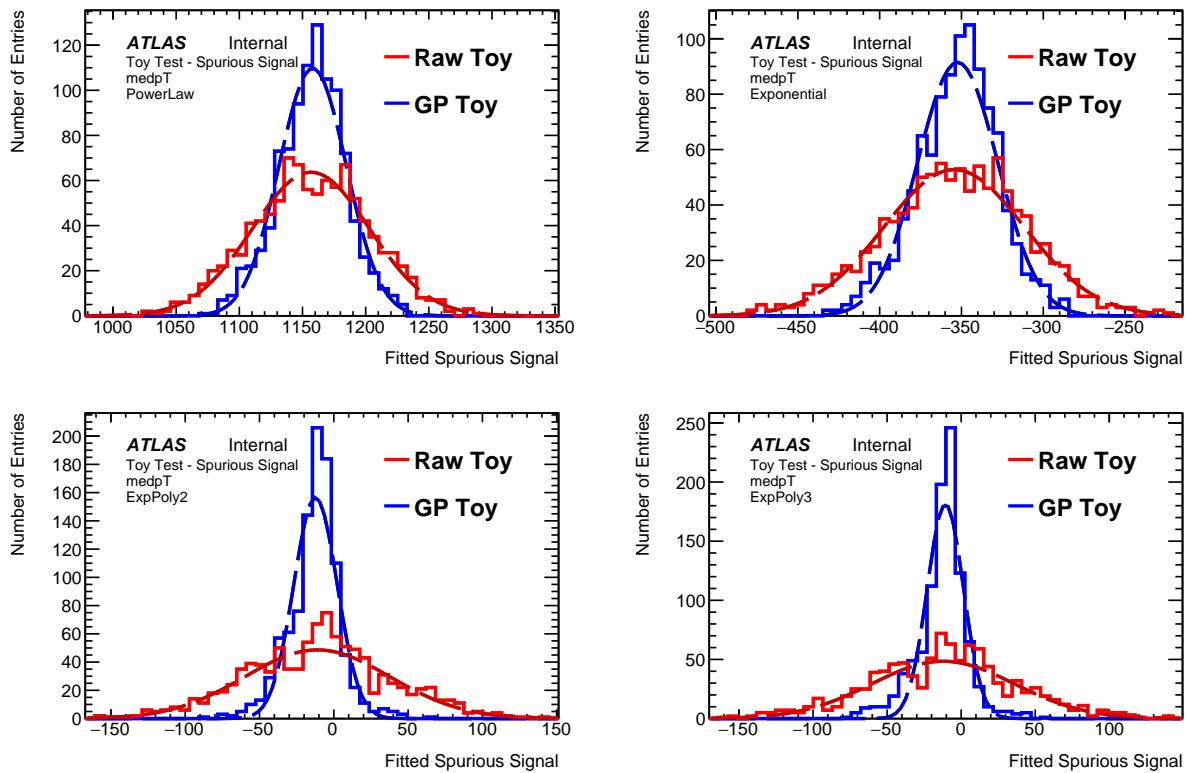


Figure A.157: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10,000,000 events.

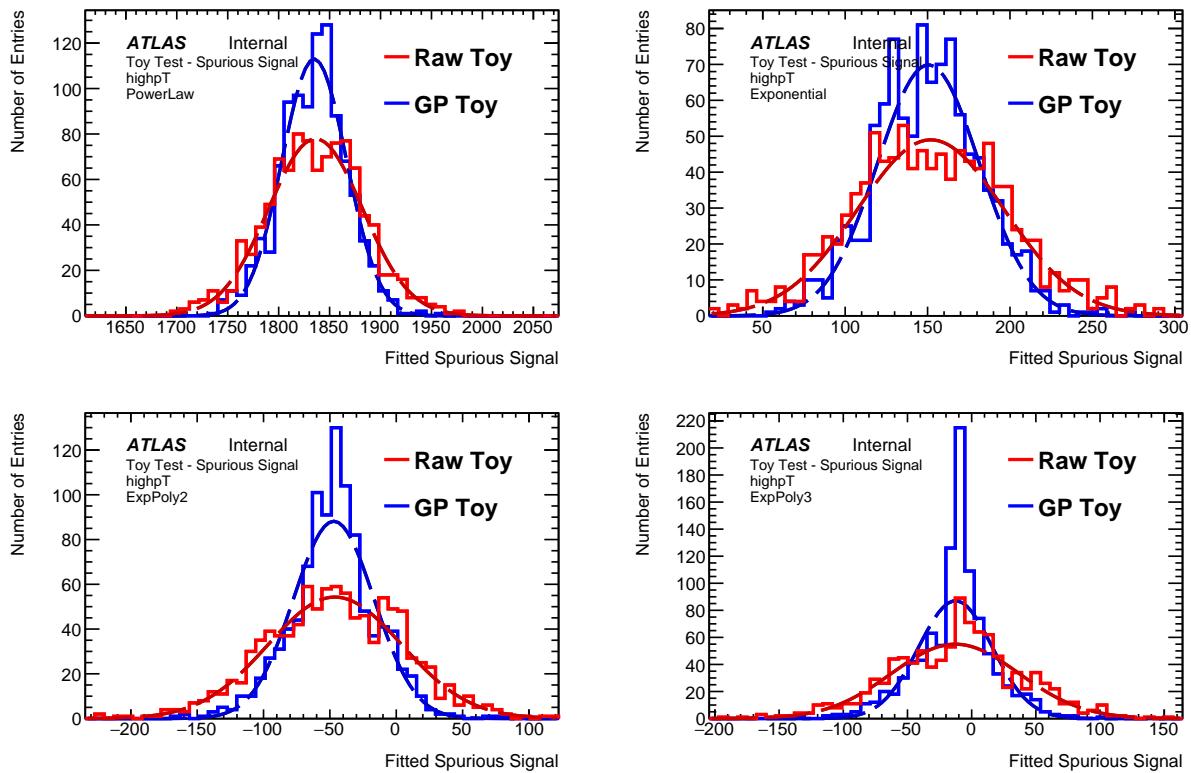


Figure A.158: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10,000,000 events.

Nominal	N_sig	Unweighted Events	Bkg events weighted	Mean SS_raw	Mean SS_GPR	GP-Raw: Bias, Mean	Sigma Toy	Sigma GPR	Bias/sigma(SS_raw)
Fit Function: Exp									
GeneratingFunction: Exp2									
low10	0	10^1	0.4	-0.11	0.01	-0.12	0.19	0.04	-0.64
low100	0	10^2	4	0.00	0.00	0.00	0.15	0.15	0.00
low1k	0	10^3	4×10^1	-0.06	0.11	-0.17	0.47	0.41	-0.36
low10k	0	10^4	4×10^2	-0.68	-0.54	-0.13	1.43	0.78	-0.09
low100k	0	10^5	4×10^3	-6.57	-6.81	0.24	4.47	2.43	0.05
low1M	0	10^6	4×10^4	-65.40	-65.47	0.07	14.13	8.16	0.01
low10M	0	10^7	4×10^5	-656.74	-656.98	0.24	46.63	28.11	0.01
Fit Function: Exp2									
Generating Function: Exp3									
med10	0	10^1	0.4	-0.13	0.01	-0.13	0.19	0.04	-0.69
med100	0	10^2	4	-0.00	0.00	-0.00	0.16	0.16	-0.02
med1k	0	10^3	4×10^1	-0.02	0.02	-0.04	0.51	0.48	-0.07
med10k	0	10^4	4×10^2	0.01	-0.01	0.02	1.70	0.71	0.01
med100k	0	10^5	4×10^3	-0.09	-0.38	0.29	4.91	1.82	0.06
med1M	0	10^6	4×10^4	-1.28	-1.44	0.16	16.09	8.12	0.01
med10M	0	10^7	4×10^5	-10.49	-12.78	2.29	48.65	17.09	0.05
Fit Function: Exp2									
Generating Function: Exp3									
high10	0	10^1	0.4	-0.12	0.01	-0.13	0.19	0.04	-0.66
high100	0	10^2	4	-0.00	-0.00	-0.00	0.16	0.16	-0.00
high1k	0	10^3	4×10^1	-0.01	0.02	-0.03	0.50	0.46	-0.06
high10k	0	10^4	4×10^2	0.05	0.08	-0.03	1.62	3.91	-0.02
high100k	0	10^5	4×10^3	-0.53	-0.79	0.25	4.94	1.99	0.05
high1M	0	10^6	4×10^4	-4.75	-4.40	-0.36	15.74	8.20	-0.02
high10M	0	10^7	4×10^5	-45.75	-47.00	1.25	51.22	31.71	0.02

Table A.13: Spurious signal means and widths for the three test functional-form distributions for a range of different template statistics.

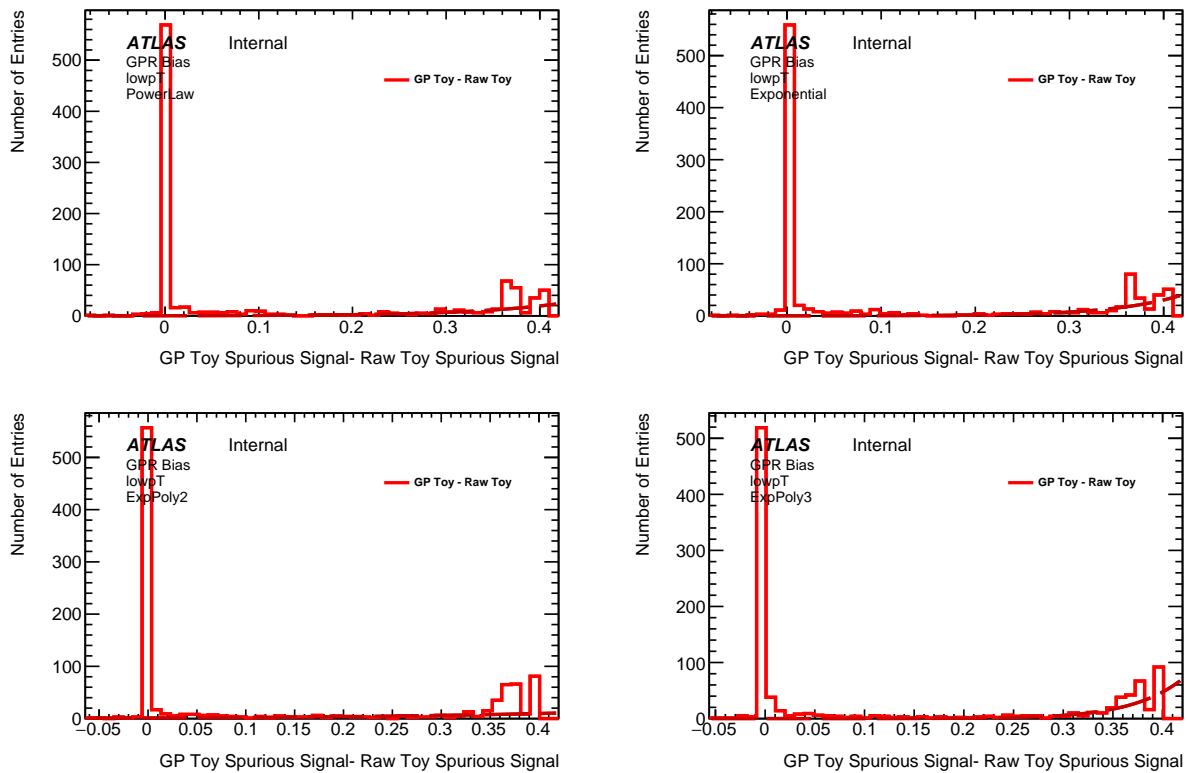


Figure A.159: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10 events.

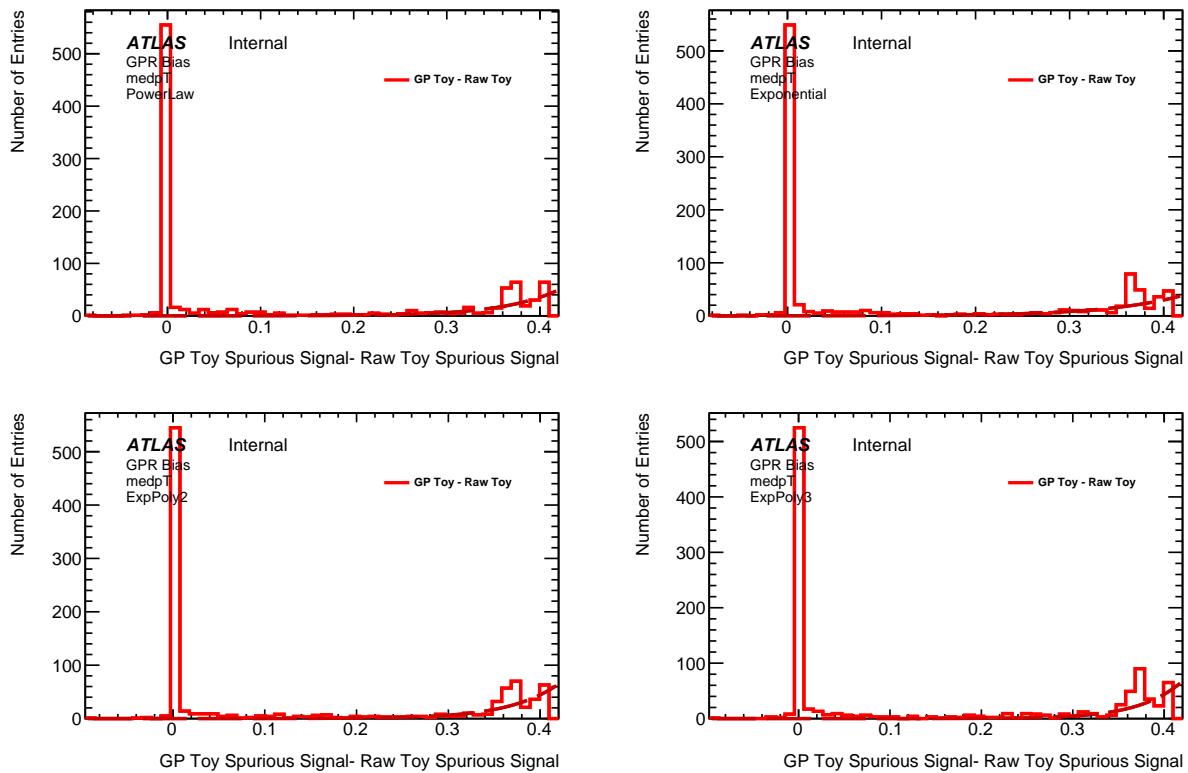


Figure A.160: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10 events.

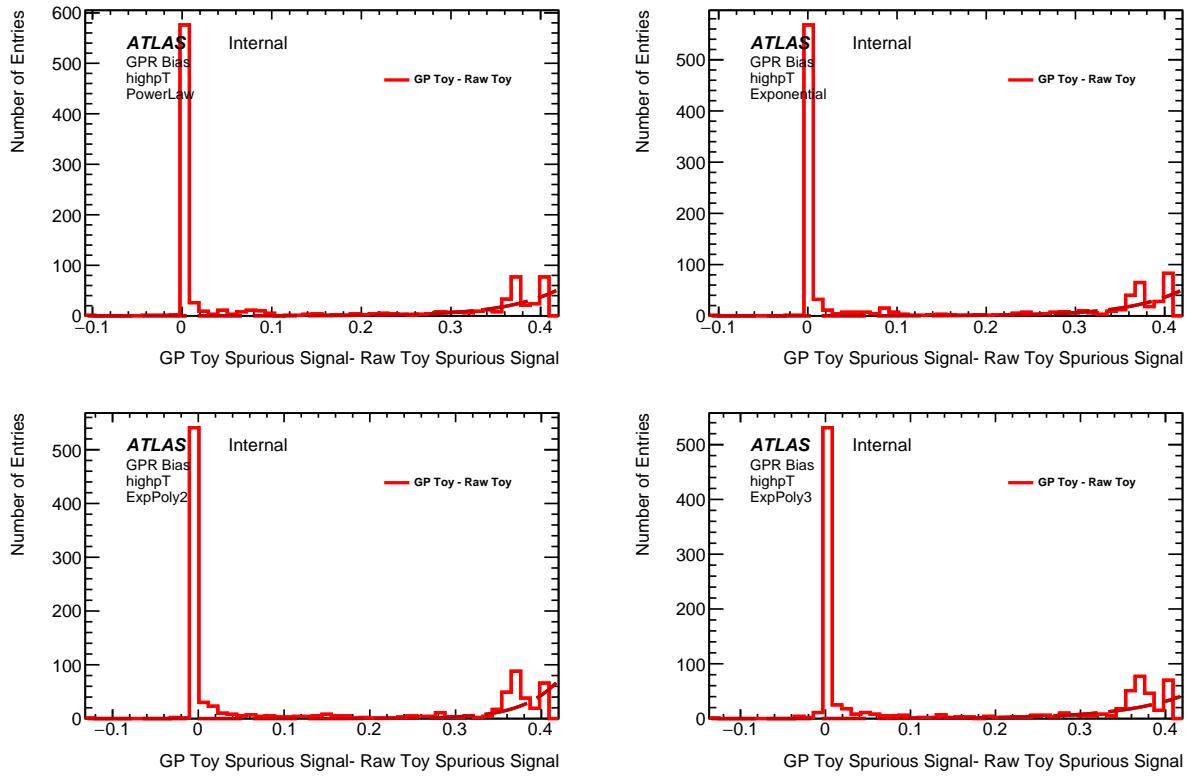


Figure A.161: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10 events.

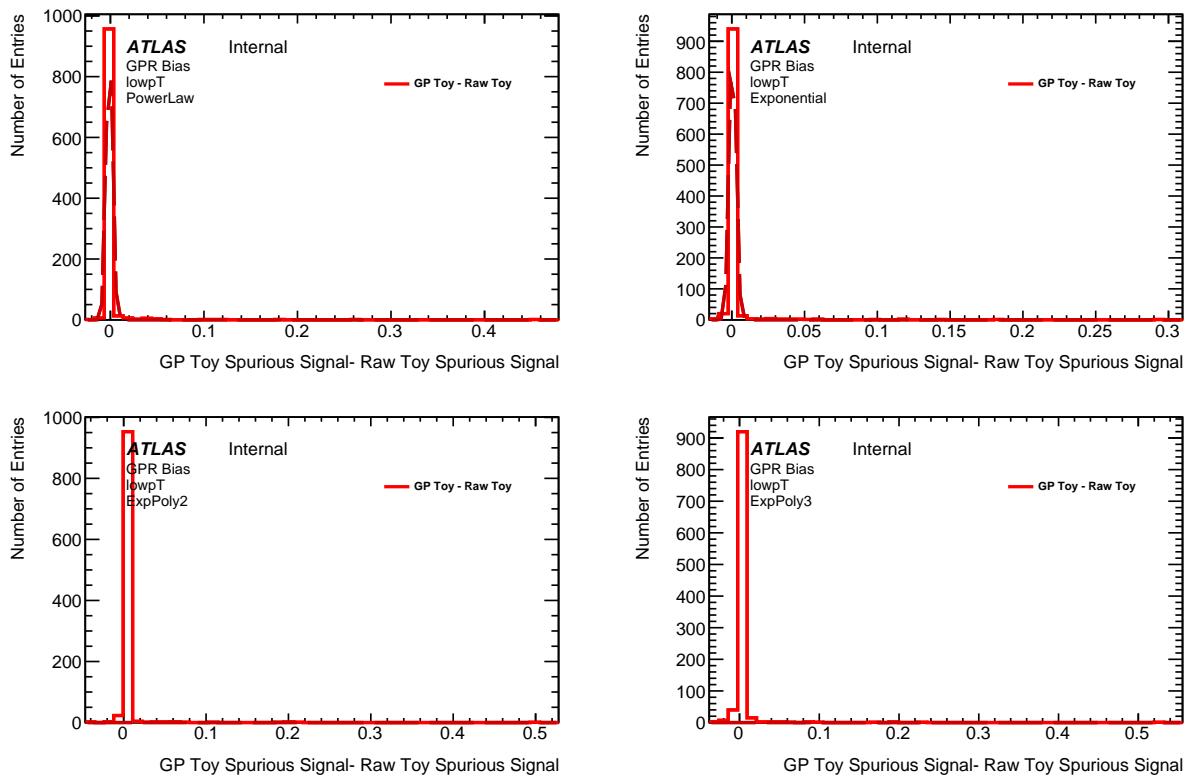


Figure A.162: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100 events.

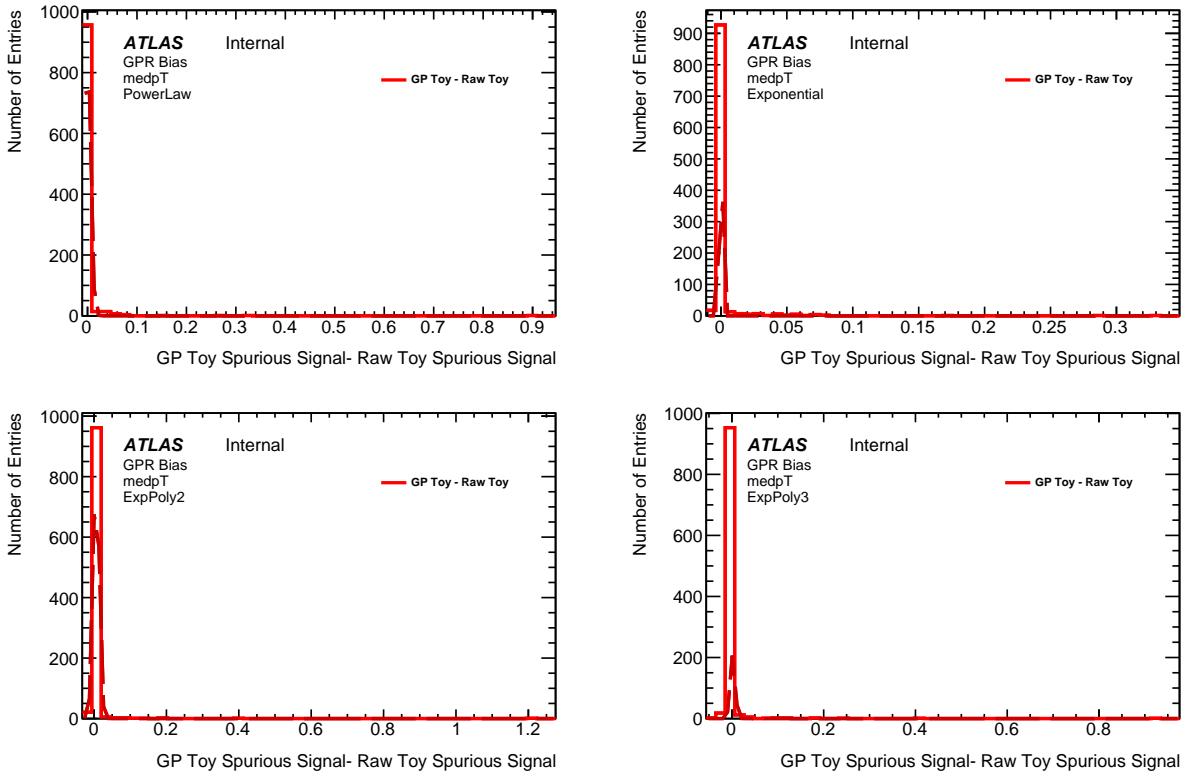


Figure A.163: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100 events.

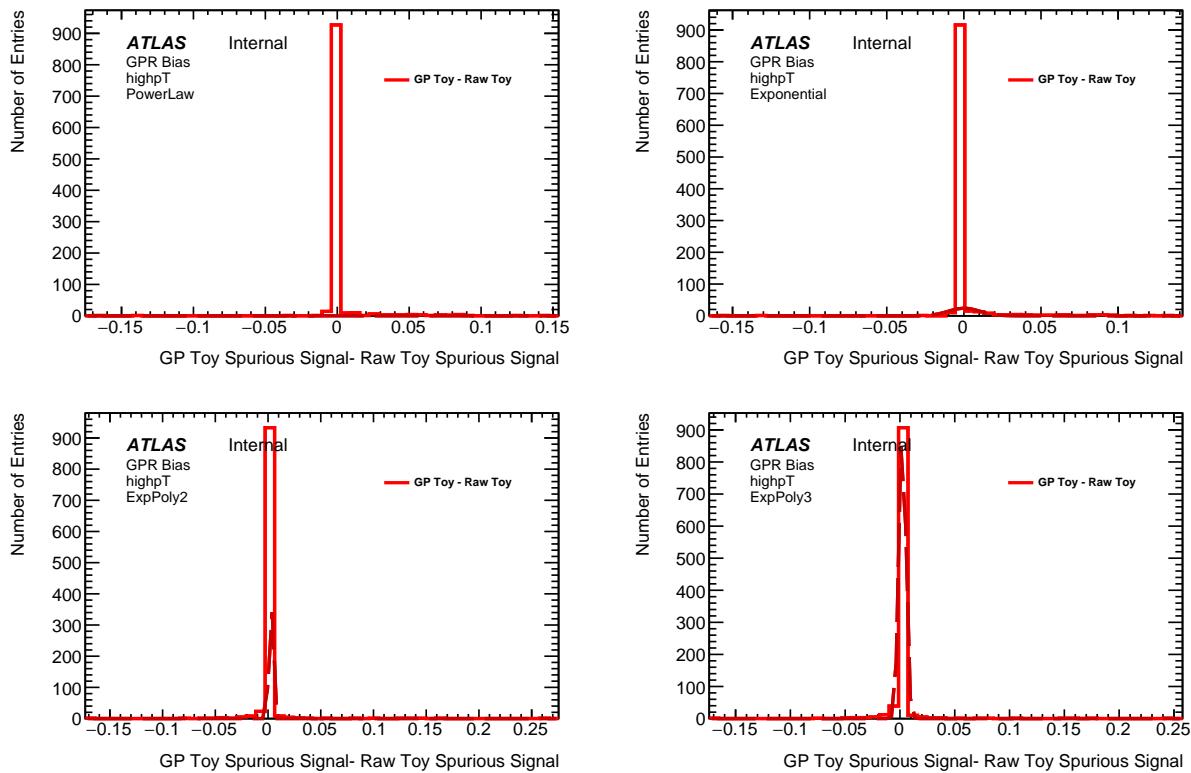


Figure A.164: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100 events.

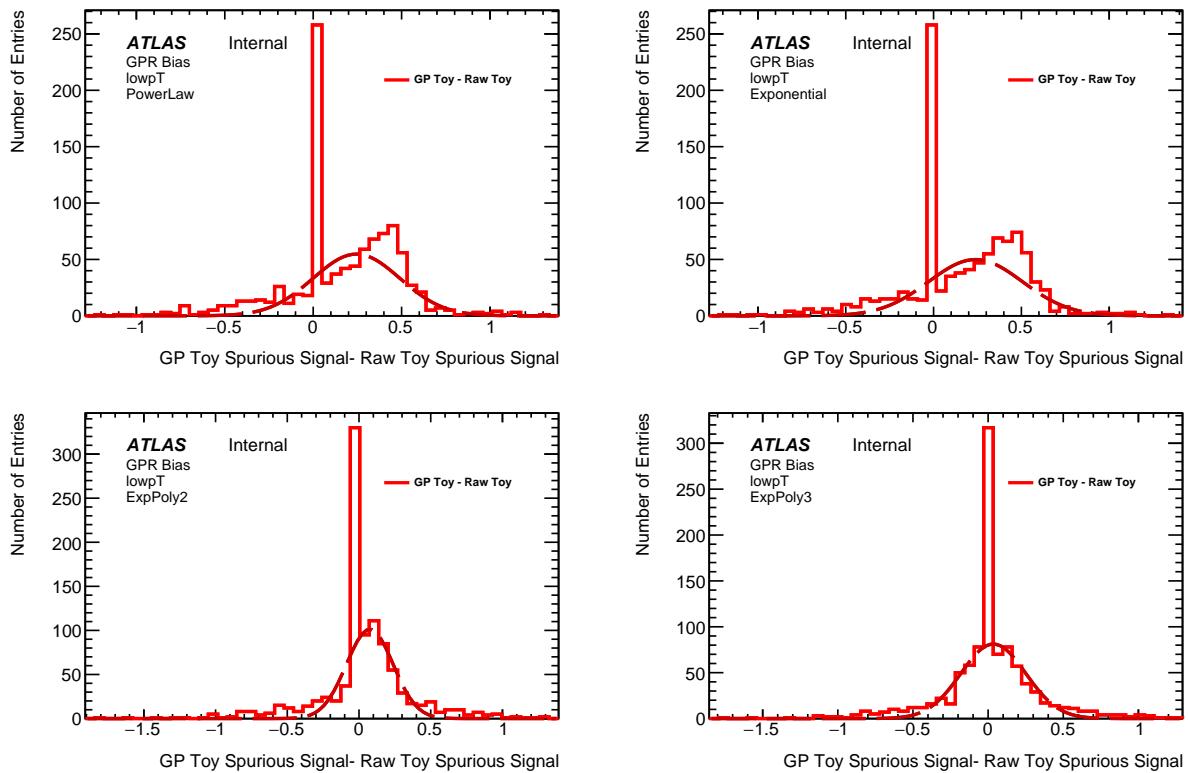


Figure A.165: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1000 events.

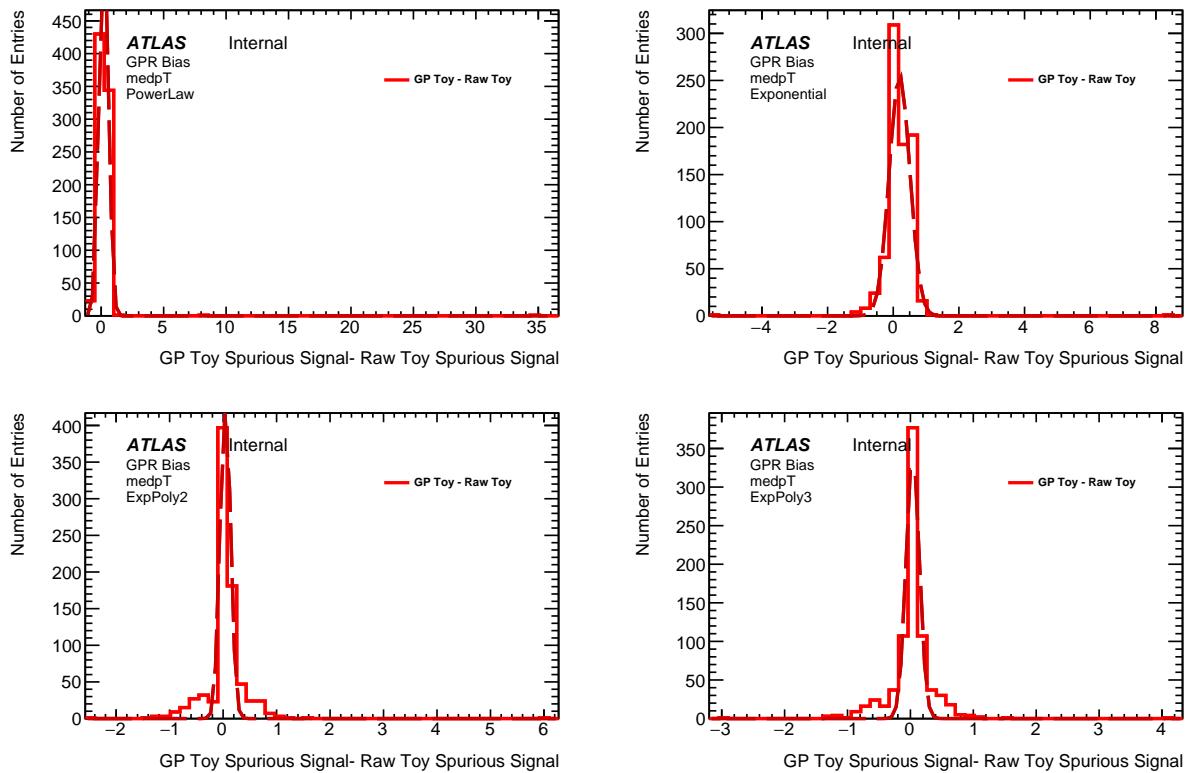


Figure A.166: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1000 events.

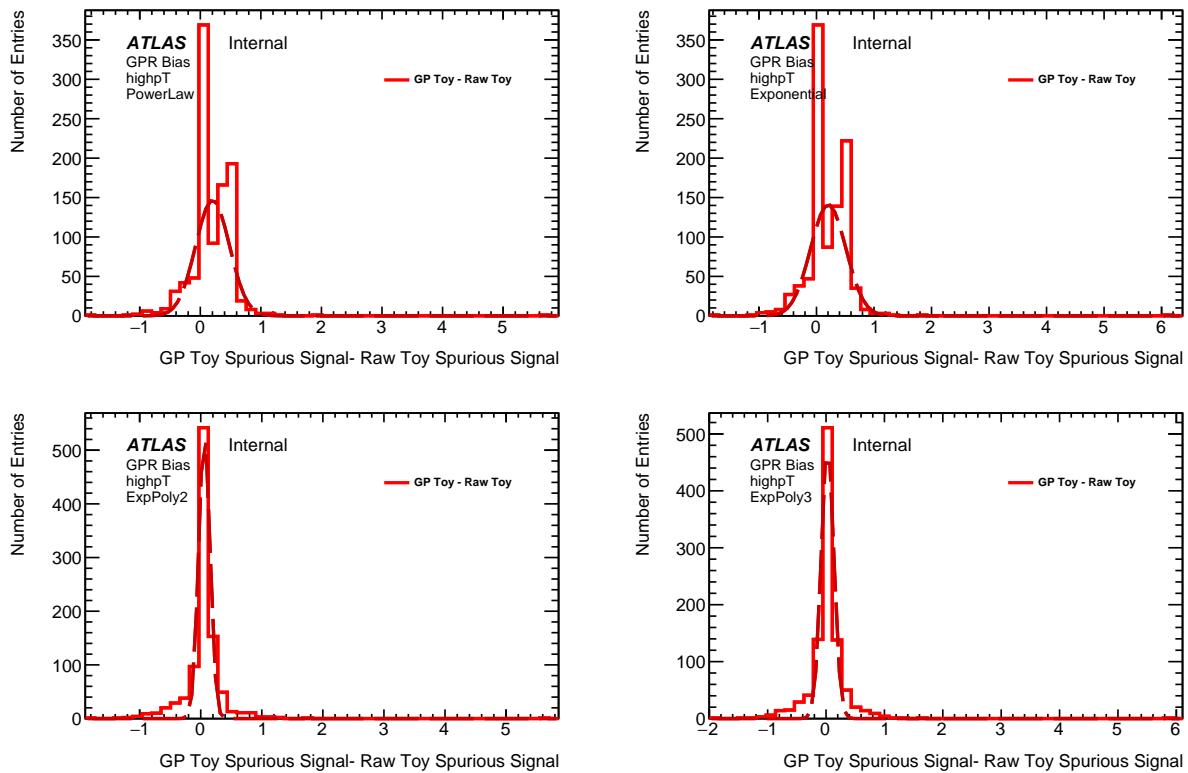


Figure A.167: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1000 events.

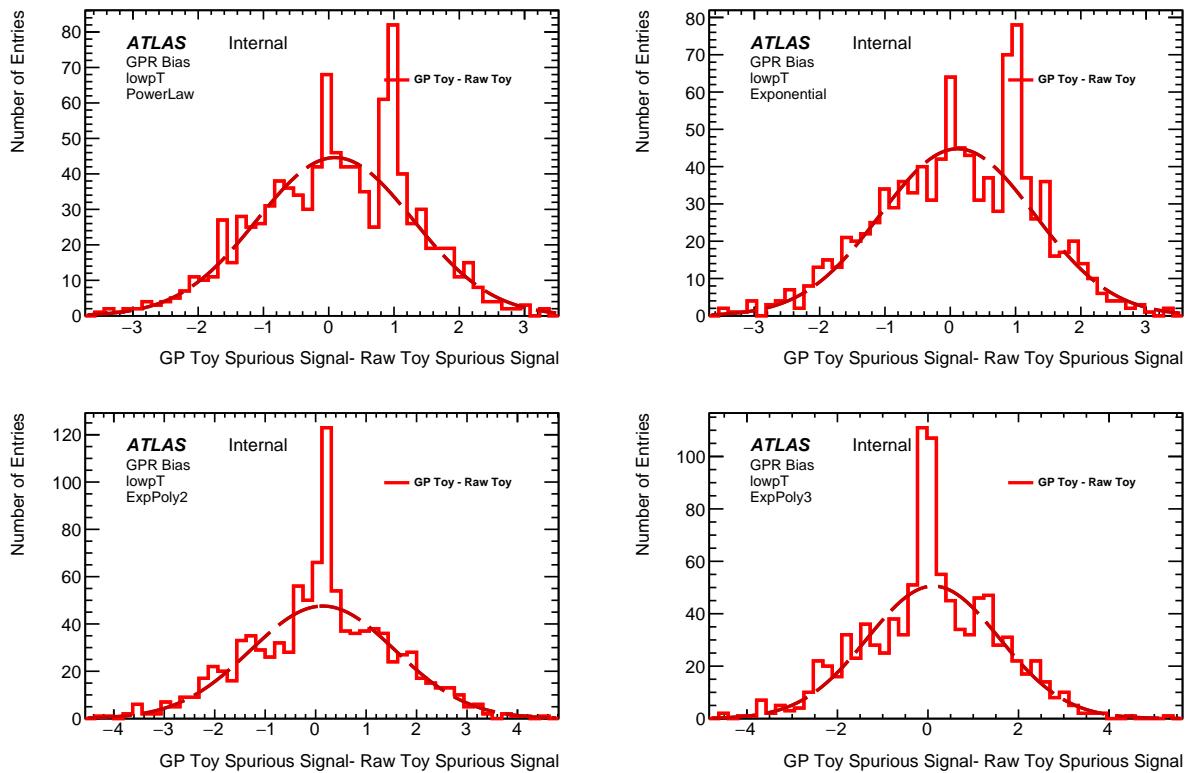


Figure A.168: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10k events.

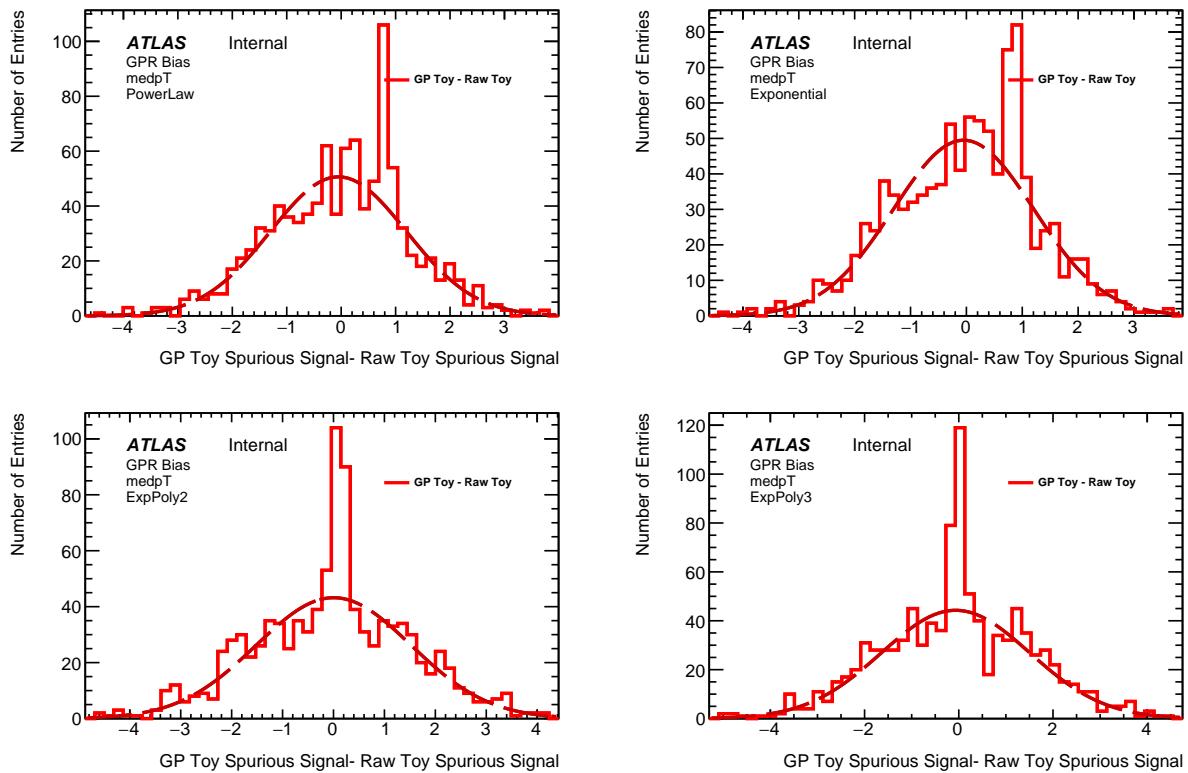


Figure A.169: The per-toy bias ($\text{GP Spurious Signal} - \text{Raw Spurious Signal}$), using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10k events.

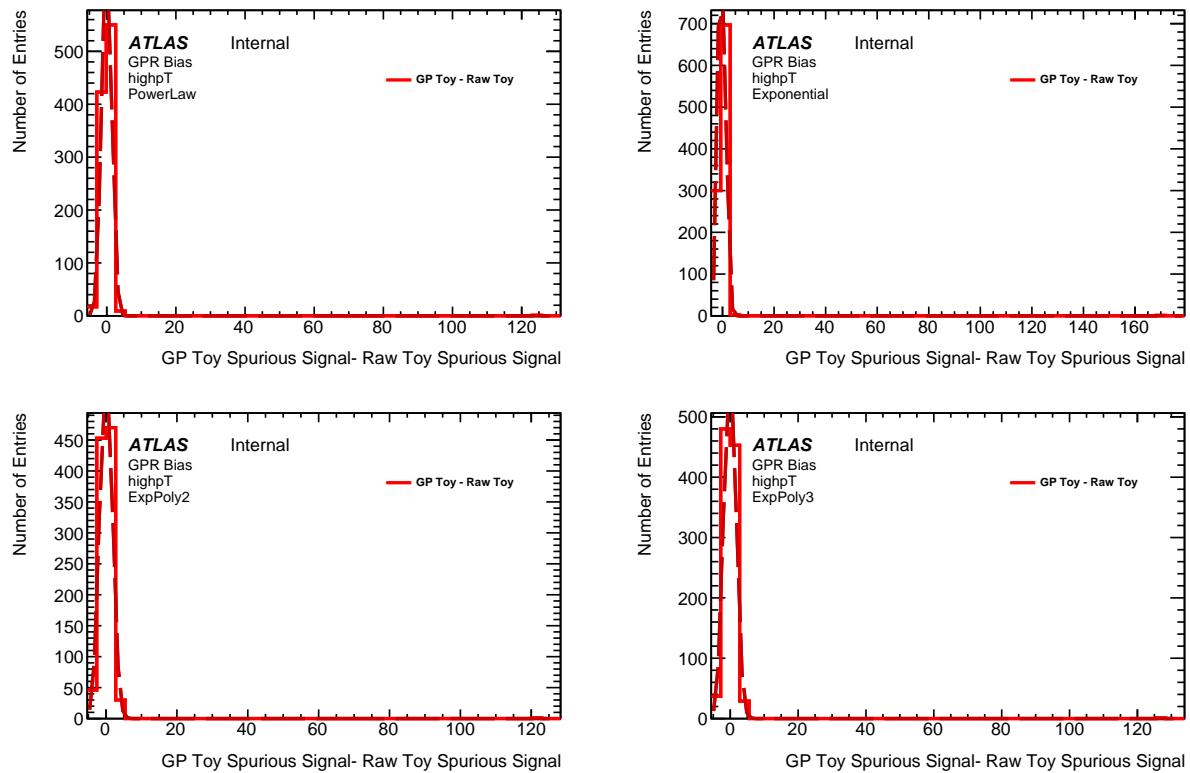


Figure A.170: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10k events.

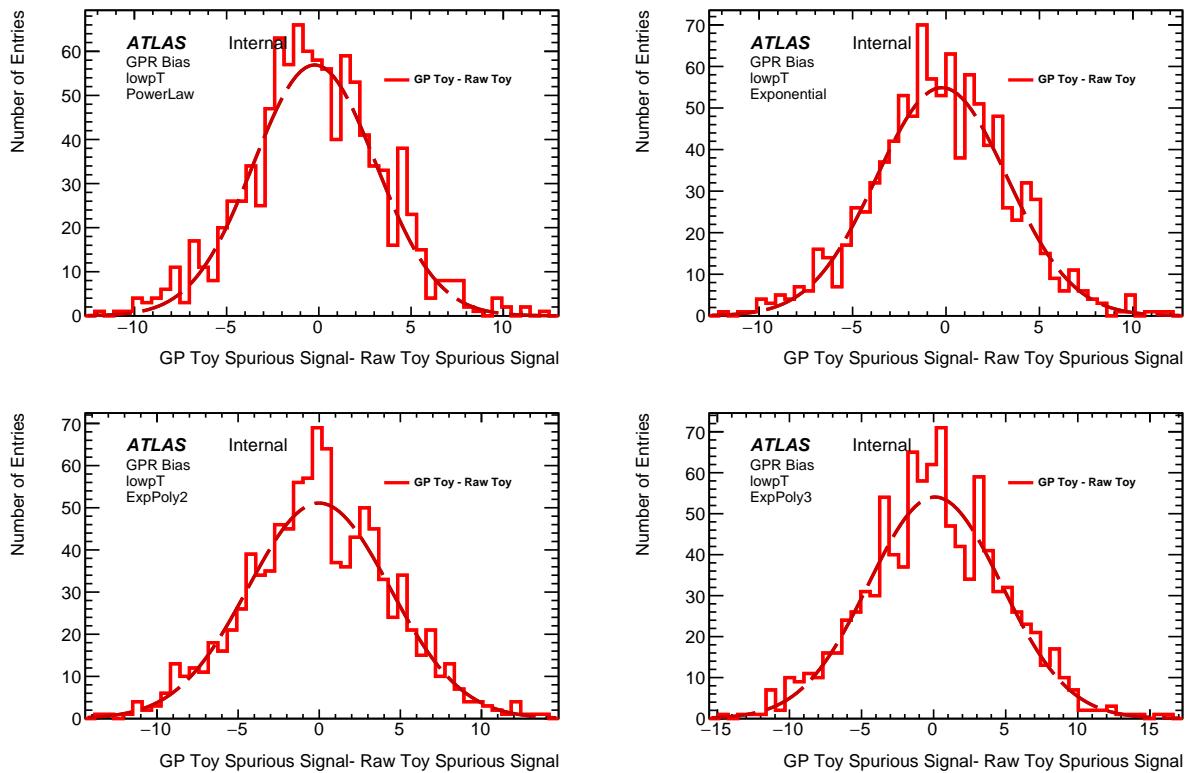


Figure A.171: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100k events.

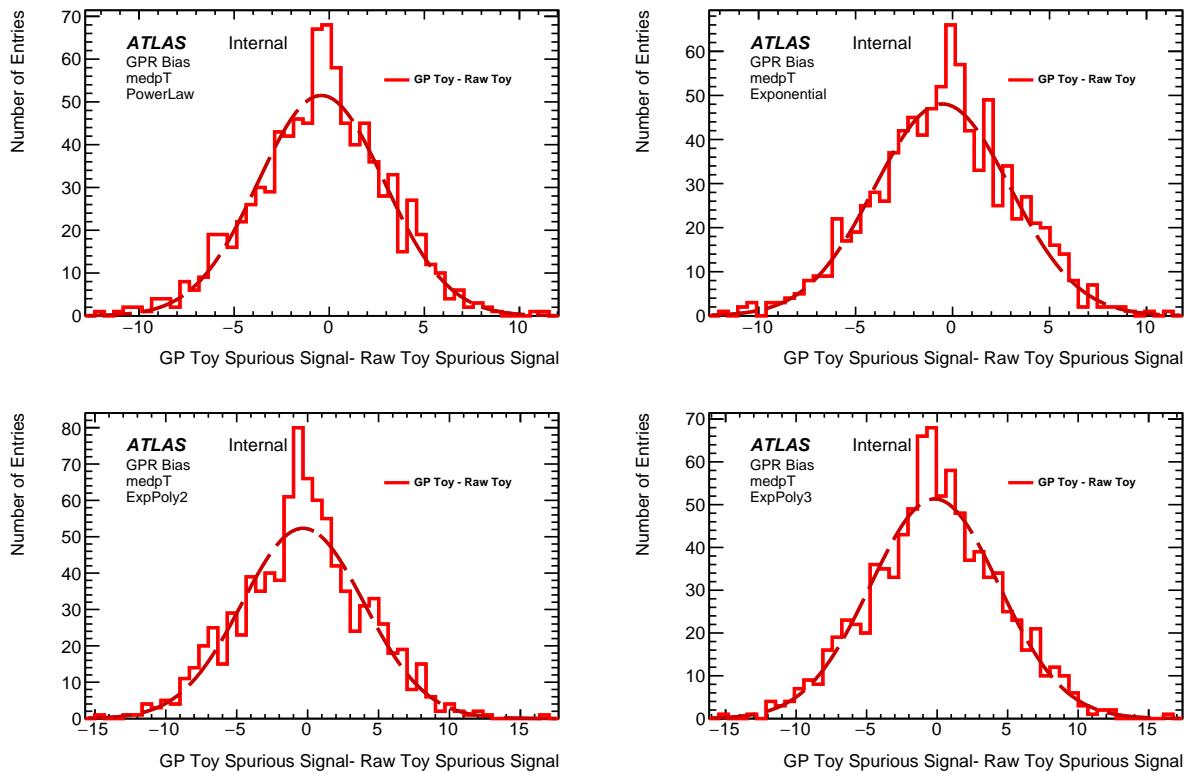


Figure A.172: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100k events.

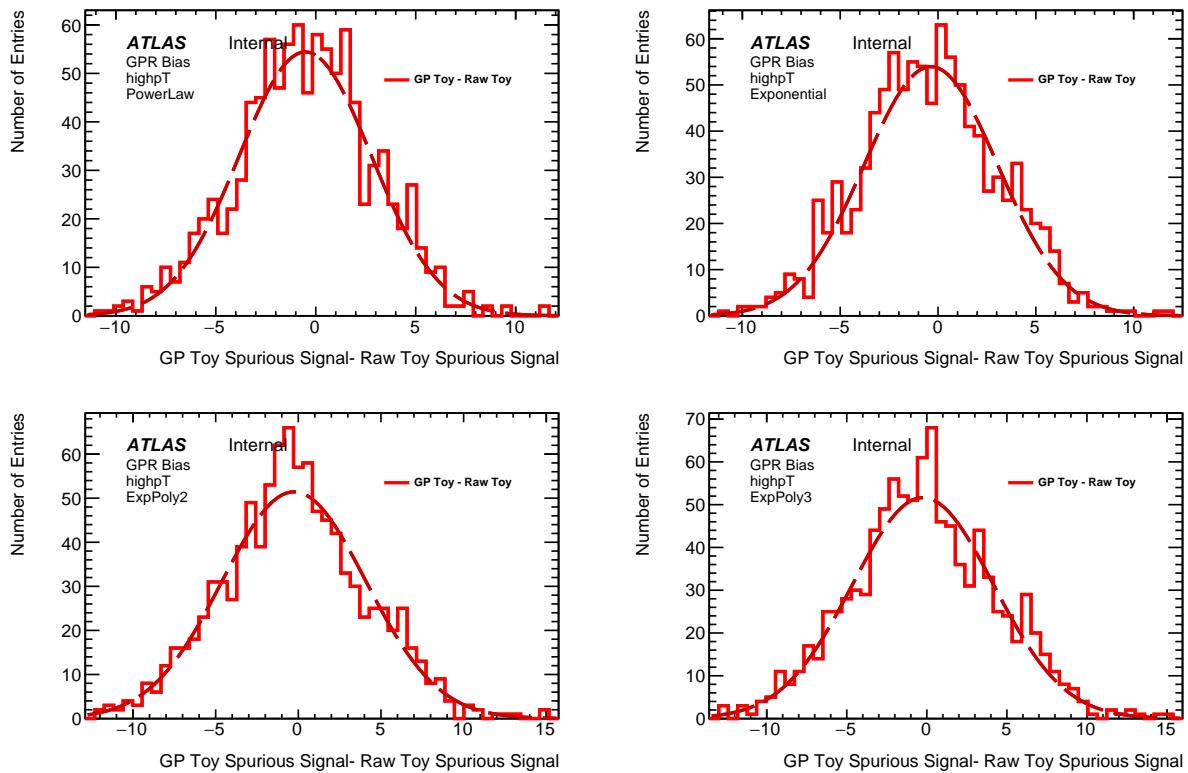


Figure A.173: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100k events.

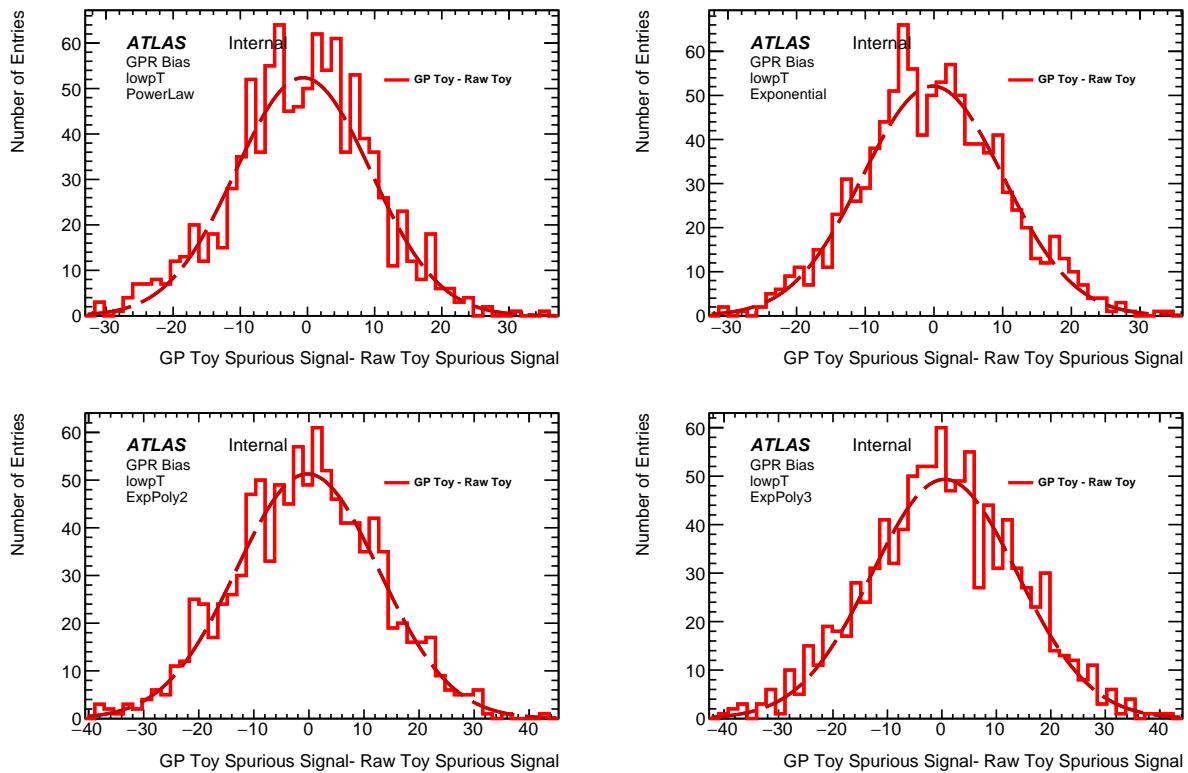


Figure A.174: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

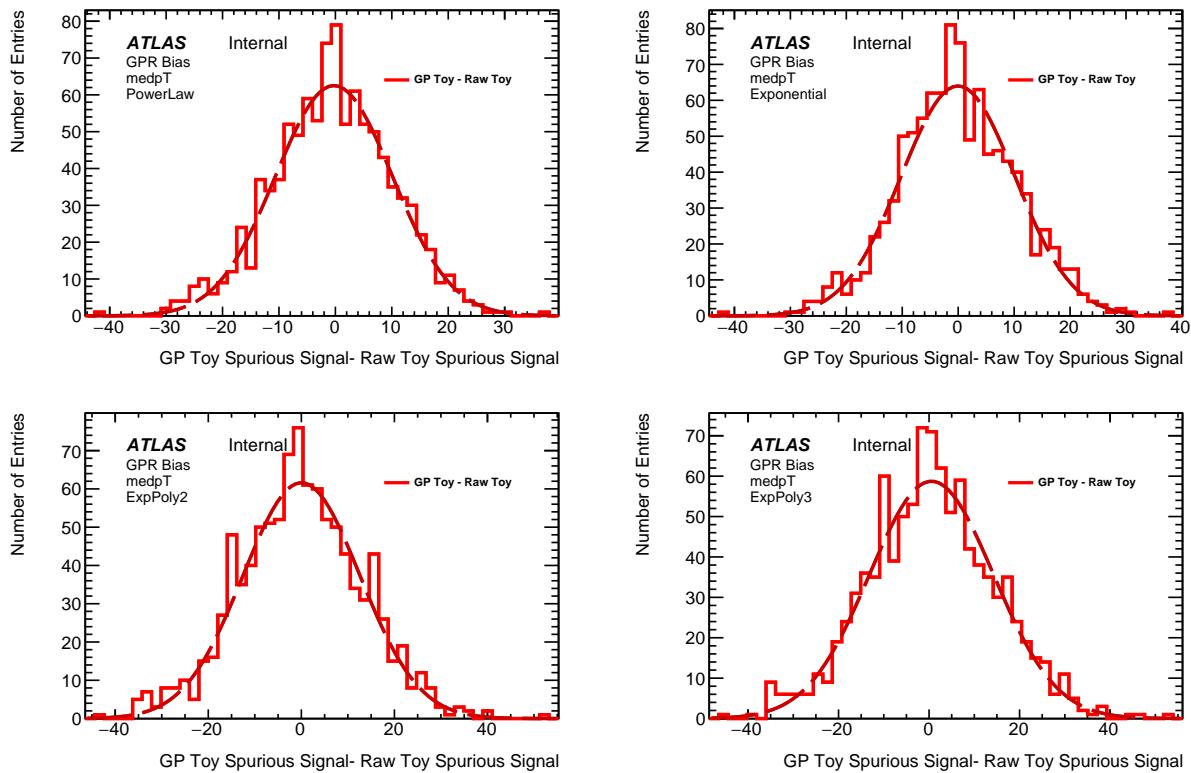


Figure A.175: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

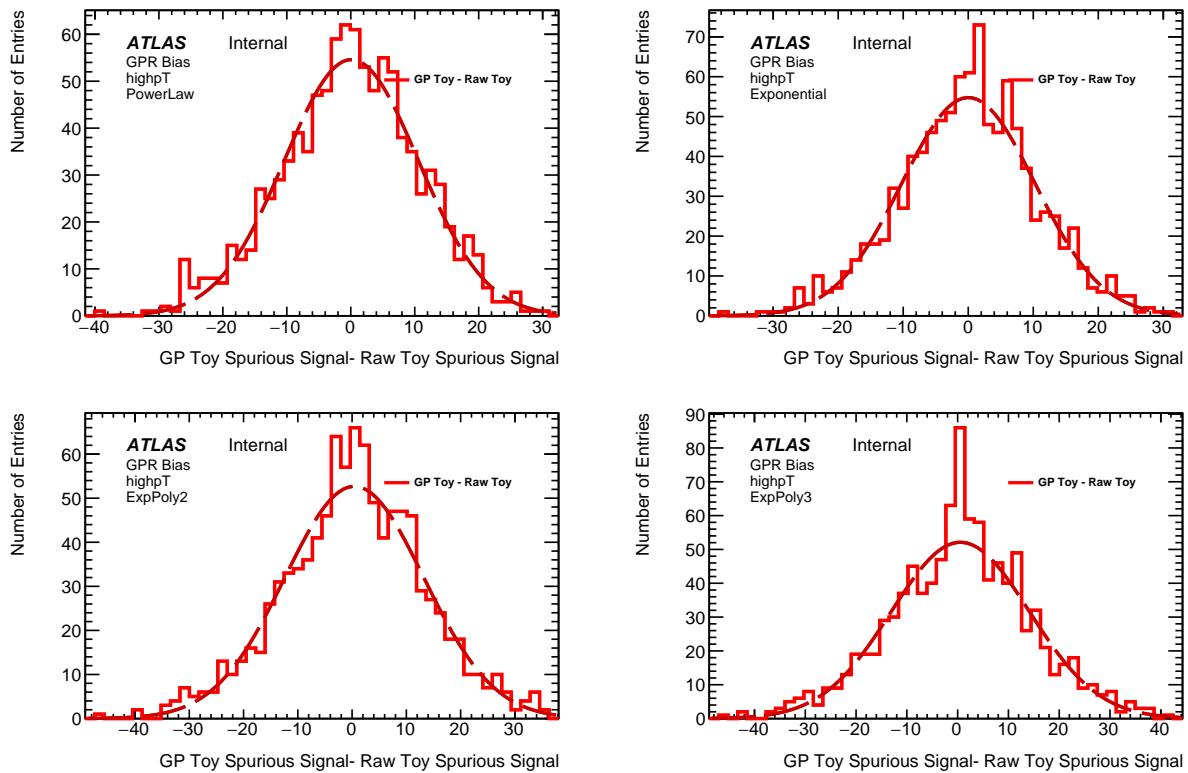


Figure A.176: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

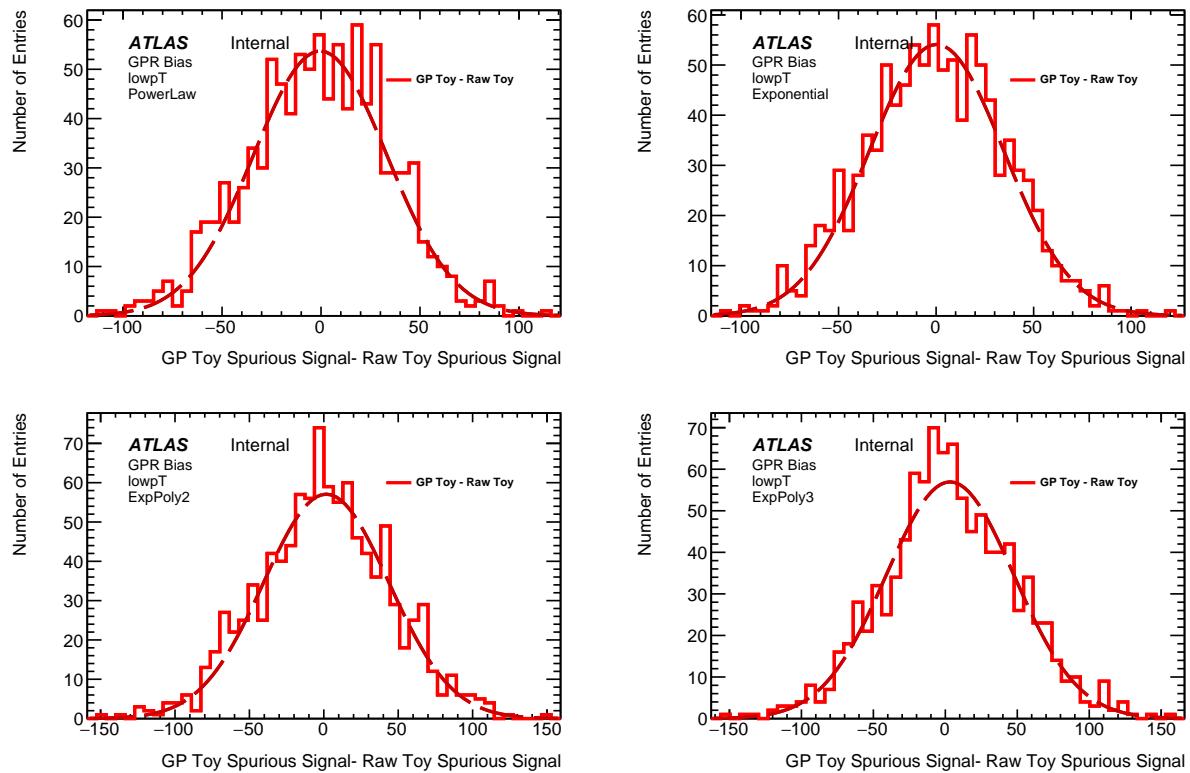


Figure A.177: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10,000,000 events.

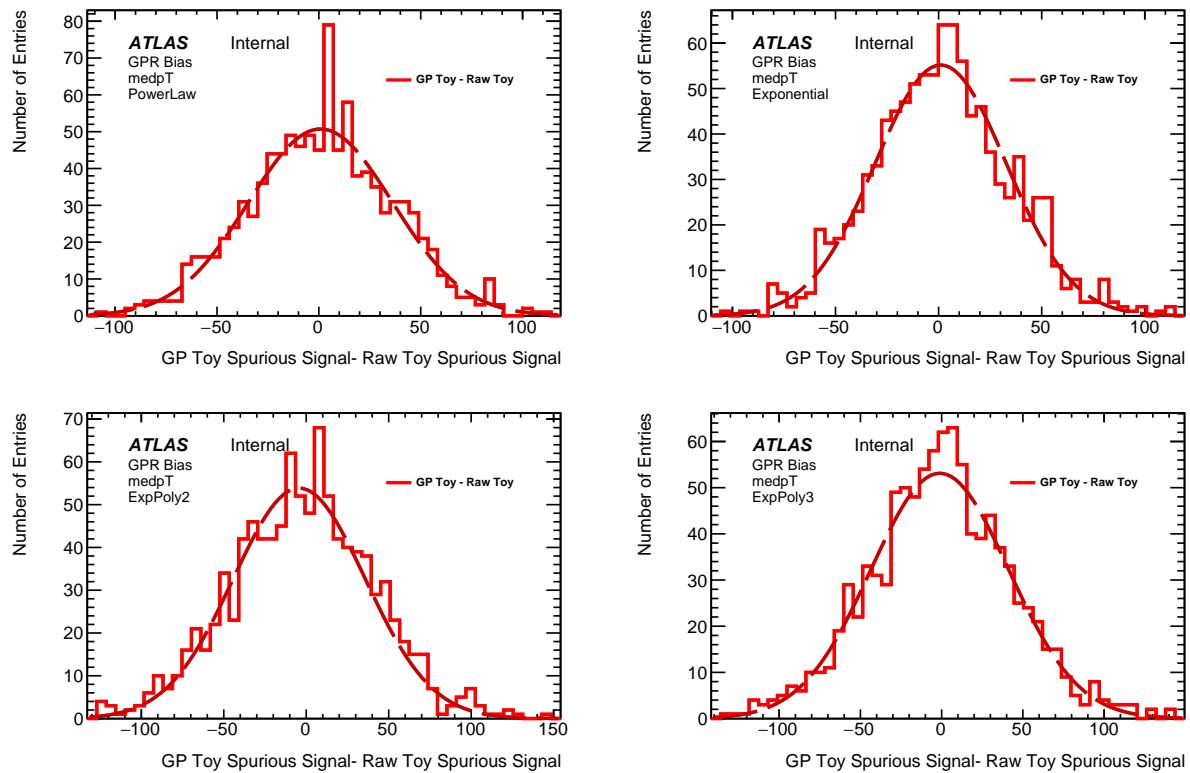


Figure A.178: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10,000,000 events.

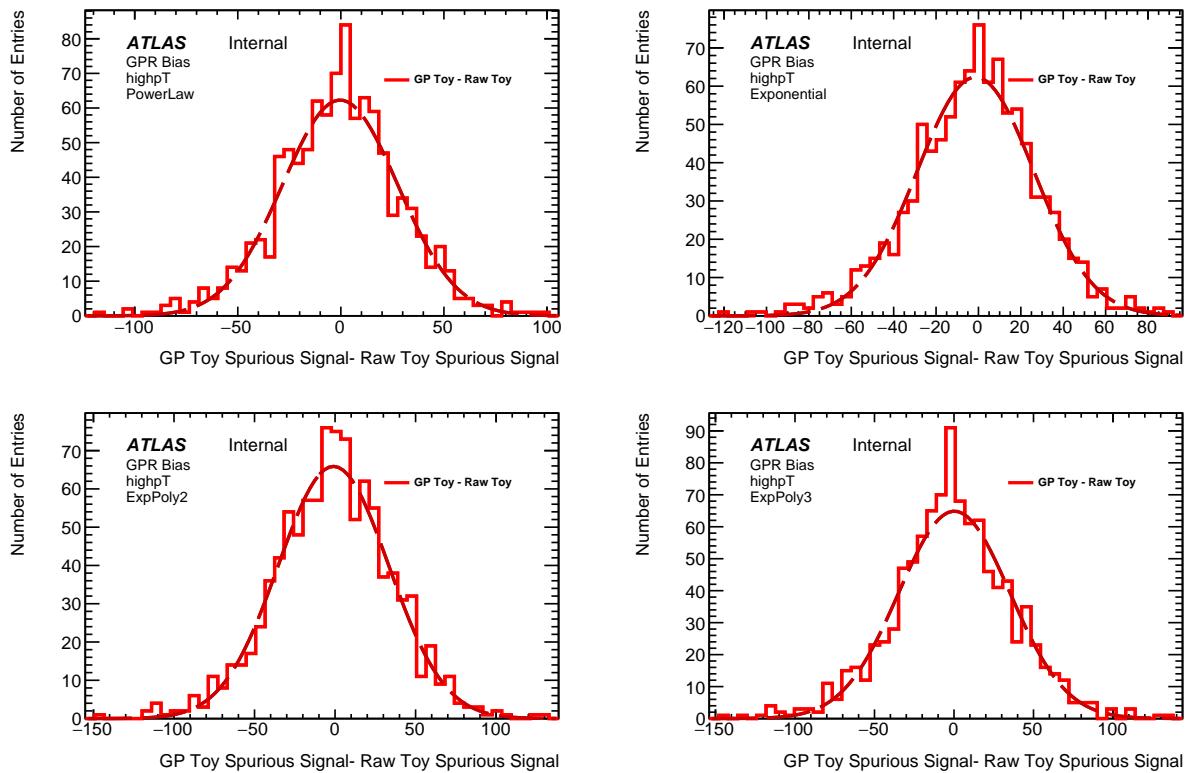


Figure A.179: The per-toy bias (GP Spurious Signal - Raw Spurious Signal), using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10,000,000 events.

To further validate the choice of 20 effective events per bin as the cutoff, we investigate some edge cases. Since the templates have 130 bins (due to the 10 bin padding on either side), we note that the 1000-event templates have just over 7 effective events per bin, while the 10000 event templates have just over 76 events per bin. We test templates with exactly 10 effective events per bin (1300 total events), slightly more than 10 effective events per bin (1400 total events), exactly 20 effective events per bin (2600 total events), and slightly more than 21 effective events per bin (2800 total events).

We note that, in the 10 event/bin regime, for templates generated with ExpPoly2 and ExpPoly3, we see no bias when fitting with ExpPoly2 and ExpPoly3, but see a bias of roughly 35% of the statistical uncertainty on the spurious signal when fitting with lower degree-of-freedom templates (i.e., Exponential and Powerlaw). Upon closer examination, we observe that this is due to the presence of more substantial edge effects in the low-mass region of this very low statistics category that cannot be appropriately modelled by the Gaussian Process.

However, by requiring at least 20 effective events per bin, we observe that this bias is reduced to less than or equal to 20% of the statistical uncertainty on the spurious signal. At the low-statistics end of this range, however, we note that the statistical uncertainty is expected to dominate (that is, spurious signal will not be a significant uncertainty), so we can conclude that the effects of the GPR bias will be minimal. We further note that, as statistics increase past 75 effective events per bin, bias drops off to less than 10% of the spurious signal uncertainty- in regimes where the spurious signal uncertainty is expected to dominate, the bias is found to be negligible.

Nominal	N_sig	Unweighted Events	Bkg events weighted	Mean SS_raw	Mean SS_GPR	GP-Raw: Bias, Mean	Sigma Toy	Sigma GPR	Bias/sigma(SS_raw)	Bias/Feature Size
Fit Function: Exp										
GeneratingFunction: Exp2										
Low1300	0	1300	10	-0.06	0.11	-0.17	0.51	0.42	-0.33	
Low1400	0	1400	10.769 230 769 230 8	-0.09	0.08	-0.16	0.54	0.42	-0.30	
Low2600	0	2600	20	-0.18	-0.03	-0.15	0.72	0.53	-0.21	
Low2800	0	2800	21.538 461 538 461 5	-0.19	-0.03	-0.15	0.76	0.52	-0.20	
Fit Function: Pow										
GeneratingFunction: Exp2										
Low1300	0	1300	10	0.10	0.26	-0.16	0.51	0.42	-0.32	
Low1400	0	1400	10.769 230 769 230 8	0.08	0.24	-0.16	0.54	0.42	-0.29	
Low2600	0	2600	20	0.13	0.28	-0.14	0.72	0.53	-0.20	
Low2800	0	2800	21.538 461 538 461 5	0.15	0.30	-0.15	0.76	0.52	-0.19	
Fit Function: Exp2										
GeneratingFunction: Exp2										
Low1300	0	1300	10	0.01	0.05	-0.04	0.56	0.37	-0.07	
Low1400	0	1400	10.769 230 769 230 8	-0.01	0.04	-0.05	0.59	0.37	-0.08	
Low2600	0	2600	20	-0.03	0.02	-0.05	0.79	0.45	-0.06	
Low2800	0	2800	21.538 461 538 461 5	-0.00	0.05	-0.05	0.83	0.42	-0.06	
Fit Function: Exp3										
GeneratingFunction: Exp2										
Low1300	0	1300	10	0.00	0.01	-0.01	0.55	0.35	-0.01	
Low1400	0	1400	10.769 230 769 230 8	-0.01	0.00	-0.01	0.58	0.35	-0.02	
Low2600	0	2600	20	-0.03	-0.03	-0.00	0.78	0.42	-0.00	
Low2800	0	2800	21.538 461 538 461 5	-0.00	0.00	-0.01	0.83	0.38	-0.01	

Table A.14: Spurious signal means and widths for all choices of fit functional-form, using the "low" template with the ExpPoly2 generating functional form, for a range of different template statistics.

Nominal	N_sig	Unweighted Events	Bkg events weighted	Mean SS_raw	Mean SS_GPR	GP-Raw: Bias, Mean	Sigma Toy	Sigma GPR	Bias/sigma(SS_raw)	Bias/Feature Size
Fit Function: Exp										
Generating Function: Exp3										
Med1300	0	1300	10	-0.02	0.14	-0.17	0.53	0.43	-0.32	
Med1400	0	1400	10.769 230 769 230 8	0.01	0.03	-0.01	0.59	0.43	-0.02	
Med2600	0	2600	20	-0.09	0.04	-0.13	0.66	0.42	-0.20	
Med2800	0	2800	21.538 461 538 461 5	-0.11	0.02	-0.13	0.75	0.54	-0.18	
Fit Function: Pow										
Generating Function: Exp3										
Med1300	0	1300	10	0.17	0.33	-0.16	0.53	0.42	-0.31	
Med1400	0	1400	10.769 230 769 230 8	0.17	0.32	-0.15	0.54	0.48	-0.27	
Med2600	0	2600	20	0.30	0.43	-0.13	0.66	0.42	-0.19	
Med2800	0	2800	21.538 461 538 461 5	0.30	0.44	-0.13	0.75	0.54	-0.18	
Fit Function: Exp3										
Generating Function: Exp3										
Med1300	0	1300	10	0.01	0.03	-0.01	0.58	0.37	-0.02	
Med1400	0	1400	10.769 230 769 230 8	0.01	-0.00	0.01	0.58	0.41	0.02	
Med2600	0	2600	20	-0.02	0.03	-0.05	0.80	0.35	-0.07	
Med2800	0	2800	21.538 461 538 461 5	-0.01	-0.00	-0.01	0.83	0.41	-0.01	
Fit Function: Exp2										
Generating Function: Exp3										
Med1300	0	1300	10	0.01	0.05	-0.04	0.36	0.38	-0.11	
Med1400	0	1400	10.769 230 769 230 8	0.01	0.03	-0.01	0.59	0.43	-0.02	
Med2600	0	2600	20	0.01	0.08	-0.08	0.79	0.36	-0.10	
Med2800	0	2800	21.538 461 538 461 5	-0.01	0.03	-0.04	0.82	0.44	-0.05	
Fit Function: Exp2										
Generating Function: Exp3										
High1300	0	1300	10	-0.04	-0.02	-0.03	0.54	0.49	-0.05	
High1400	0	1400	10.769 230 769 230 8	-0.00	0.03	-0.04	0.61	0.58	-0.06	
High2600	0	2600	20	-0.03	0.01	-0.04	0.76	0.41	-0.05	
High2800	0	2800	21.538 461 538 461 5	-0.01	0.02	-0.02	0.82	0.44	-0.03	

Table A.15: Spurious signal means and widths for all choices of fit functional-form, using the "medium" template with the ExpPoly3 generating functional form and the "high" template with the ExpPoly3 generating functional form, for a range of different template statistics.

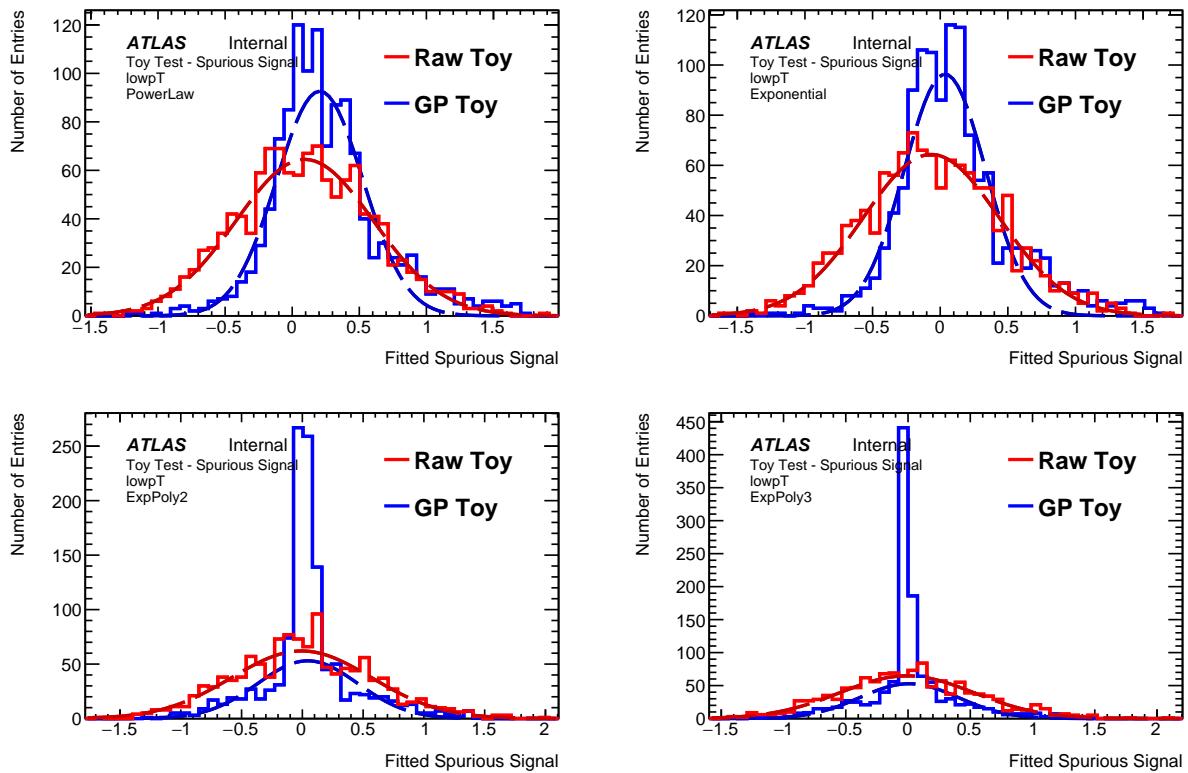


Figure A.180: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1300 events.

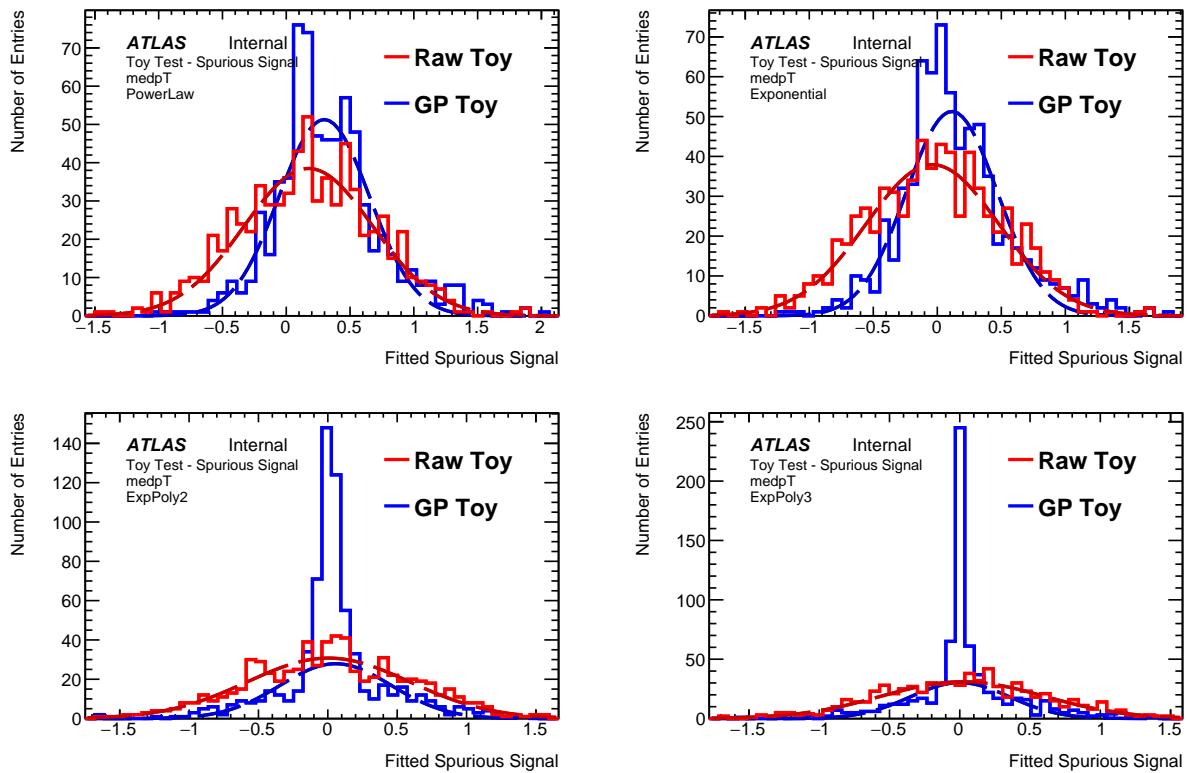


Figure A.181: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1300 events.

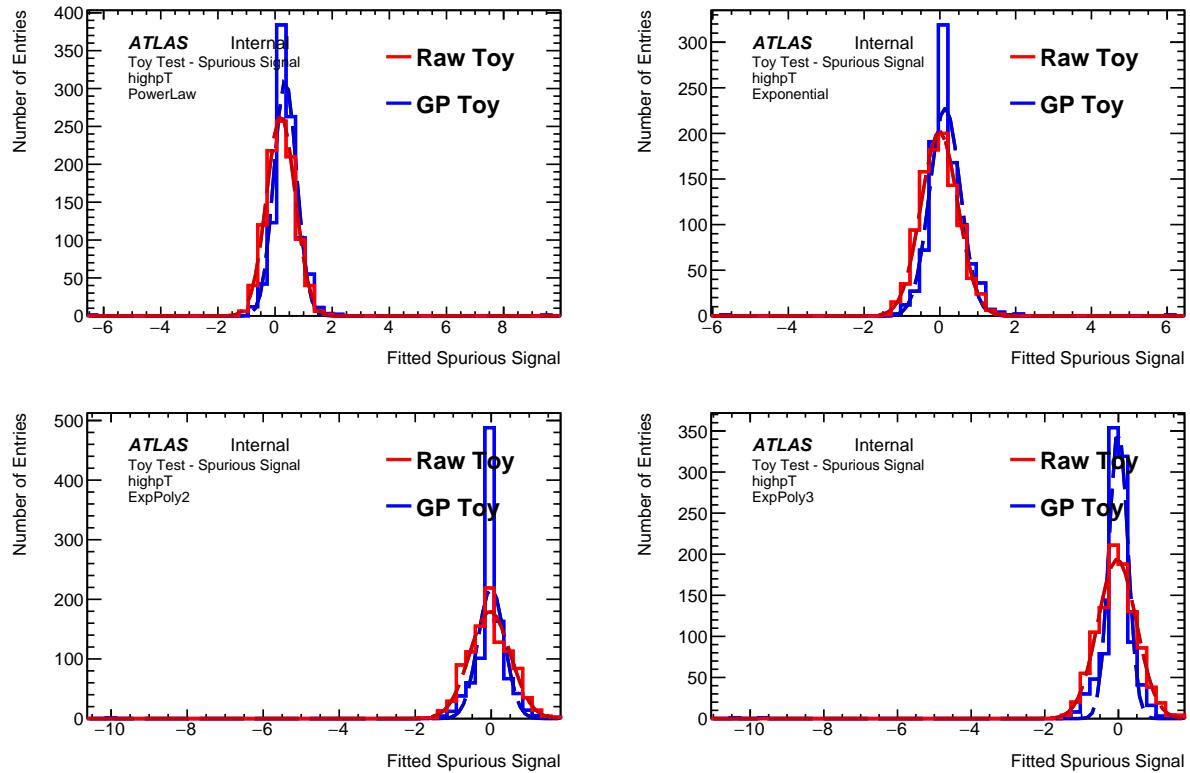


Figure A.182: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1300 events.

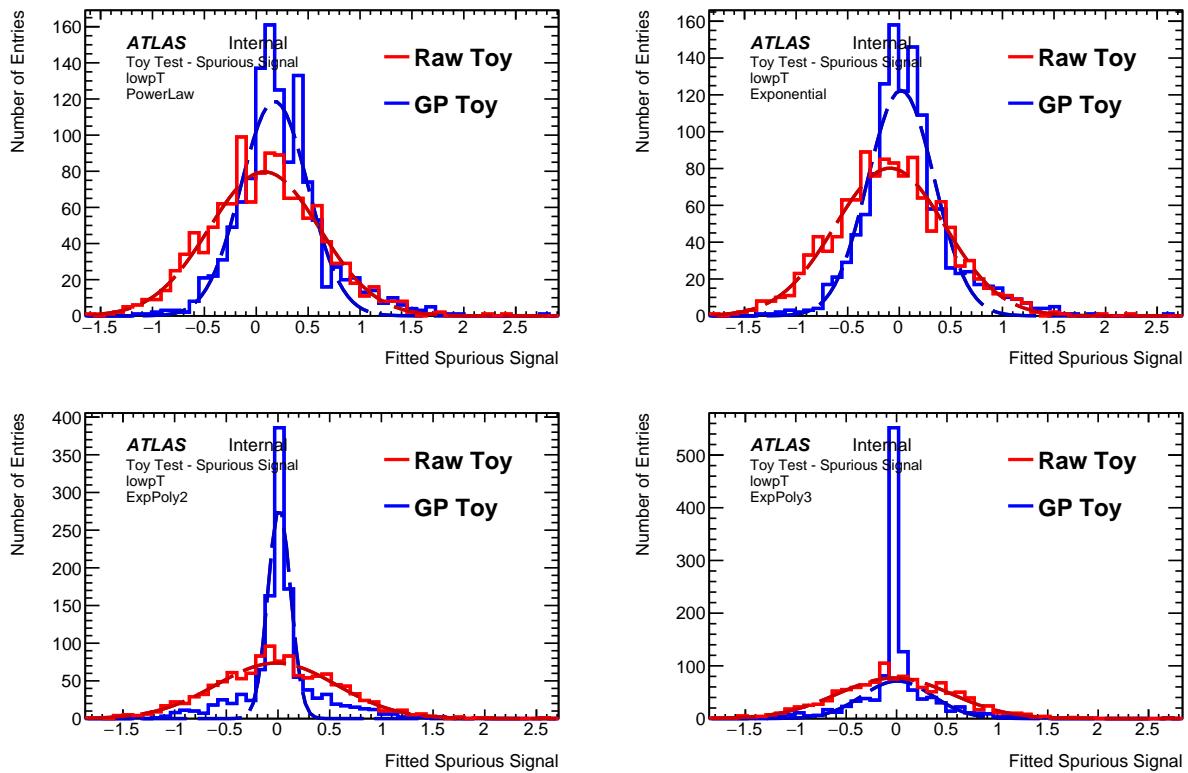


Figure A.183: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1400 events.

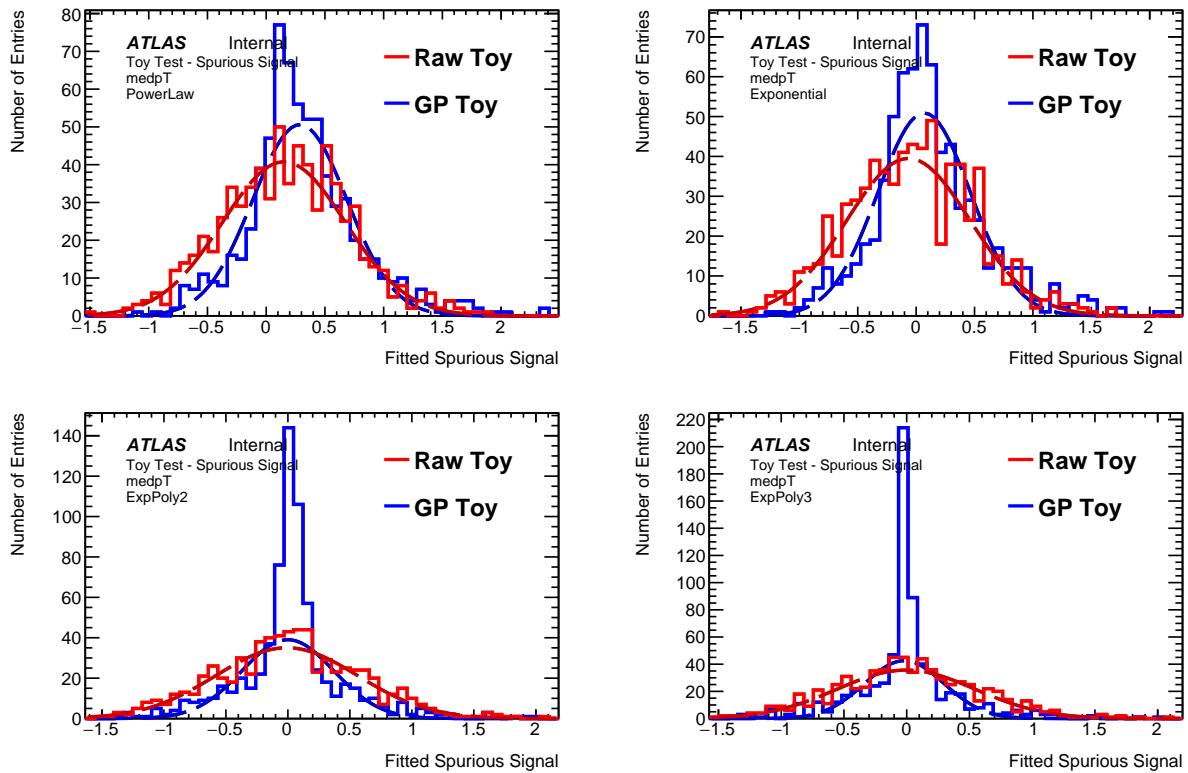


Figure A.184: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1400 events.

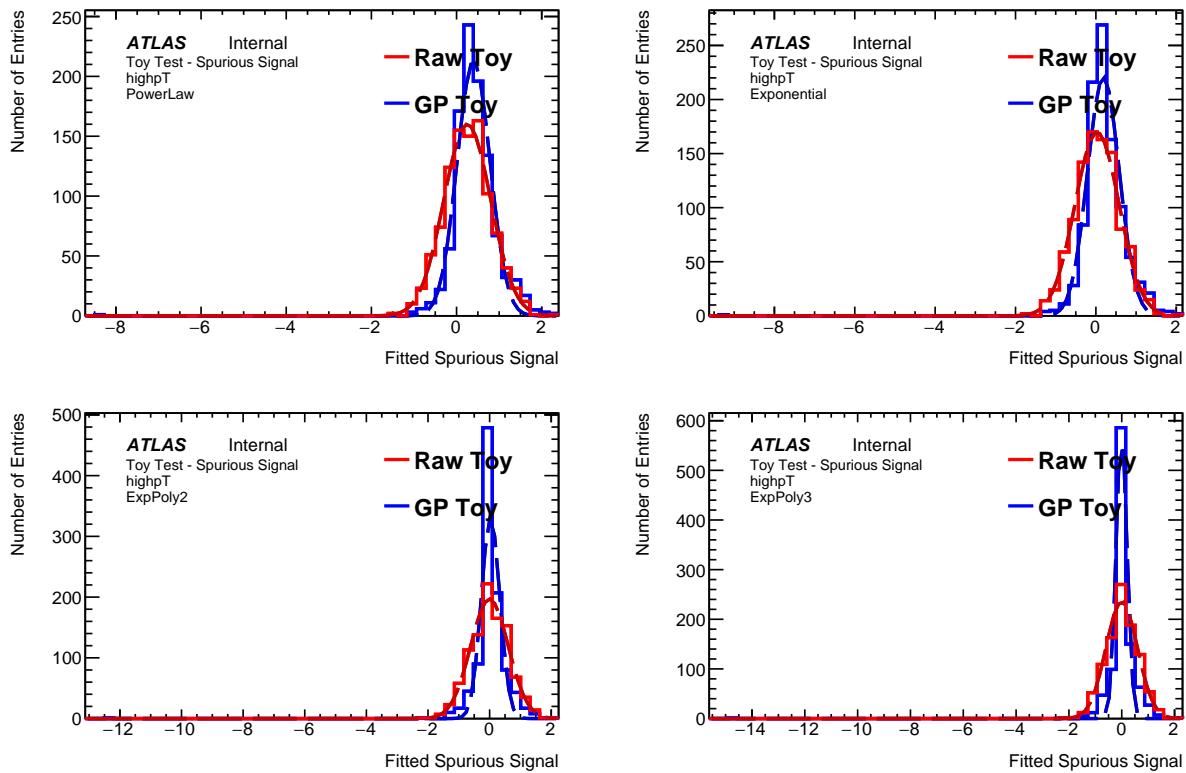


Figure A.185: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1400 events.

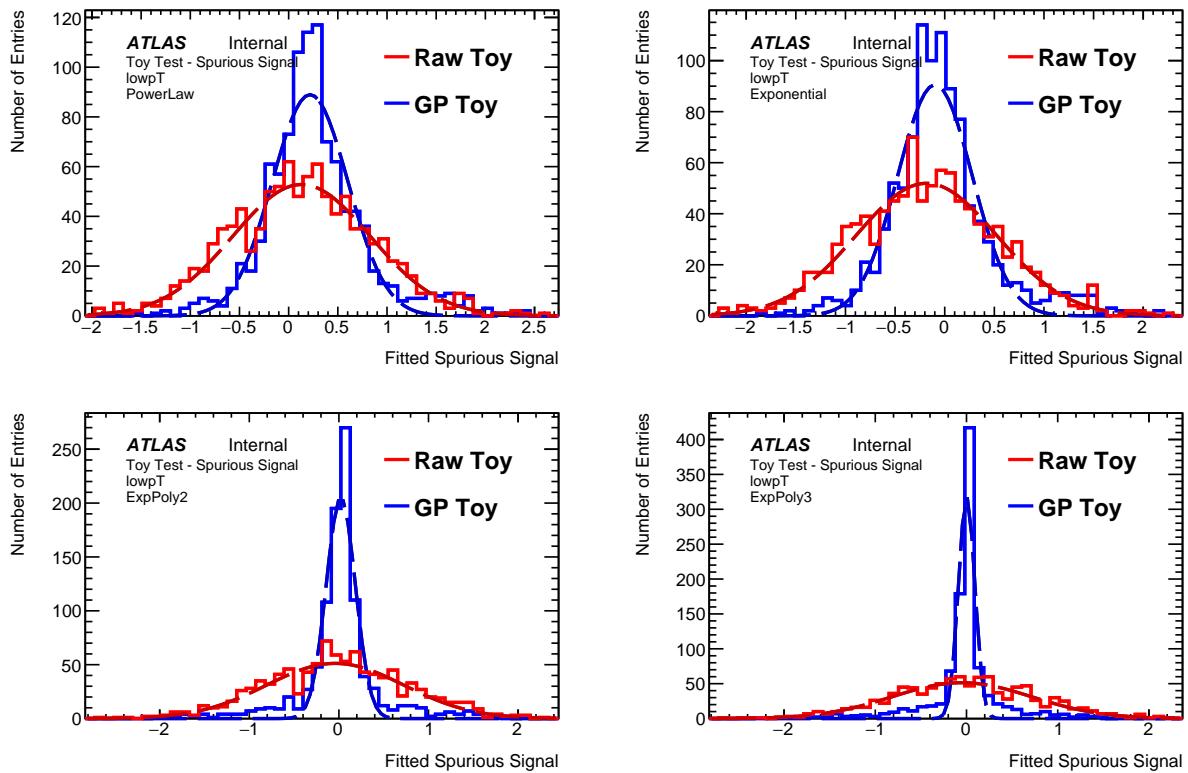


Figure A.186: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 2600 events.

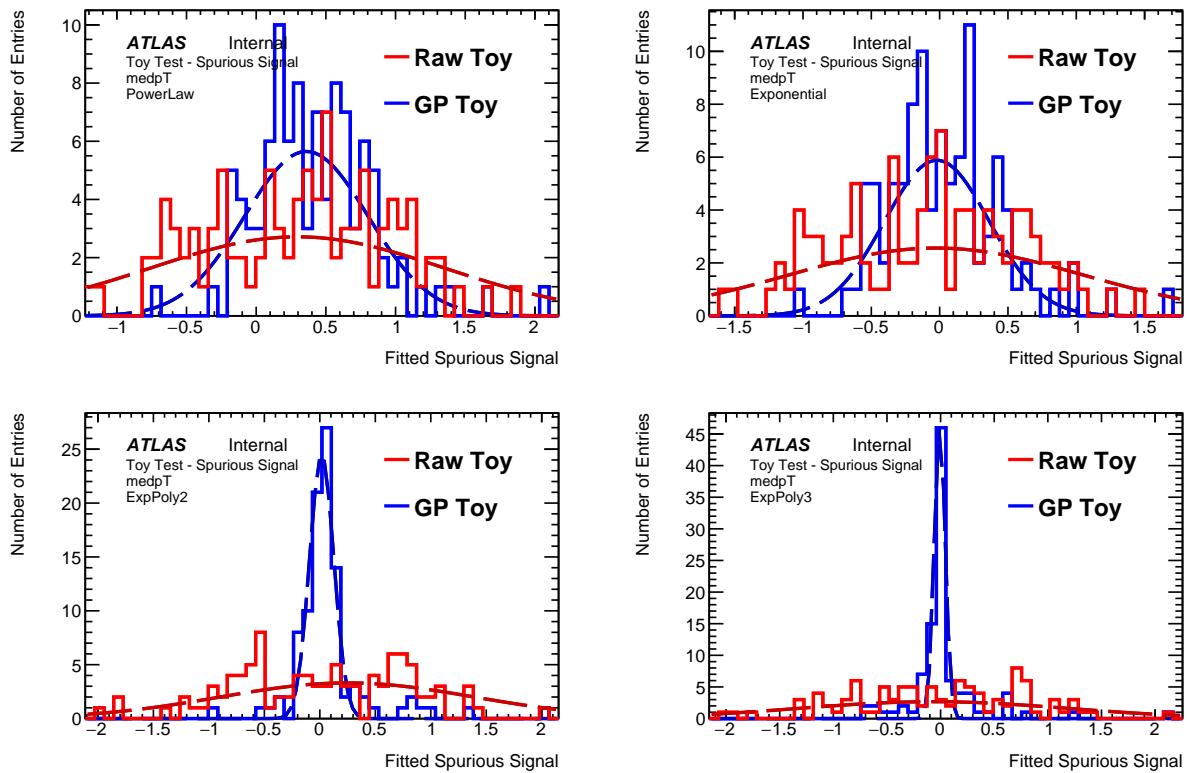


Figure A.187: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 2600 events.

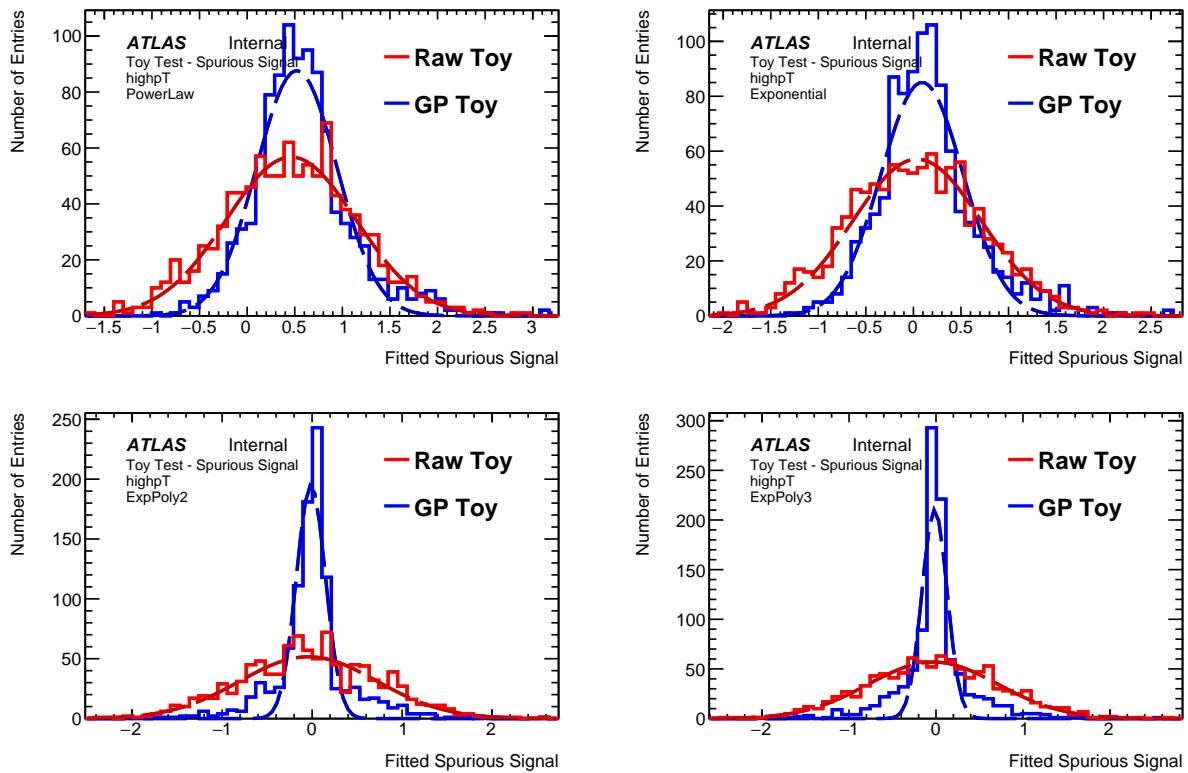


Figure A.188: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 2600 events.

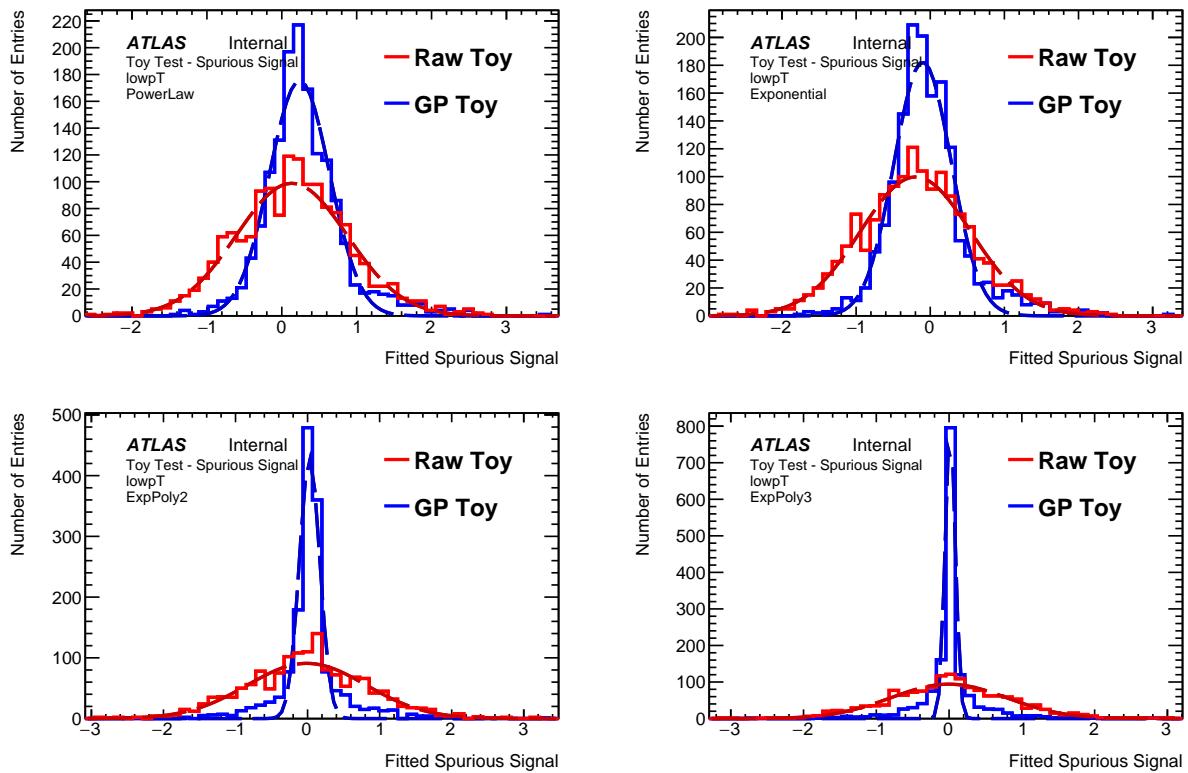


Figure A.189: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 2800 events.

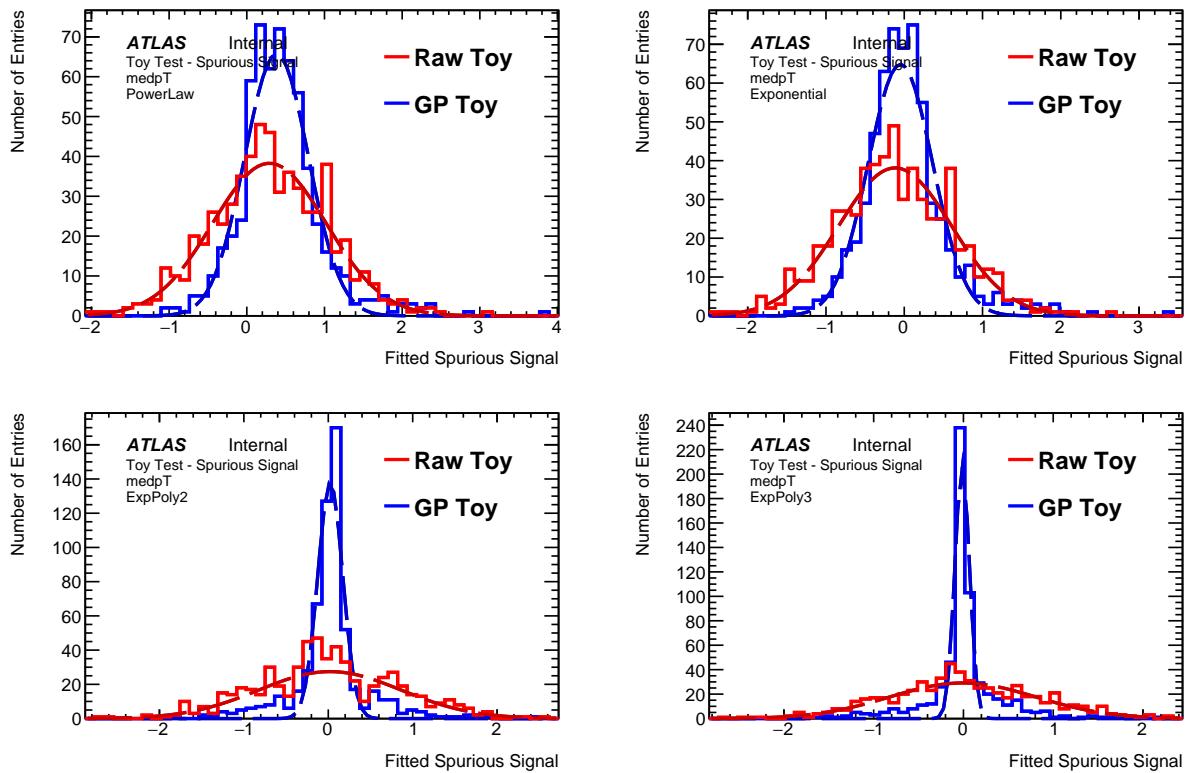


Figure A.190: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 2800 events.

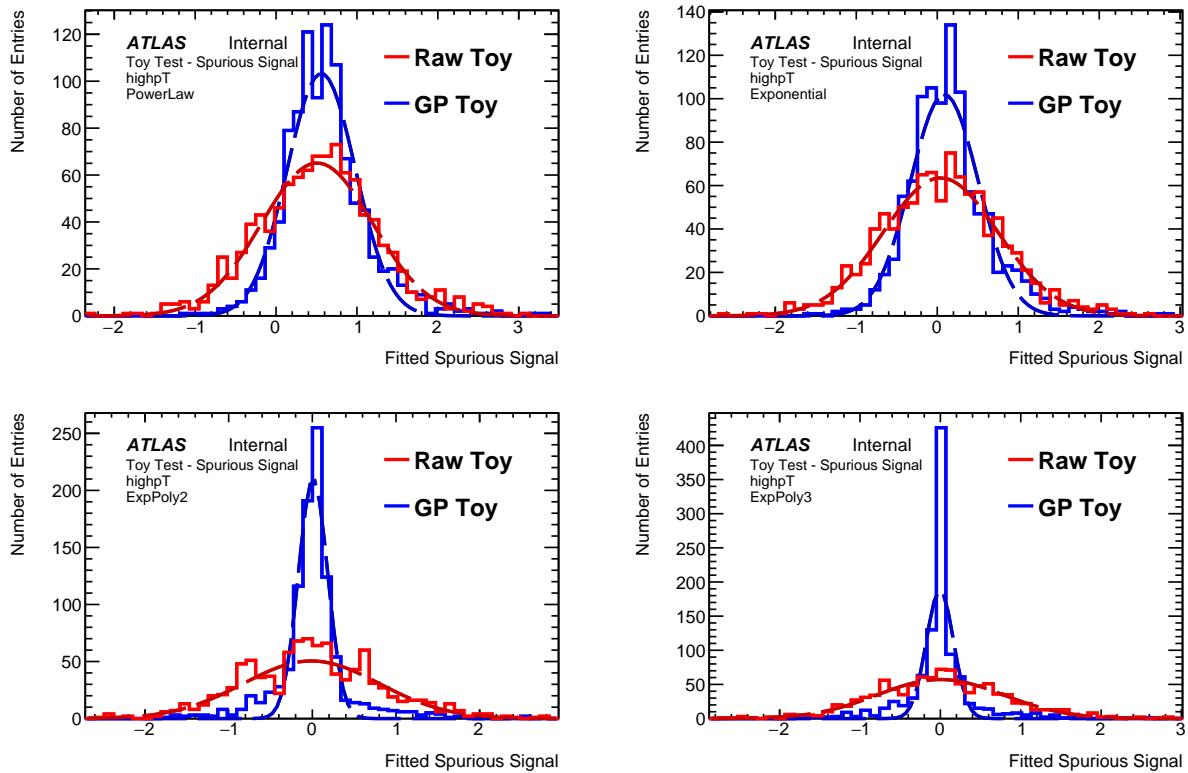


Figure A.191: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 2800 events.

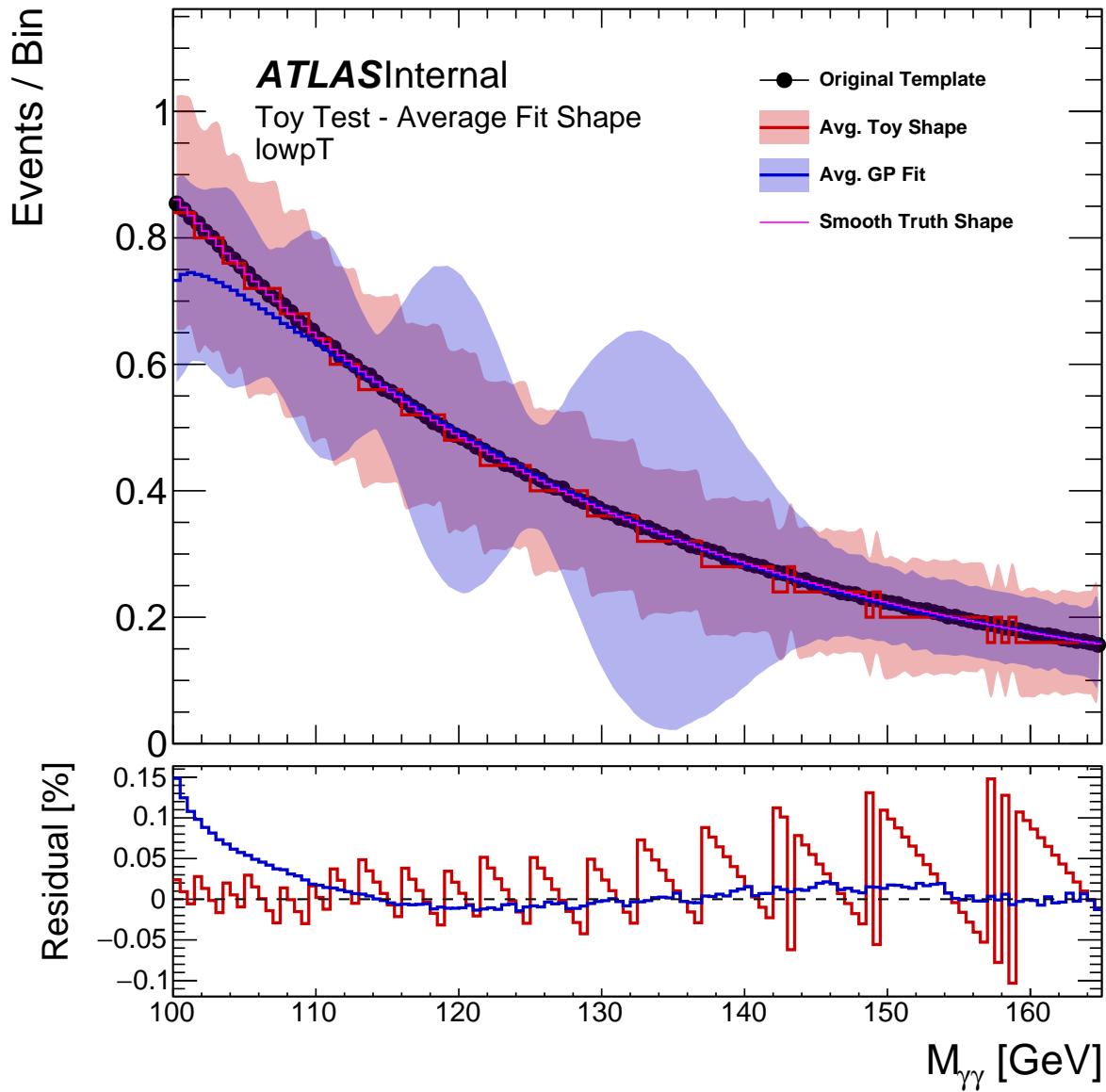


Figure A.192: The average toy distribution for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1300 events.

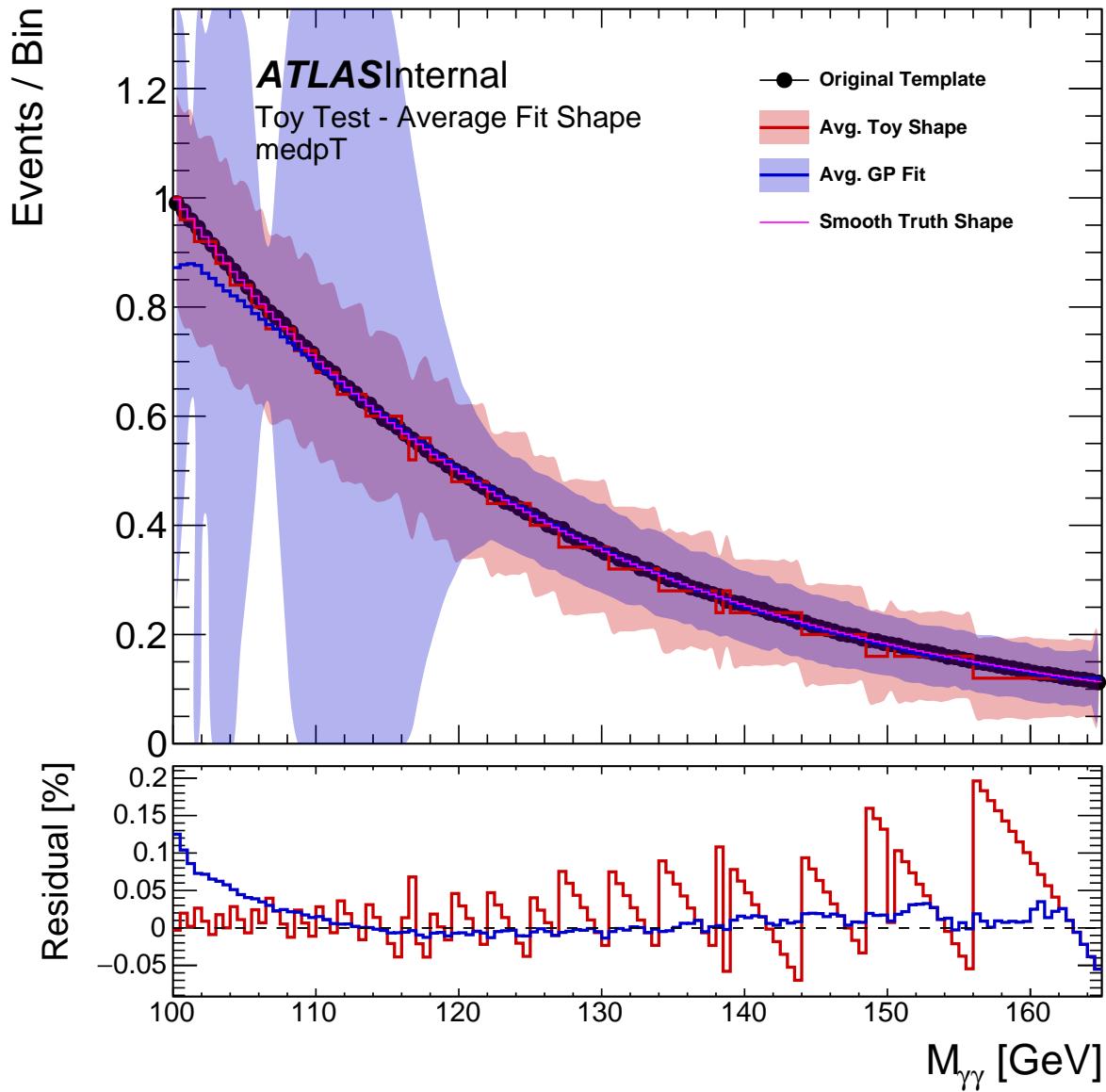


Figure A.193: The average toy distribution for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1300 events.

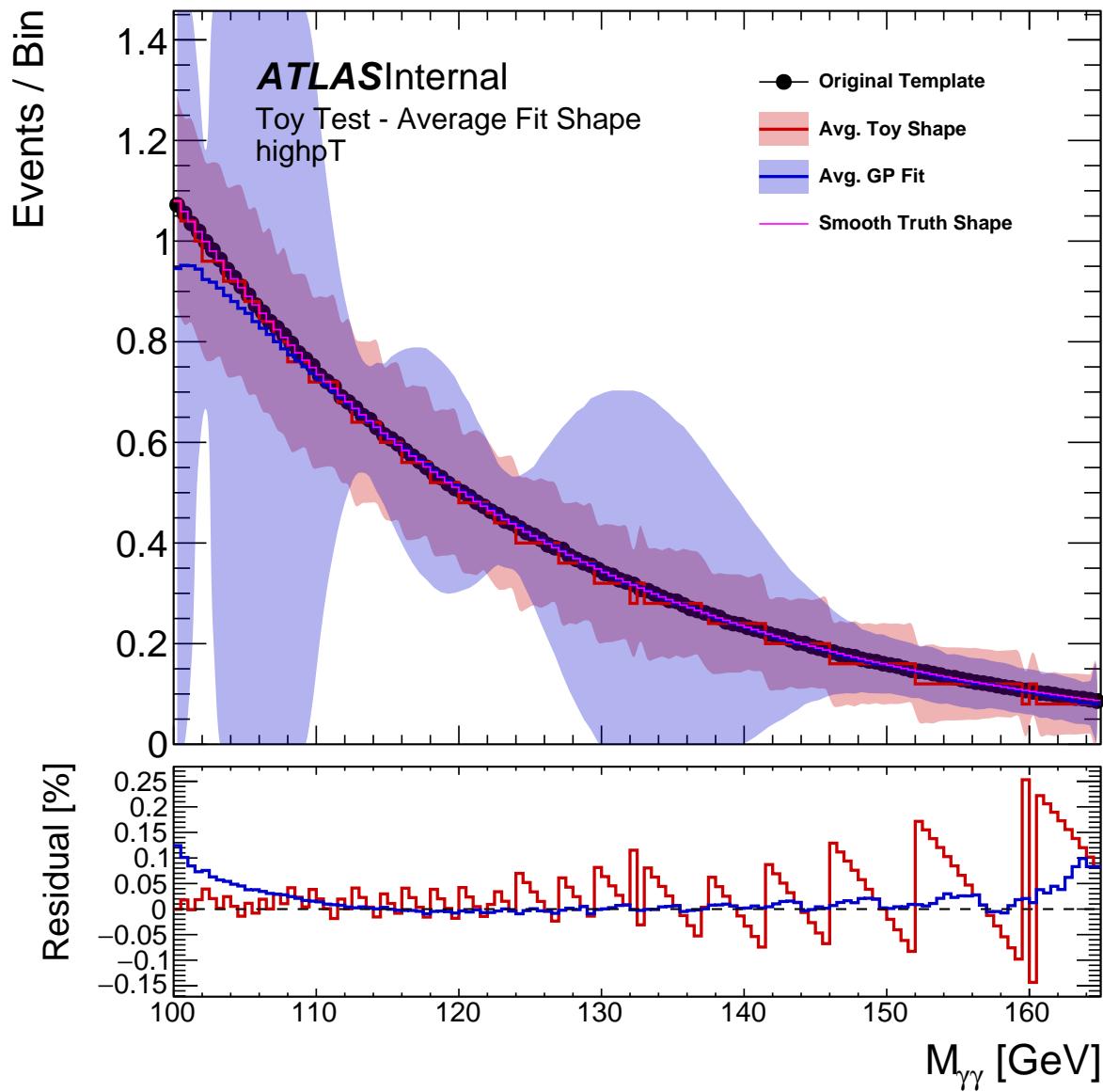


Figure A.194: The average toy distribution for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1300 events.

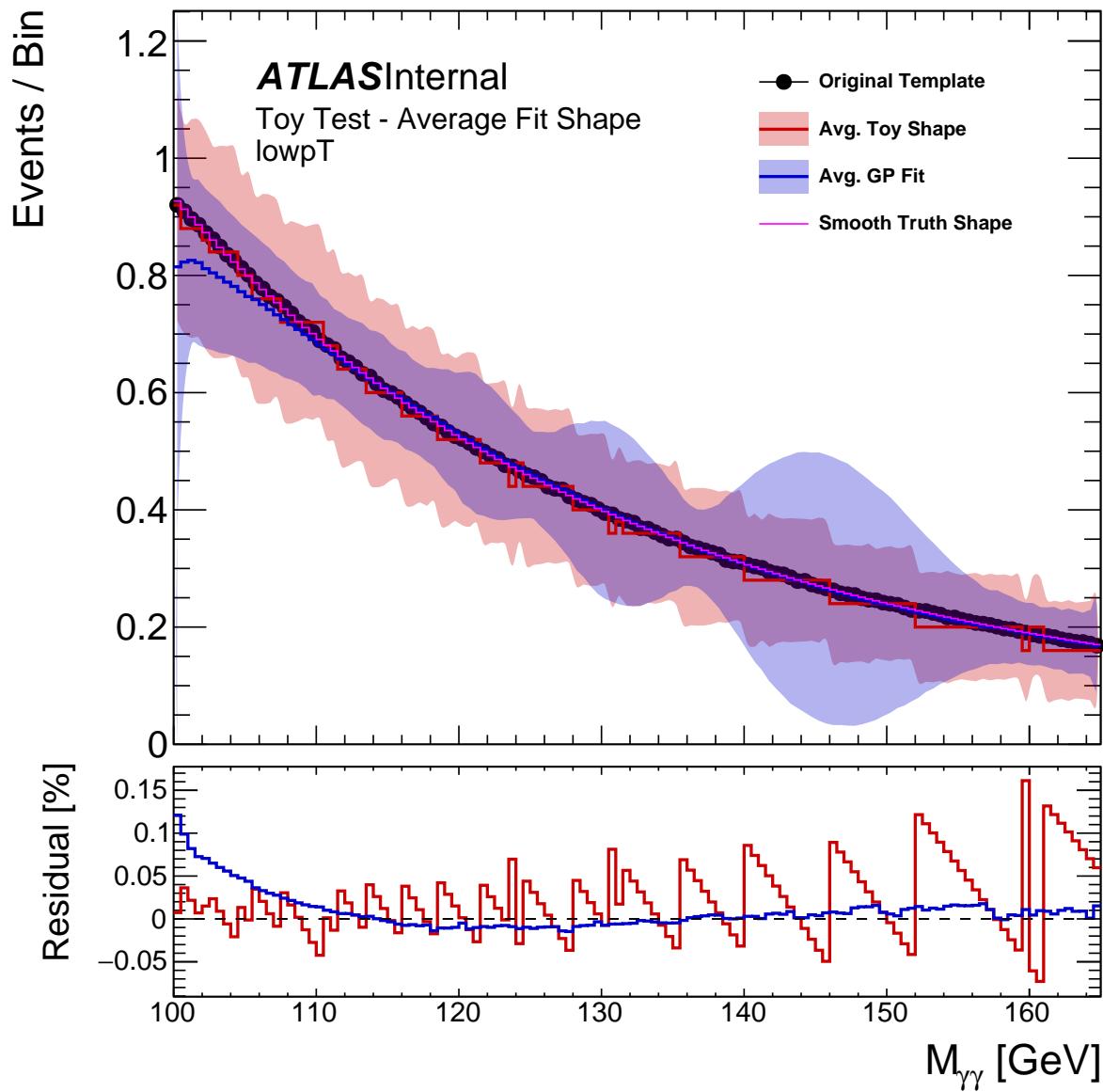


Figure A.195: The average toy distribution for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1400 events.

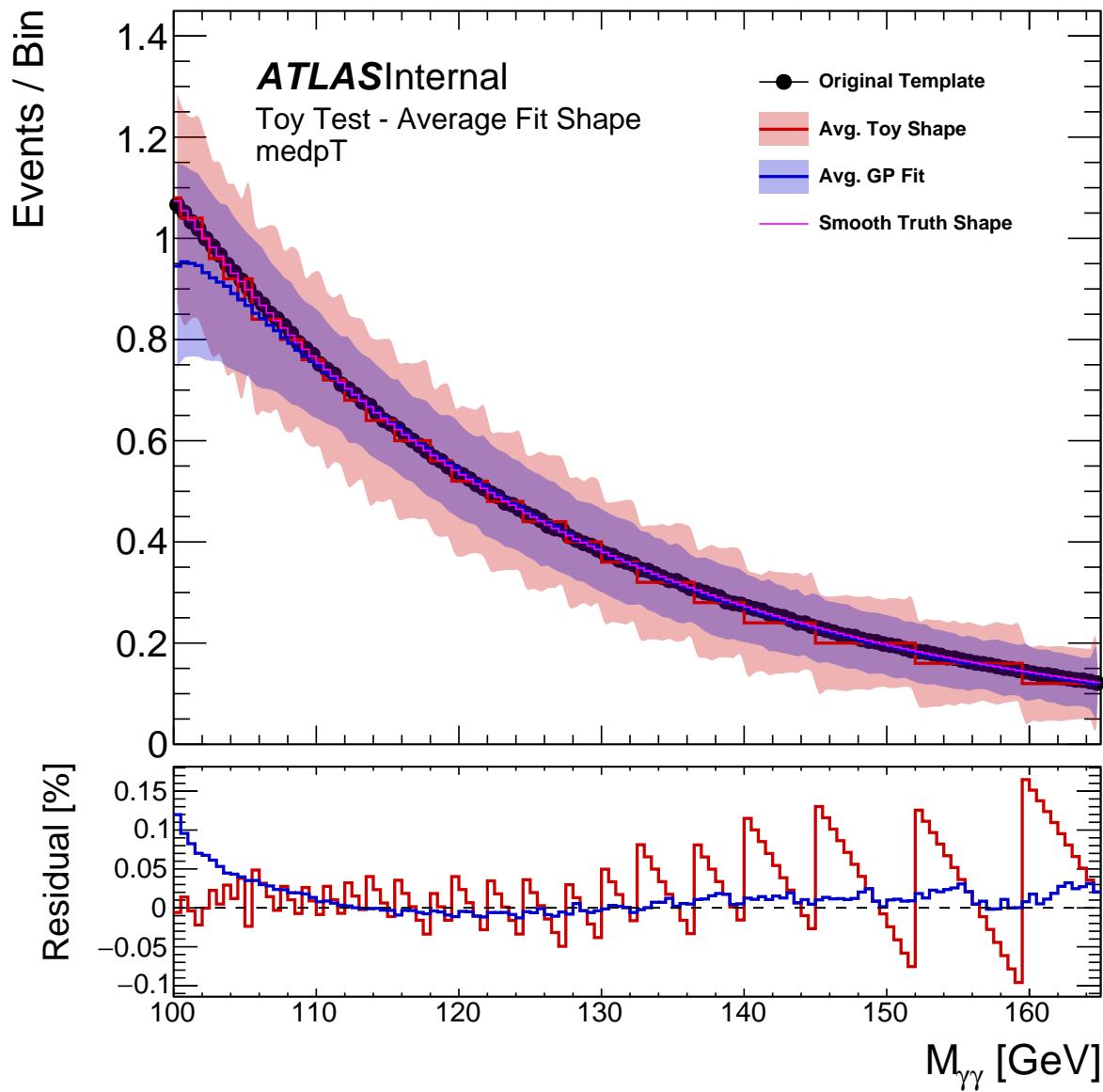


Figure A.196: The average toy distribution for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1400 events.

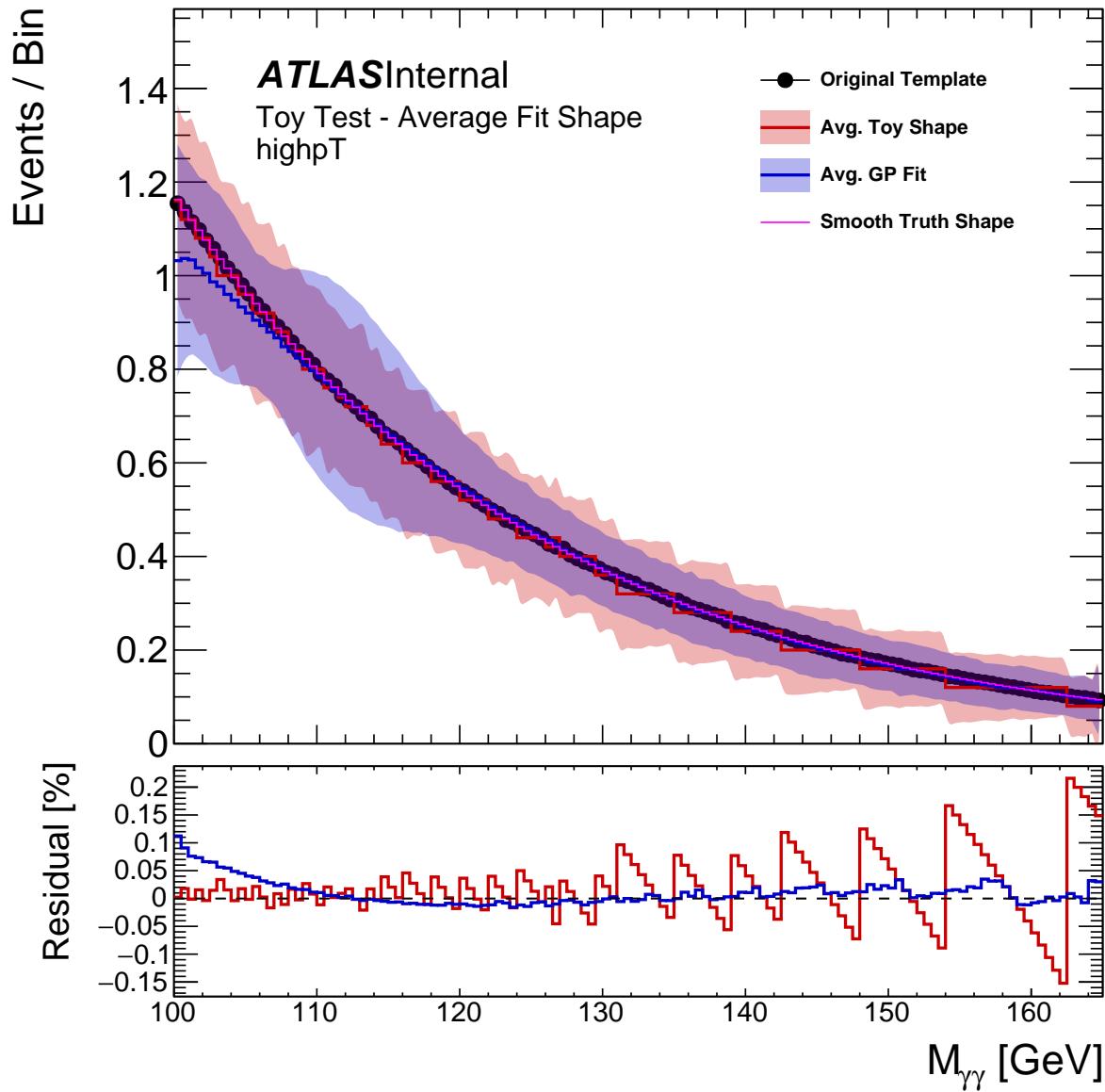


Figure A.197: The average toy distribution for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1400 events.

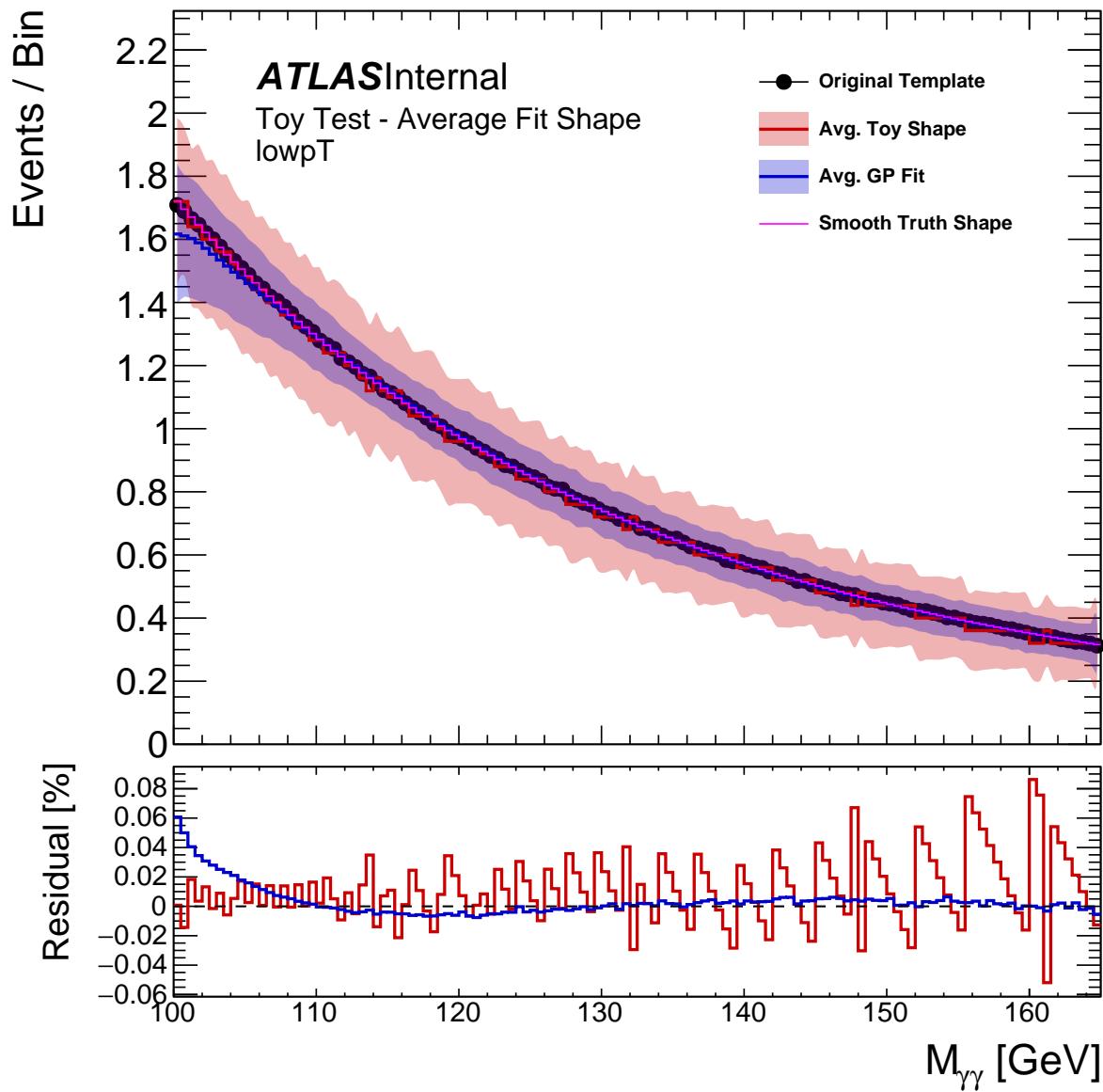


Figure A.198: The average toy distribution for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 2600 events.

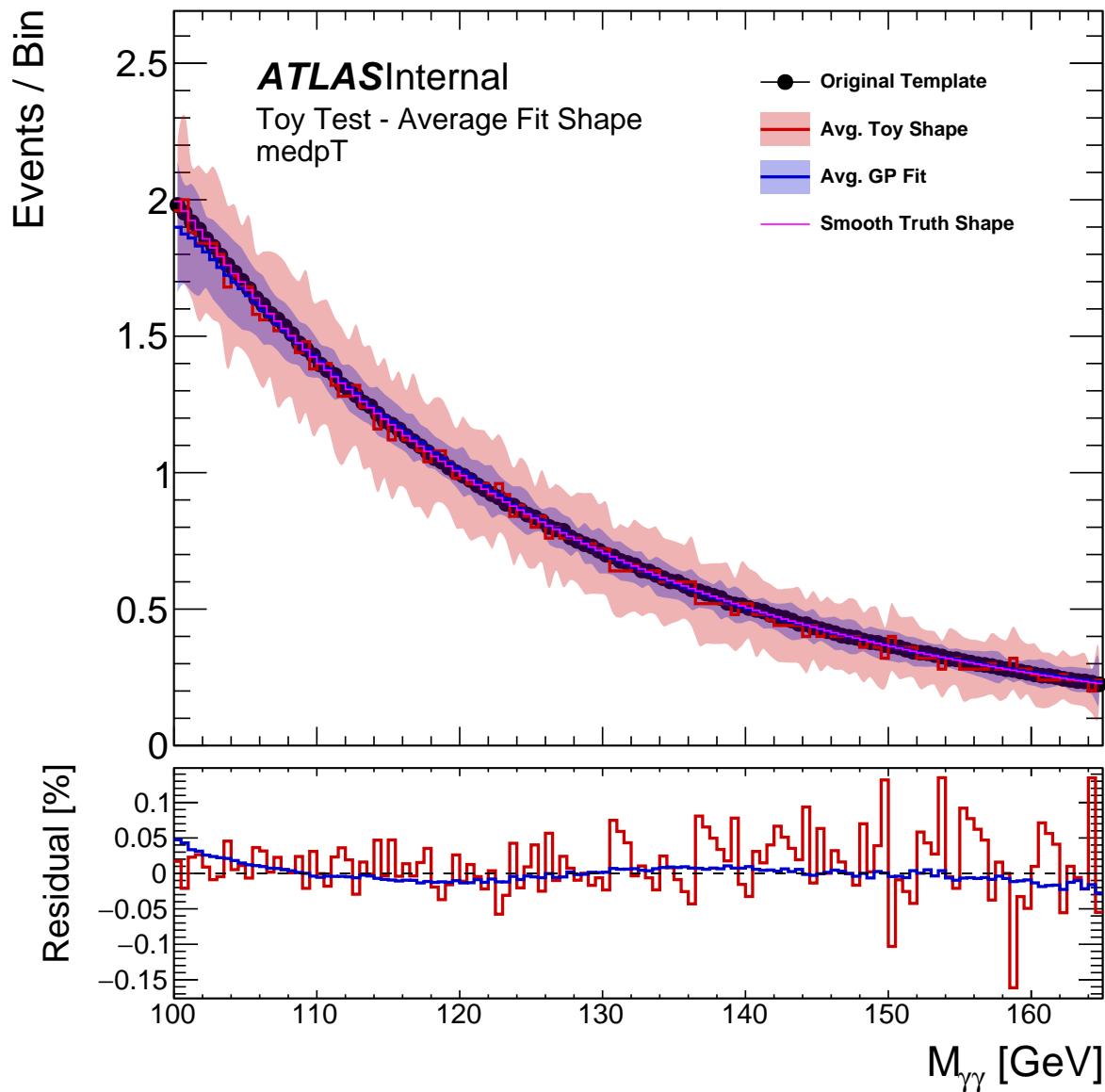


Figure A.199: The average toy distribution for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 2600 events.

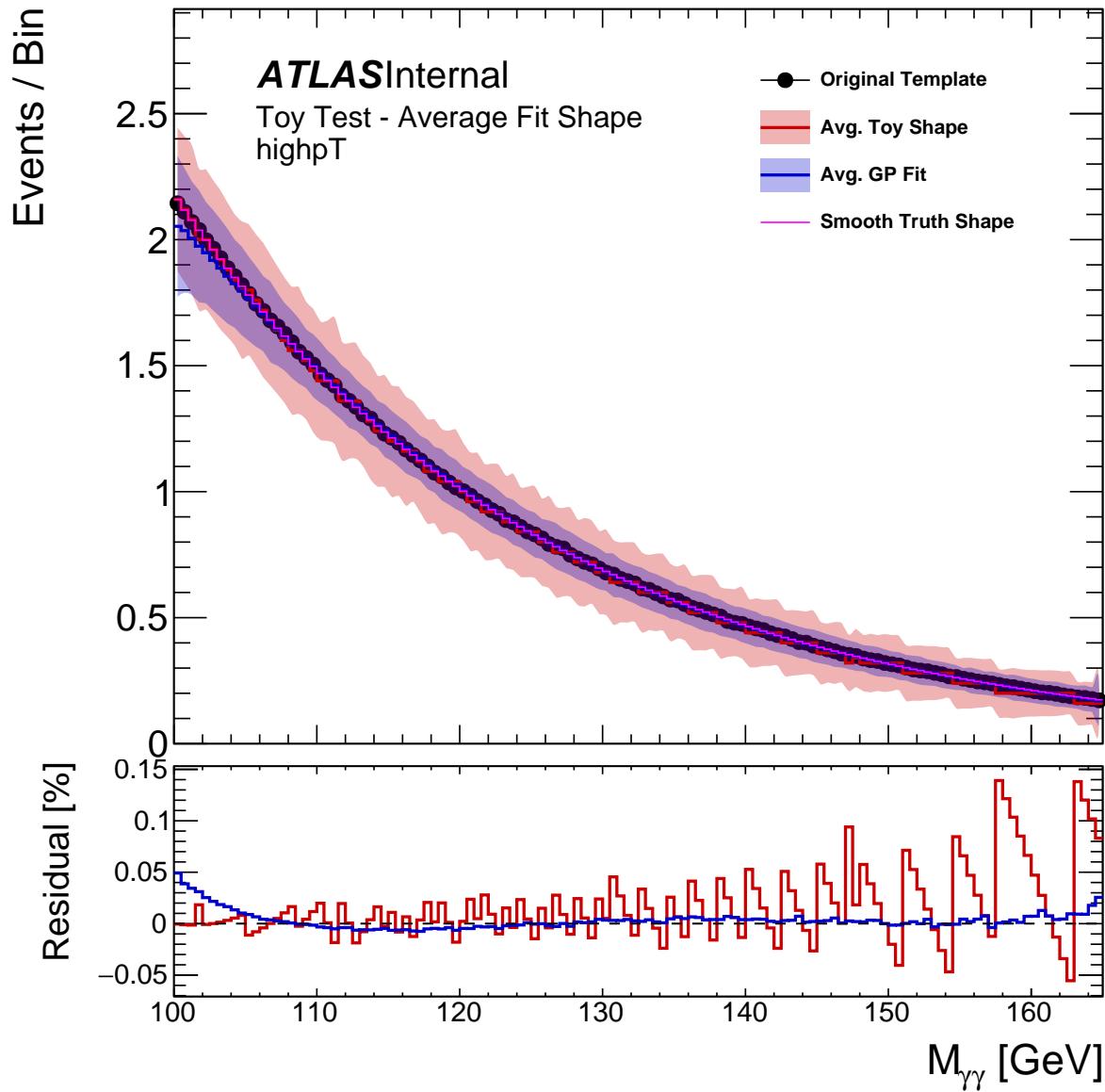


Figure A.200: The average toy distribution for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 2600 events.

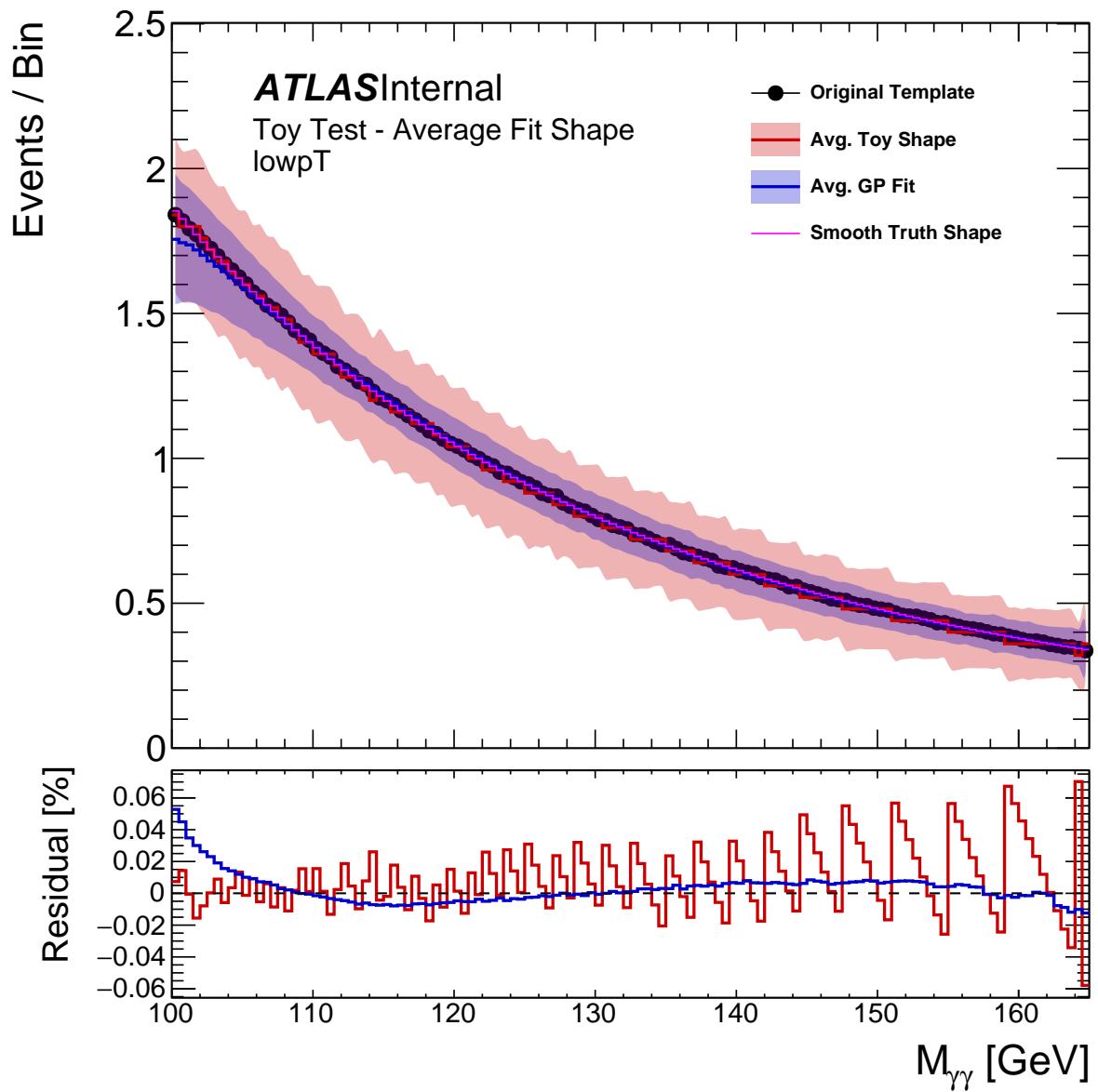


Figure A.201: The average toy distribution for both the GPR and raw template, using an expPoly2-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 2800 events.

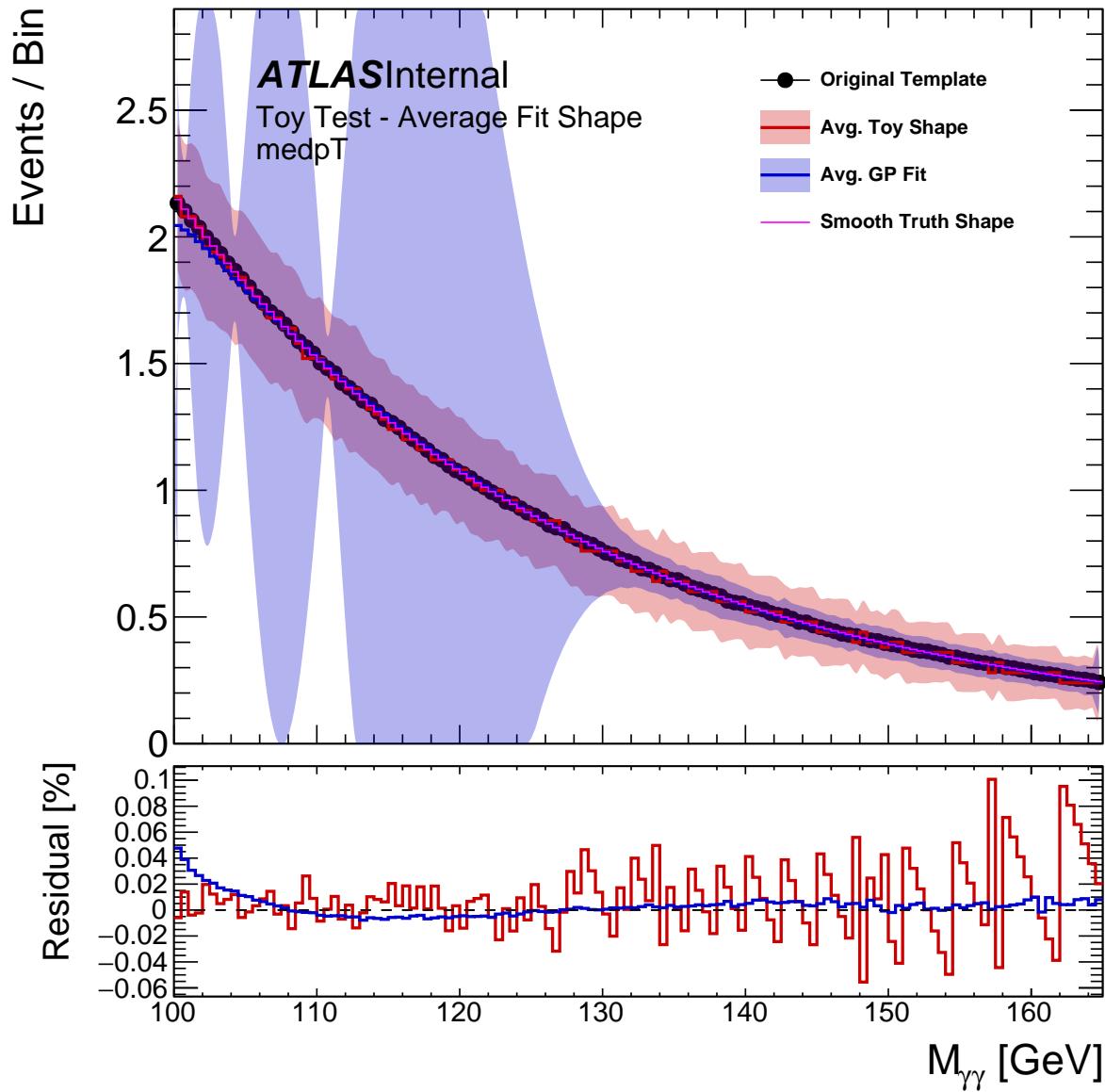


Figure A.202: The average toy distribution for both the GPR and raw template, using an expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 2800 events.

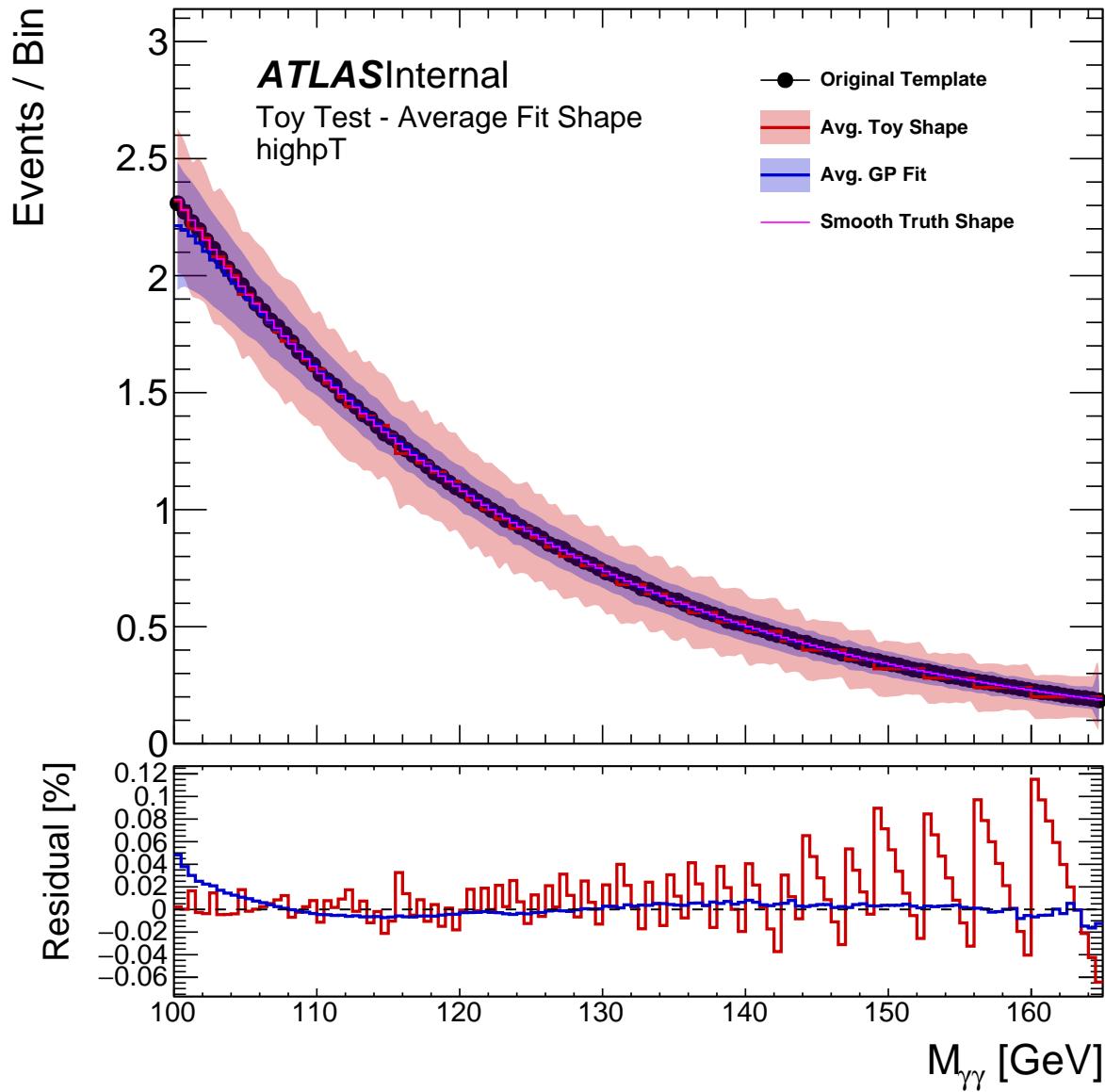


Figure A.203: The average toy distribution for both the GPR and raw template, using a different expPoly3-derived template extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 2800 events.

A.5.4 Feature Injection Study

In addition to showing that, with the correct choice of noise kernel, GPR introduces no bias in templates without features, we show that GPR is not severely hampered when real features are present in the data. Though this study is not relevant in the context of the Couplings analysis, as we assume our underlying templates are smooth, it provides further proof of GPR’s power and applicability in different contexts.

For feature-injection tests, we inject a signal-like feature to the toy template. We then smooth the template using the GPR and extract the spurious signal from both the smoothed and unsmoothed toy, and compare them to see whether the two results differ substantially.

For the preliminary feature-width study, we use the templates from Category 1 and Category 31 of the Couplings analysis (GG2H_0J_PTH_0_10_0 and GG2H_PTH_300_450_0). We perform a functional form fit to each and use these as the basis for drawing toys; the normalization and statistics of each toy are fixed to those of the original template. We perform the feature-injection and the no-feature injection tests using these templates. We conclude that the bias increases as template statistics decrease and as feature width decreases; however, we note that the presence of true narrow background features in our templates is highly unphysical (as we assume our underlying distributions are generally smooth). Thus, we conclude that, for the estimation of the feature-injection bias, a three-times-signal-width feature (i.e., approximately 3 GeV) provides a decent trade-off between a conservative estimate of the bias and the ability to model a potential realistic feature. However, in these feature-injection tests, the width of the signal model used for the spurious signal test is kept at the nominal value of 1 GeV.

For this initial study, the error passed to the GP fit is the input error of the initial template (i.e., the linear noise kernel is not used) and no padding is used. Additionally, because we are simply attempting to choose an optimal width for injection, we fit with the same function we have used to generate the template.

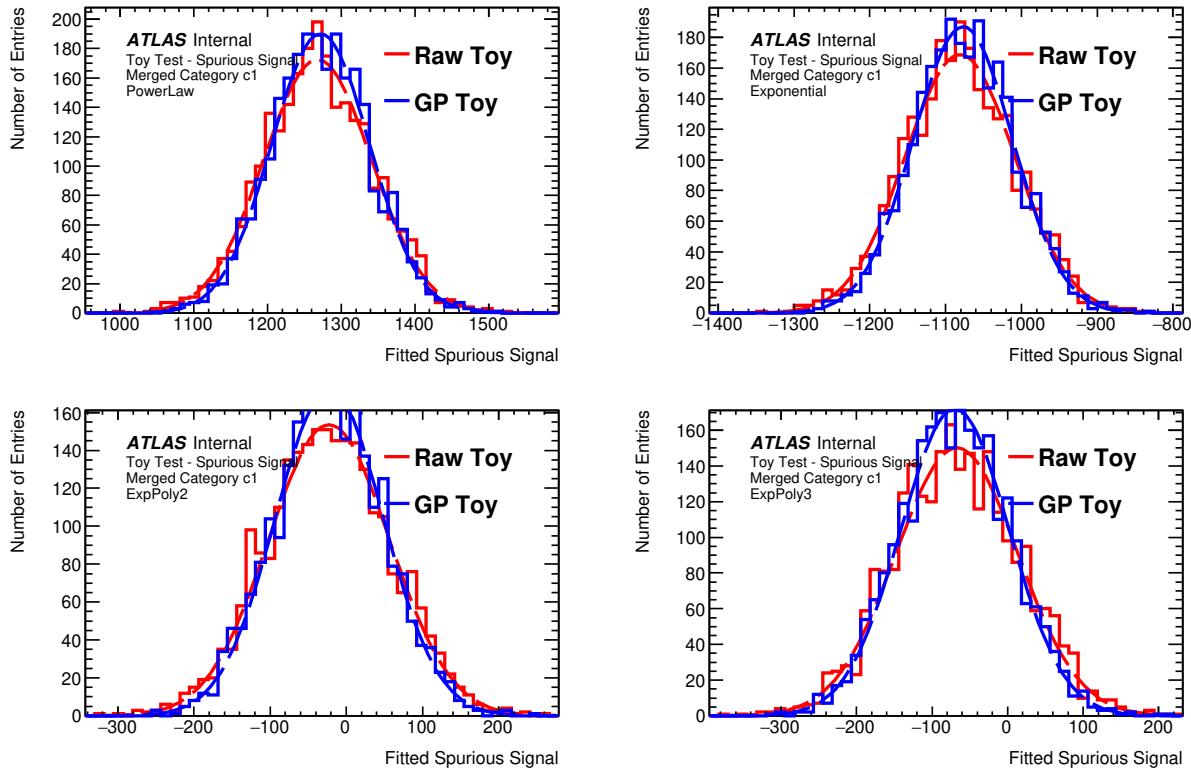


Figure A.204: The distribution of spurious signal for various functions for both the GPR and raw template, using the c1 template as a basis.

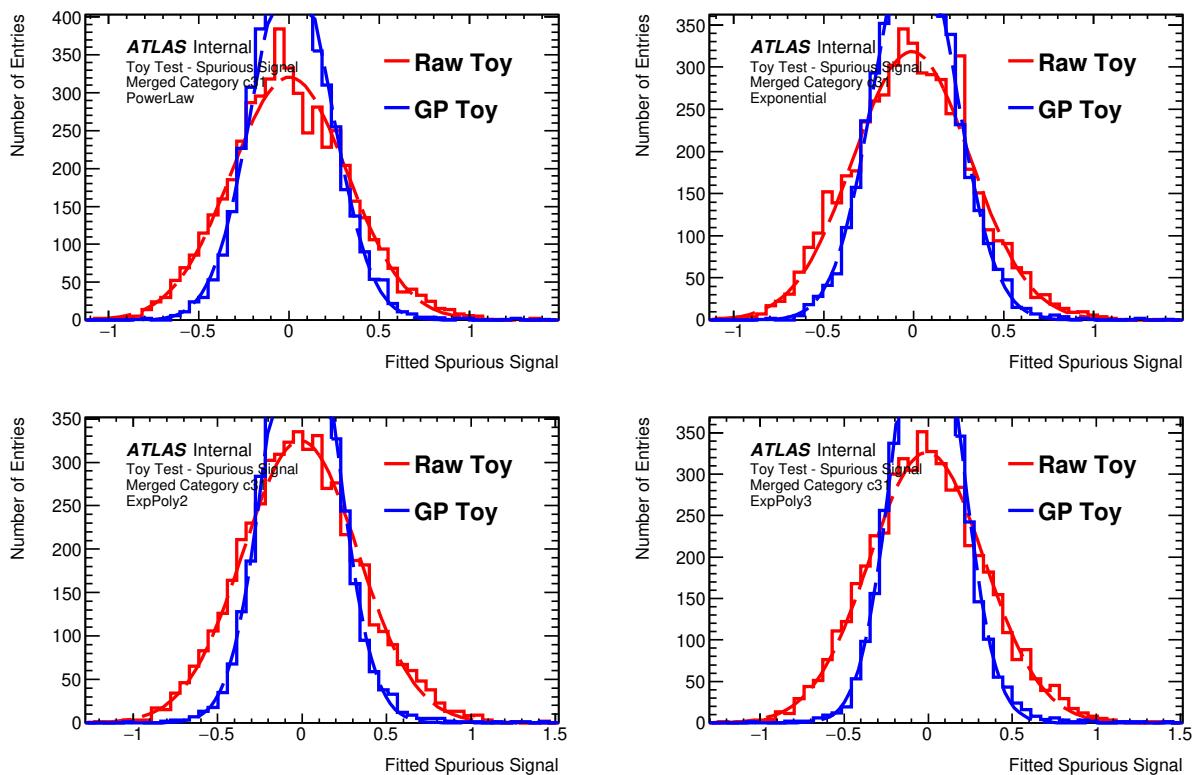


Figure A.205: The distribution of spurious signal for various functions for both the GPR and raw template, using the c31 template as a basis.

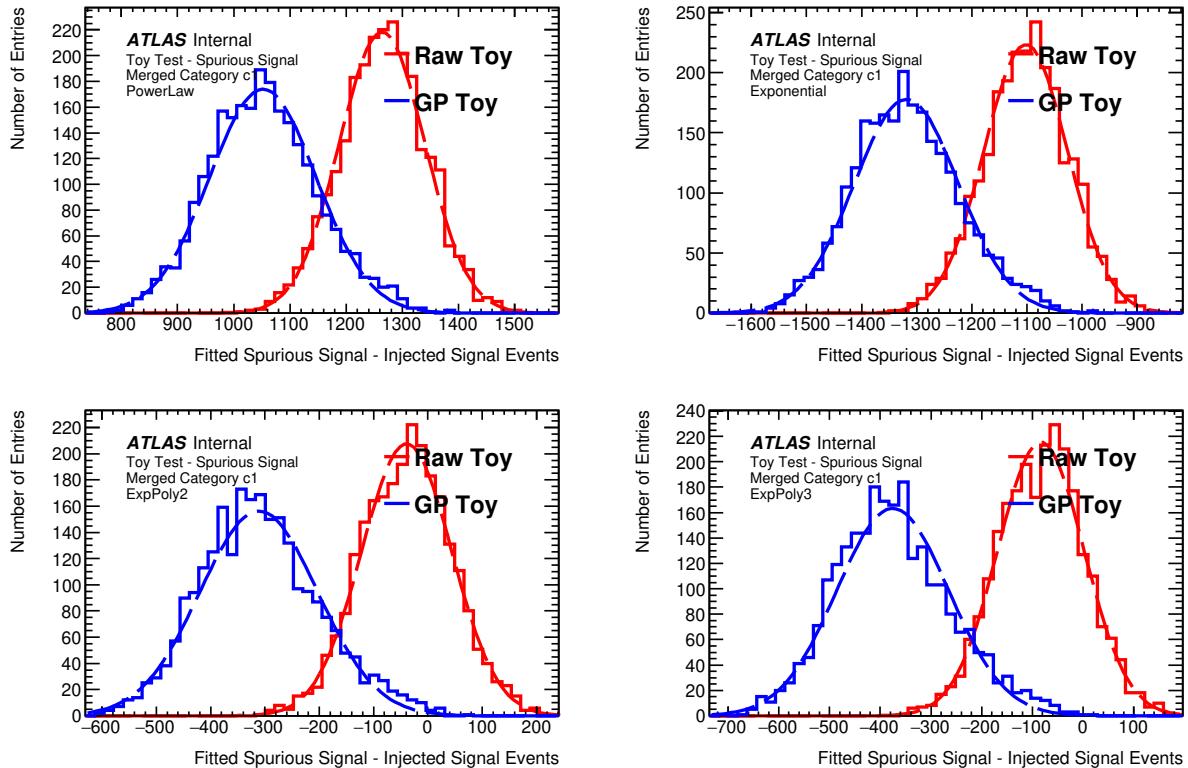


Figure A.206: The distribution of spurious signal for various functions for both the GPR and raw template, using the c1 template with a 1GeV wide feature injected as a basis.

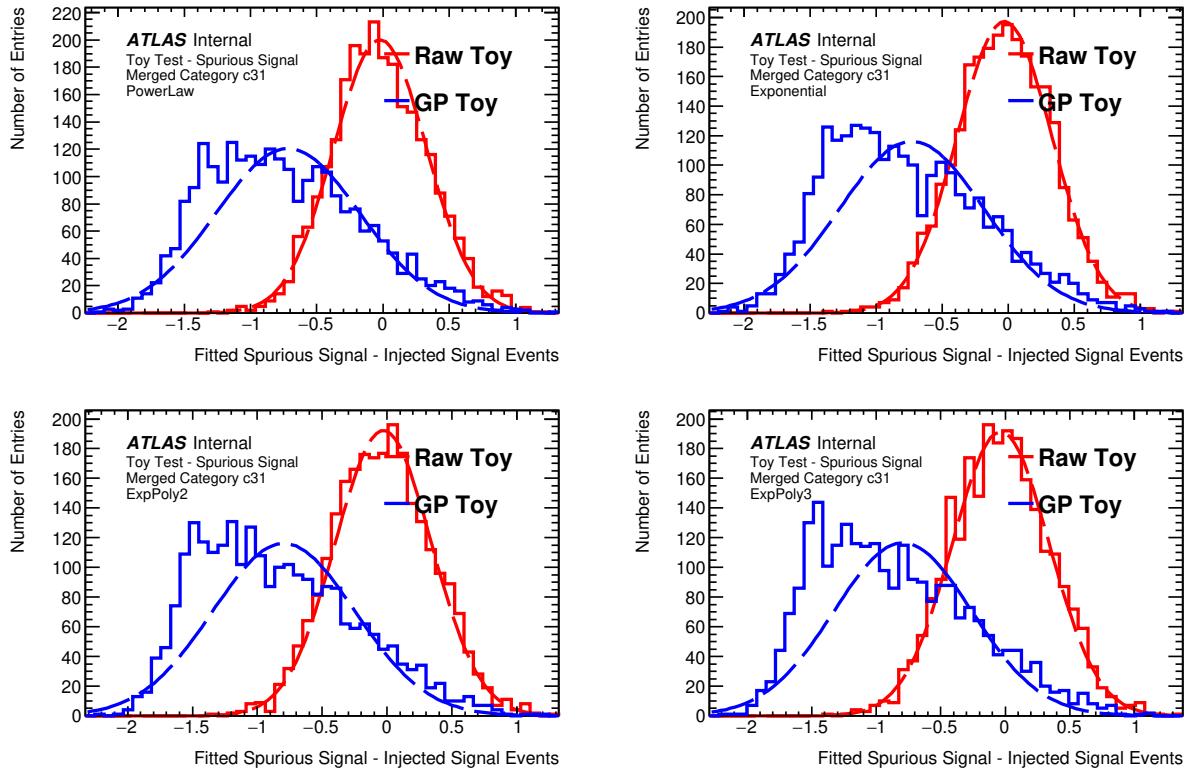


Figure A.207: The distribution of spurious signal for various functions for both the GPR and raw template, using the c31 template with a 1GeV-wide feature injected as a basis.

Category and Function	Nsig	Signal Width	Median SS, GPR	Median SS, Raw Toy	Bias	(GPR- Raw)/ NSig
C1, Exp2	0	1.02	-19.644	-19.776	0.132	N/A
C1, Exp2	2405.177	1.02	2084.351	2363.8	-279.449	-0.116
C1, Exp2	2405.177	3.062	631.187	632.204	-1.017	-0.0004
C1, Exp2	2405.177	5.104	-640.038	-644.867	4.829	0.002
C31, Exp	0	1.818	0.004	-0.009	0.013	N/A
C31, Exp	2.02	1.818	1.182	1.989	-0.807	-0.400
C31, Exp	2.02	5.453	0.583	0.829	-0.246	-0.122
C31, Exp	2.02	9.088	0.385	0.443	-0.058	-0.029

Table A.16: The median spurious signal extracted from a distribution of 1000 toys (in this study, we use the median rather than the mean to be robust to potential outliers; however, distributions are approximately Gaussian so the two generally do not disagree) for a variety of feature-injection widths.

Following this, we set the size of a feature as 3 times the standard-model signal width (i.e., approximately 3 GeV). Additionally, we fix the size of the feature injected as 1% of the number of events in the toy template (except for toys with 10 events, for which we inject a one-event feature). We show the spurious signal extracted for 10 event through 10M event toys for the three shapes, and record the bias. The results are reported in Table ??

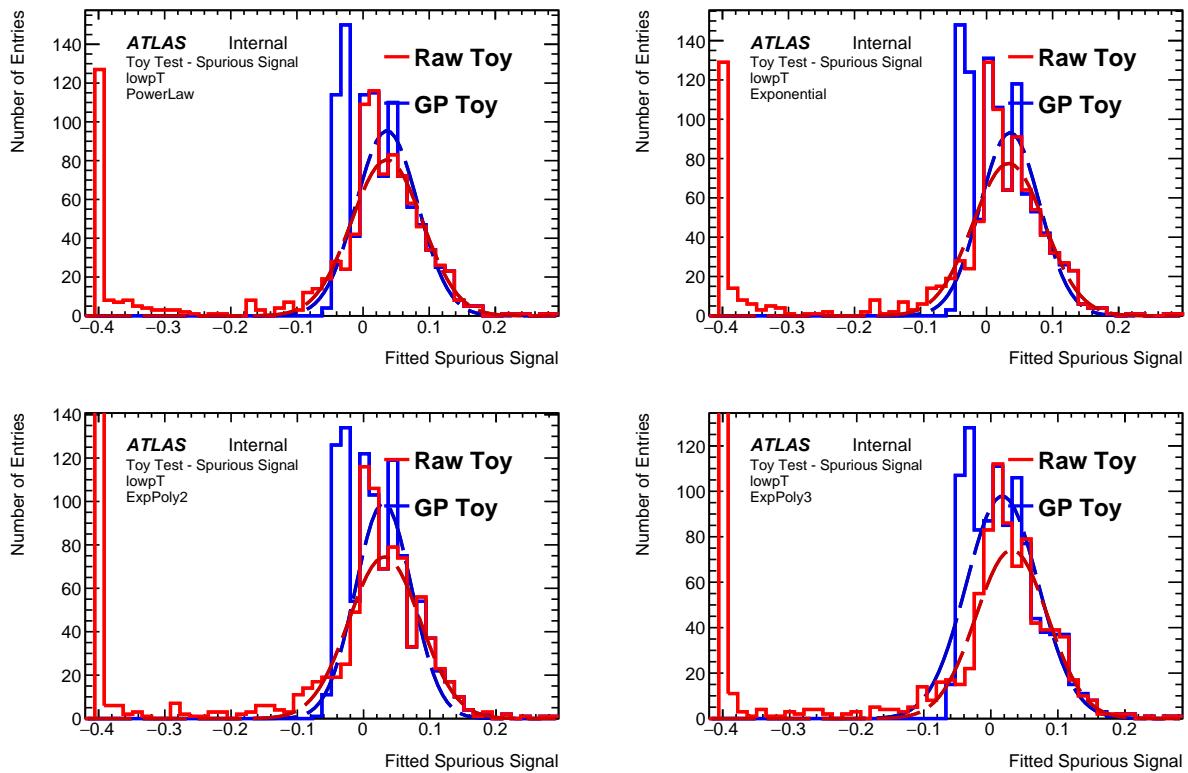


Figure A.208: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10 events.

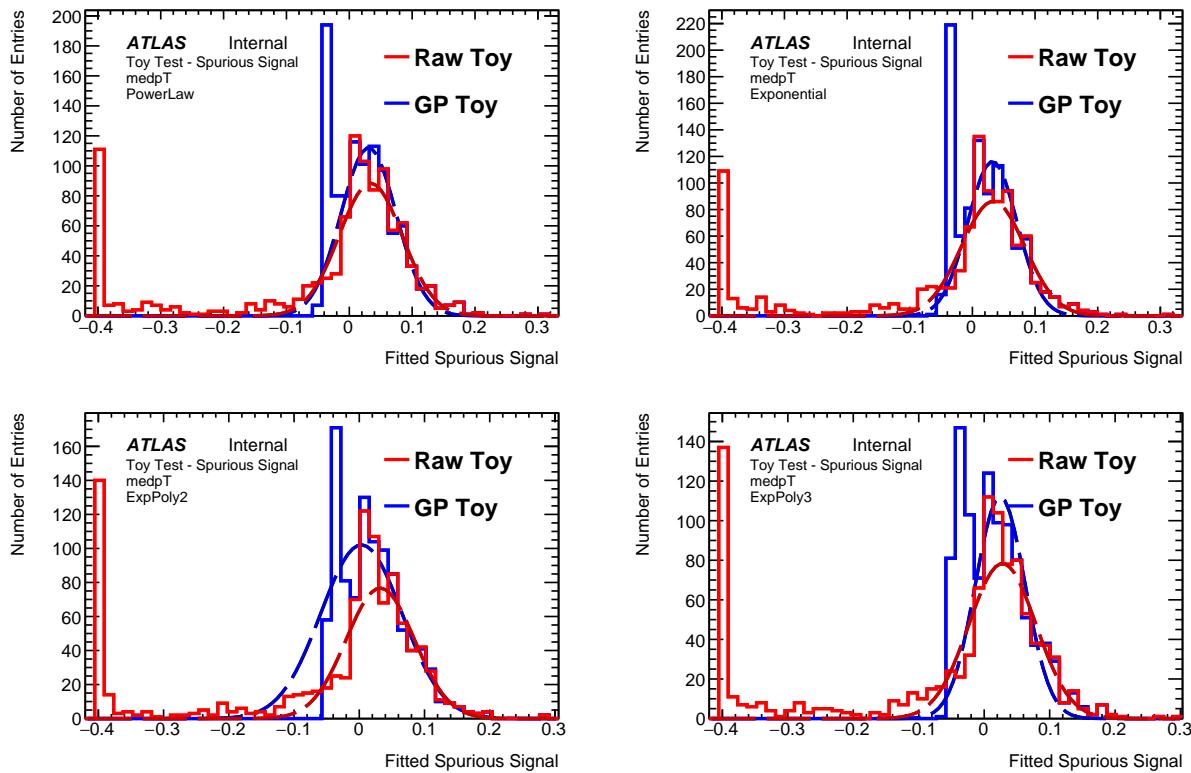


Figure A.209: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10 events.

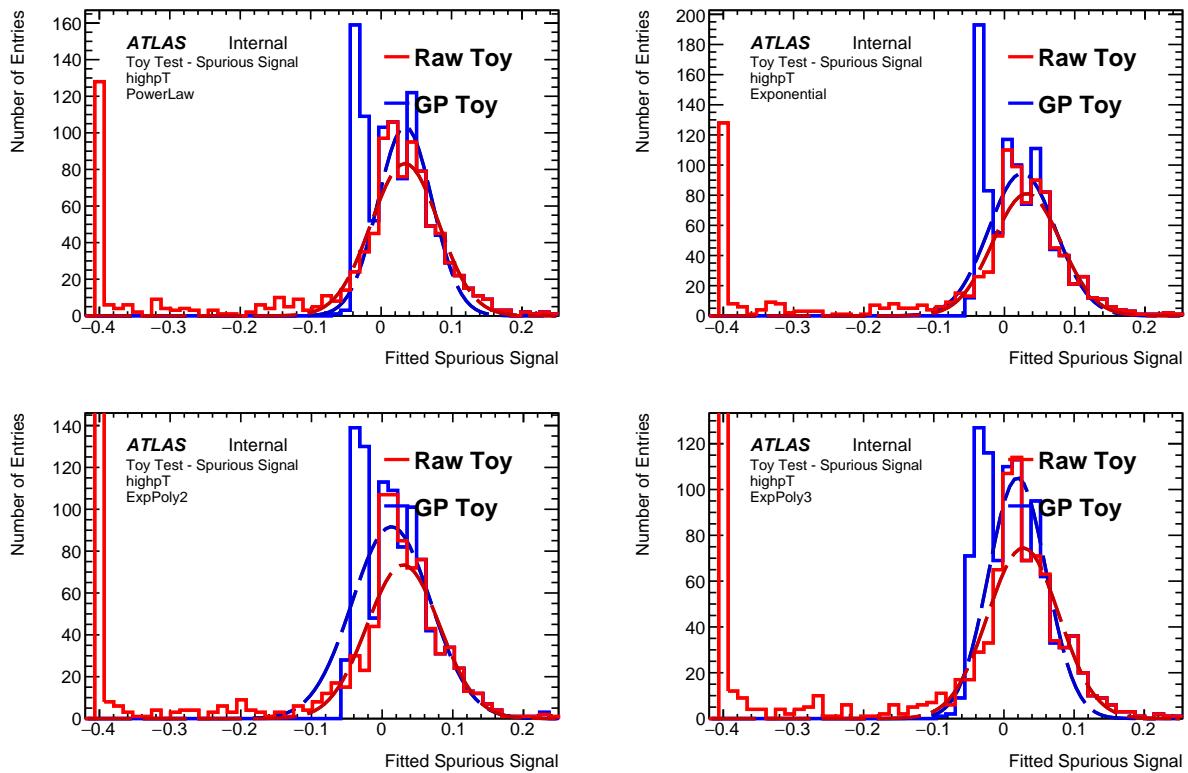


Figure A.210: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10 events.

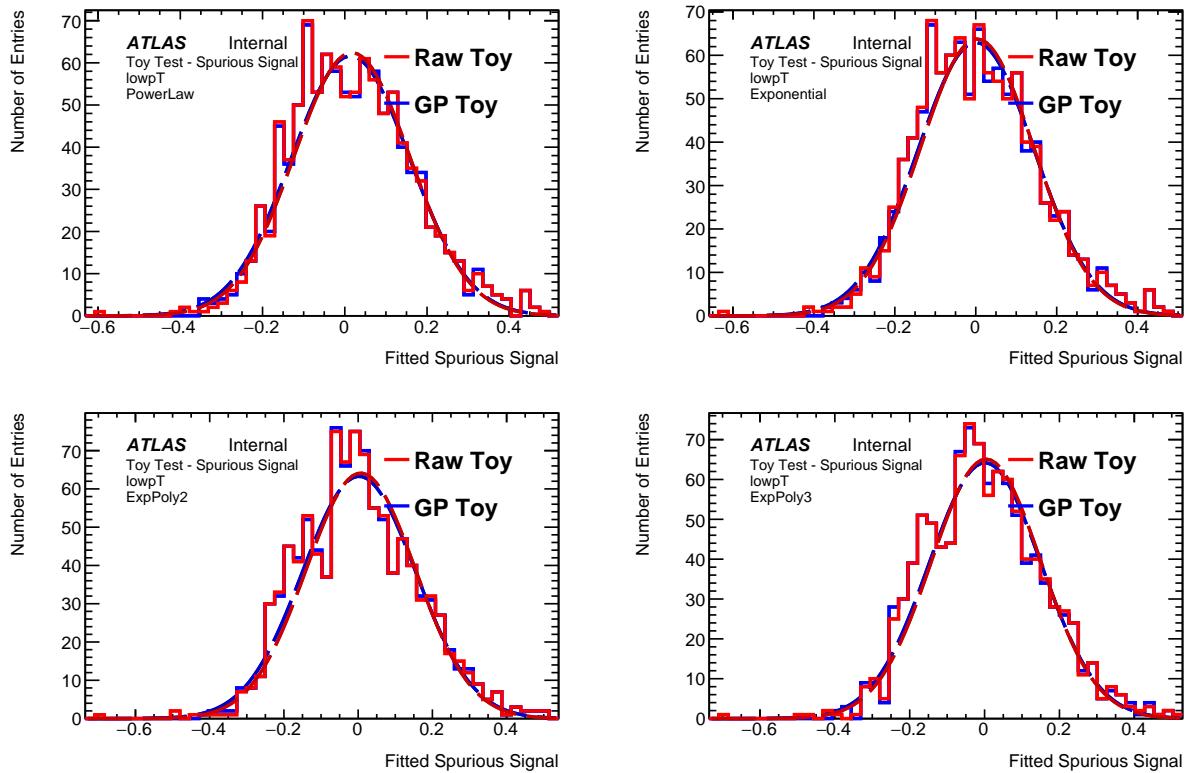


Figure A.211: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100 events.

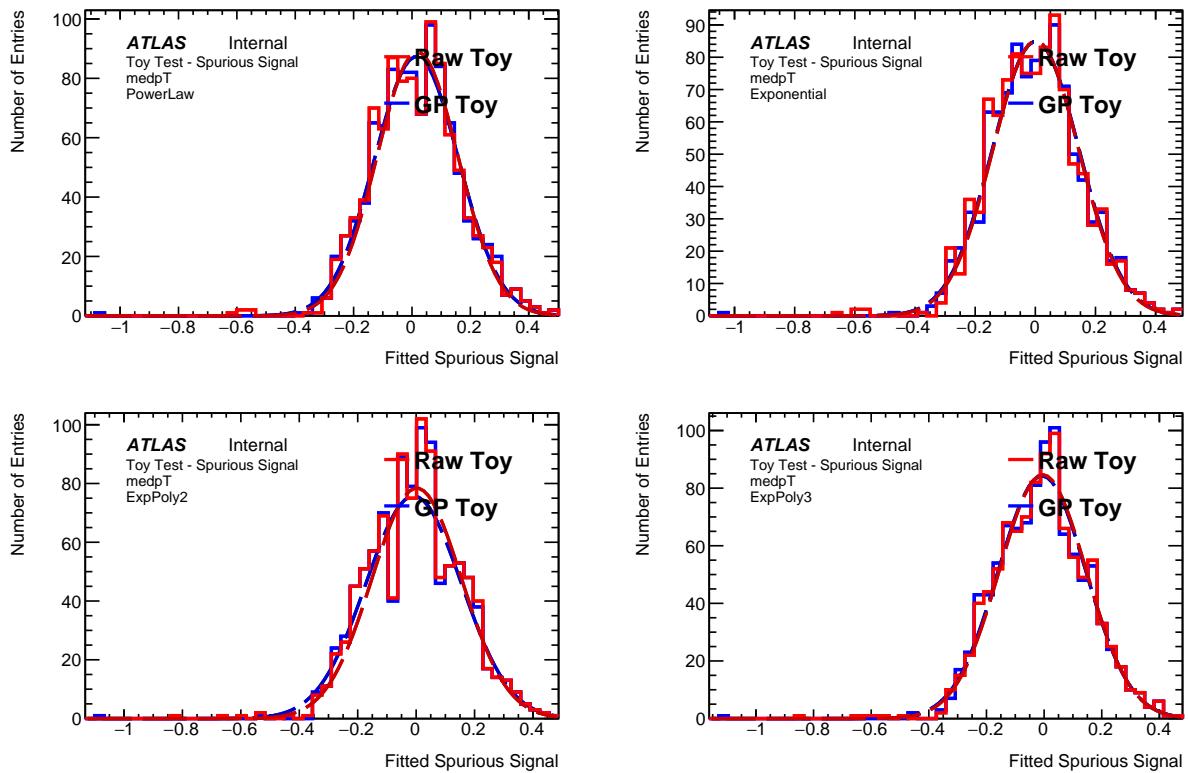


Figure A.212: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100 events.

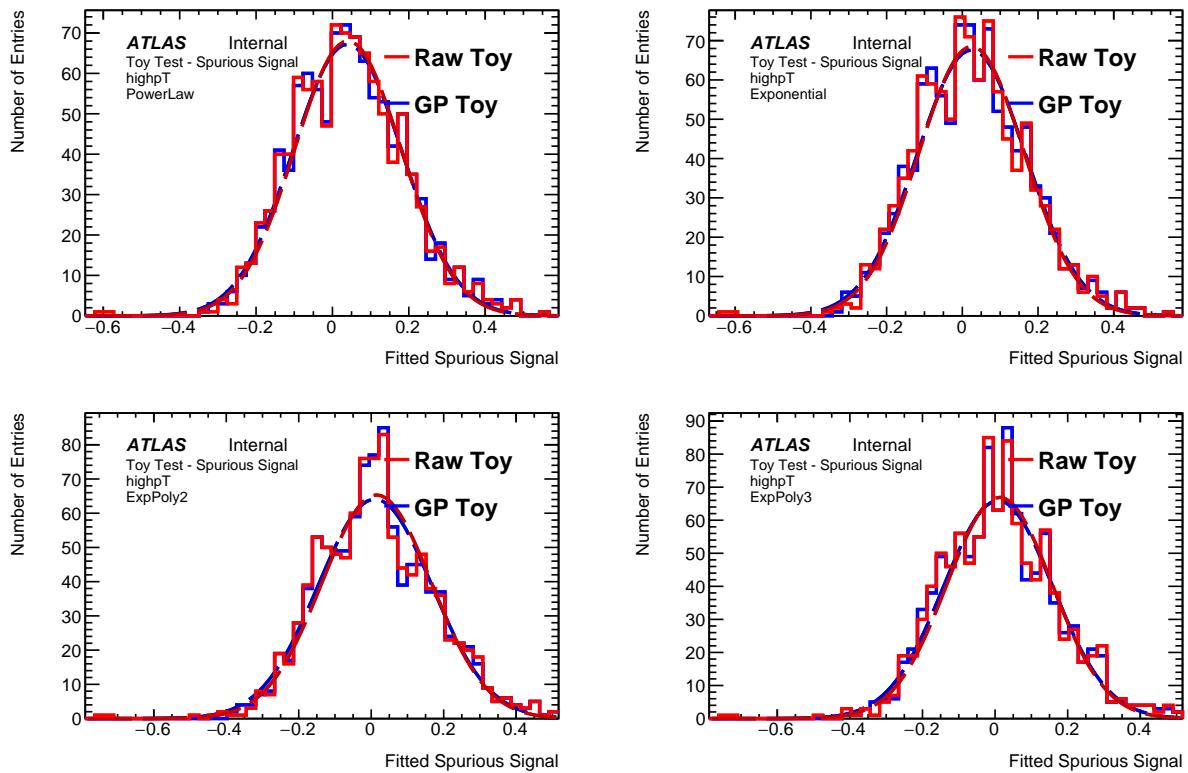


Figure A.213: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100 events.

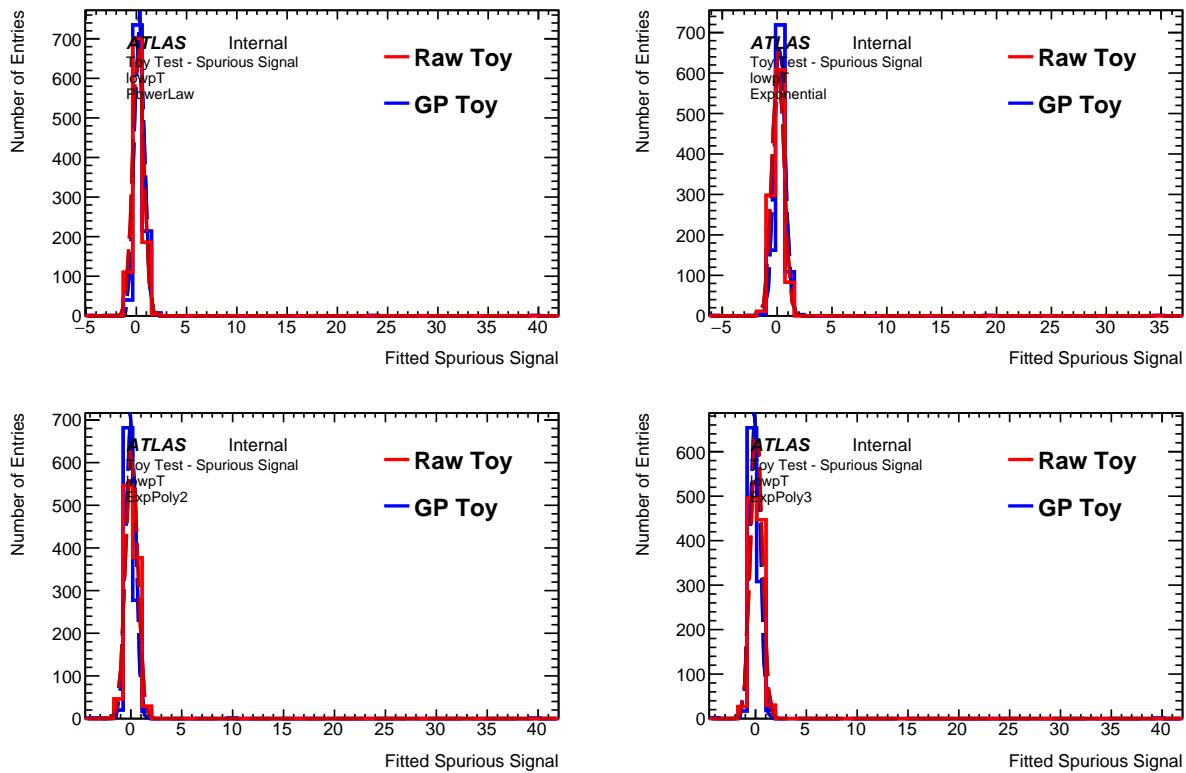


Figure A.214: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1000 events.

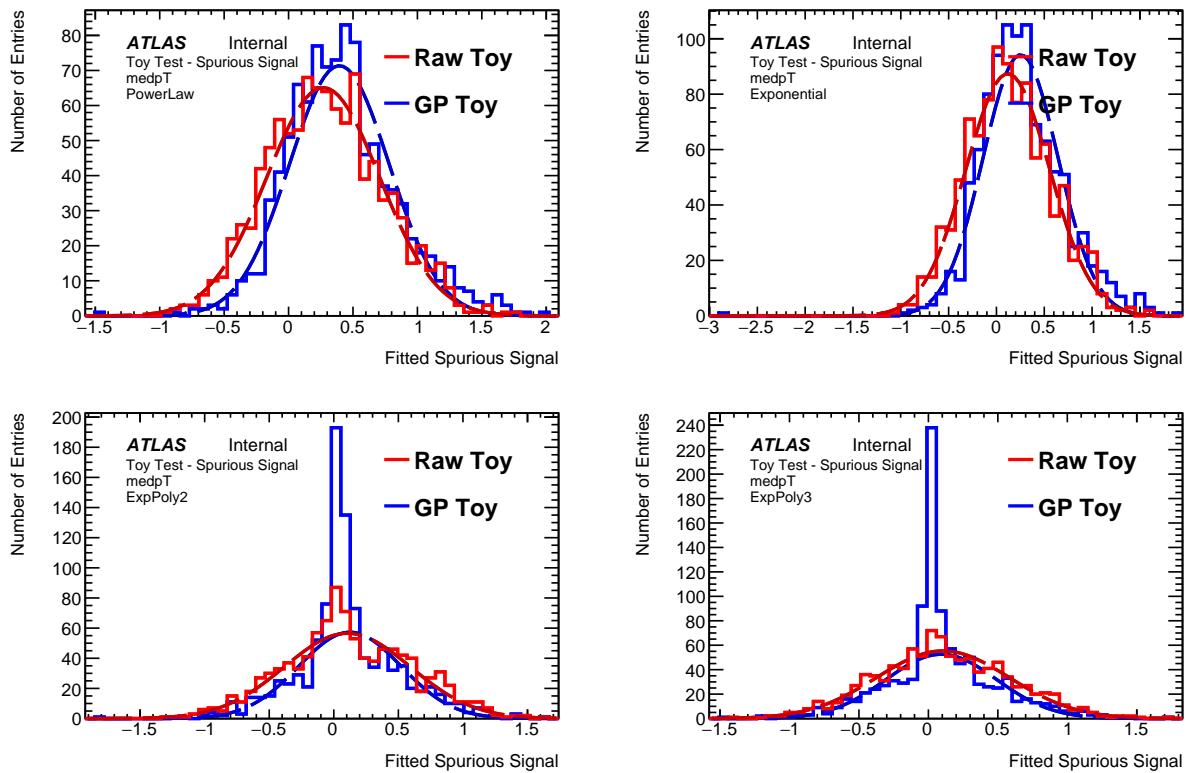


Figure A.215: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1000 events.

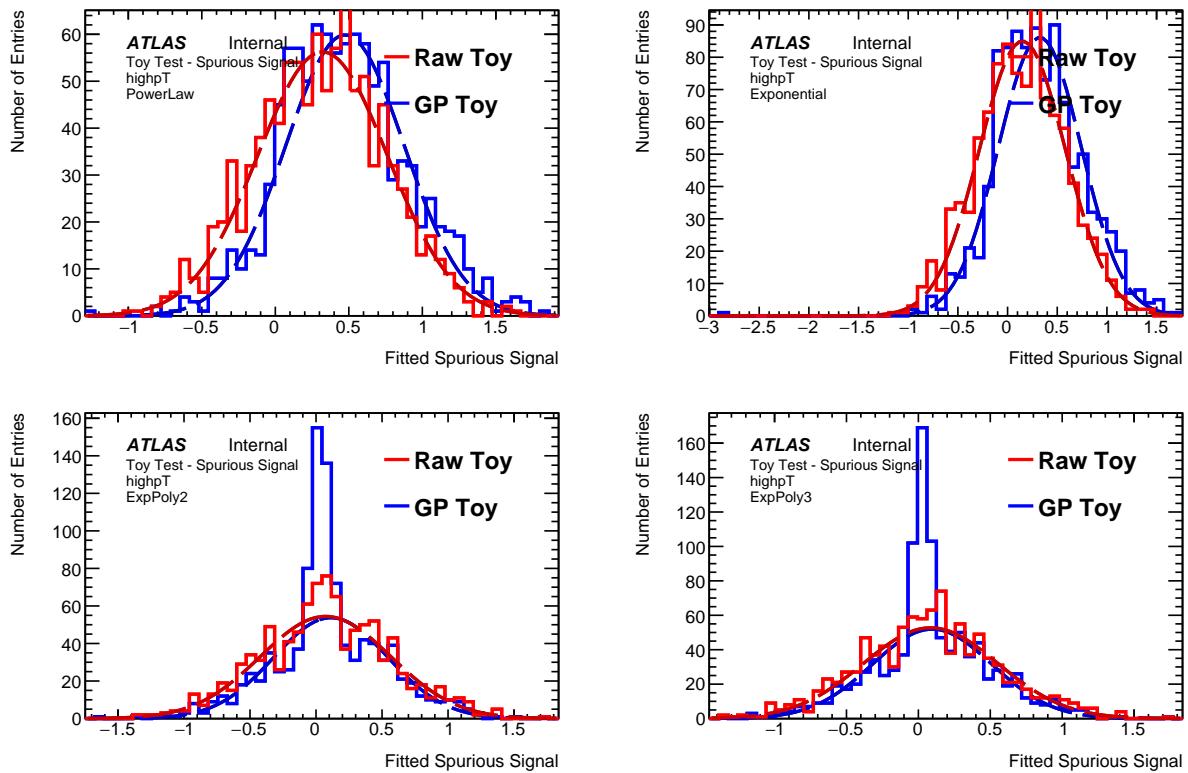


Figure A.216: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1000 events.

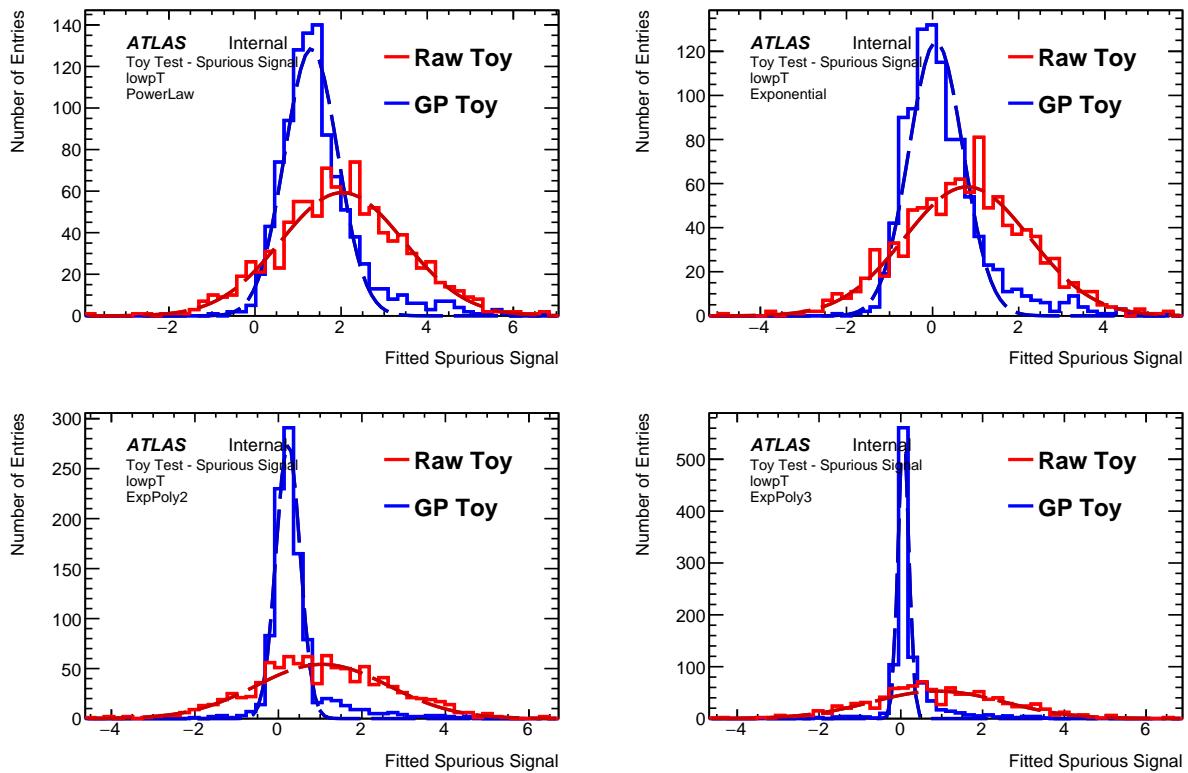


Figure A.217: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10k events.

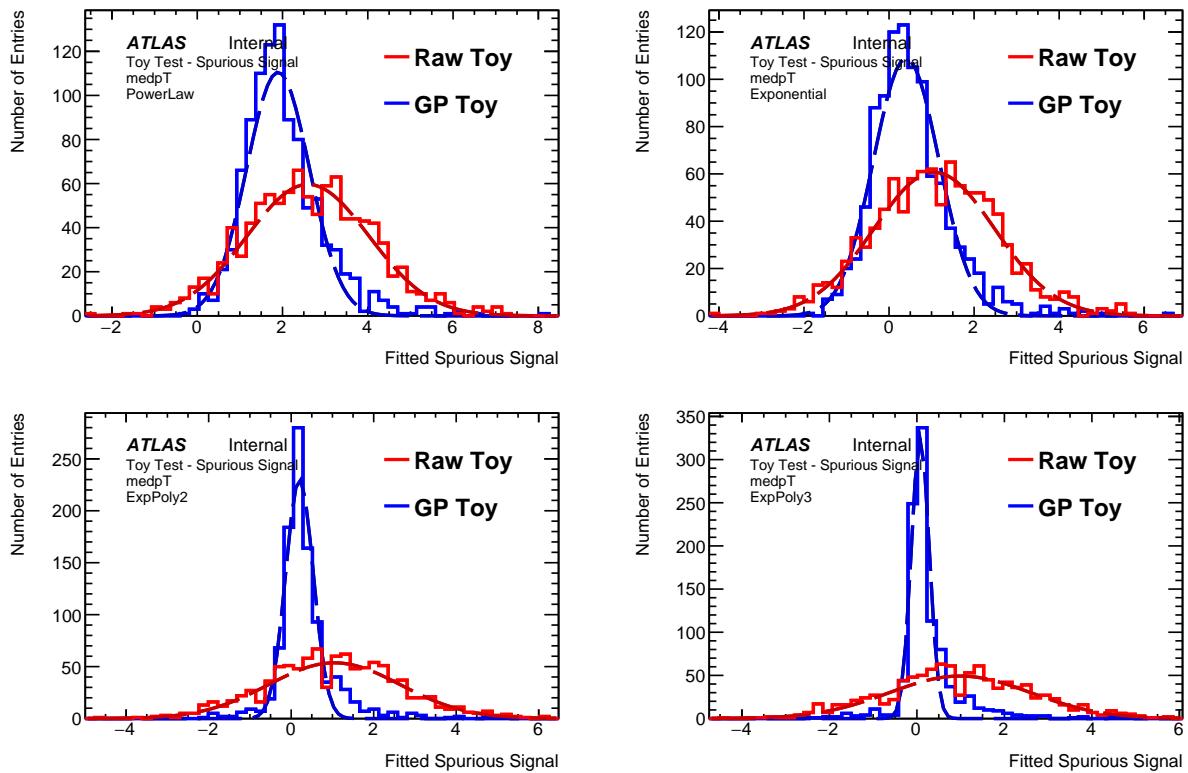


Figure A.218: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10k events.

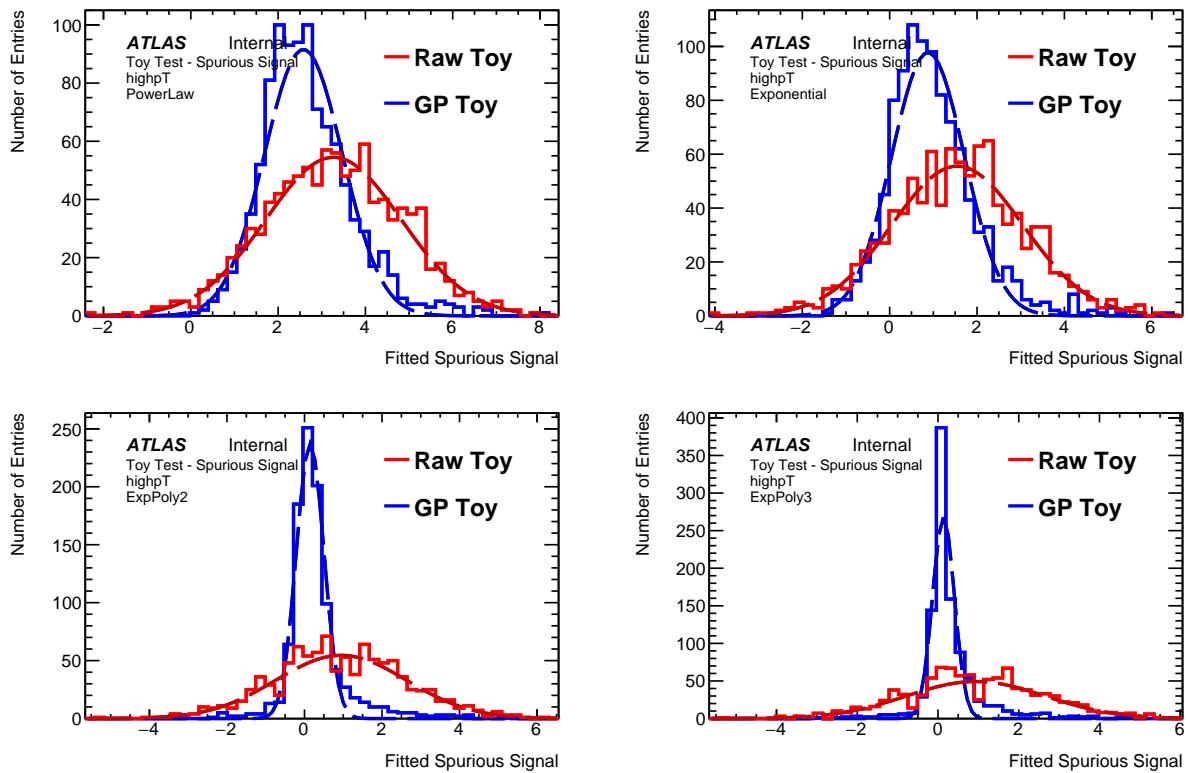


Figure A.219: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10k events.

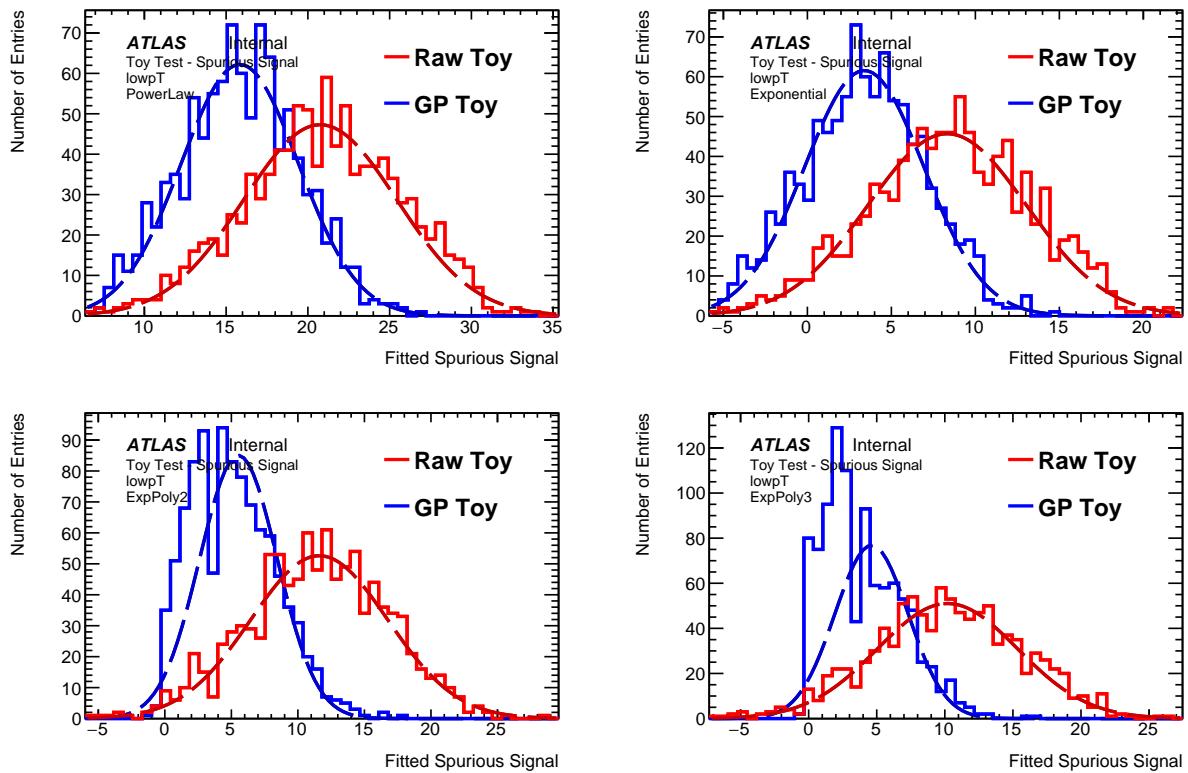


Figure A.220: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100k events.

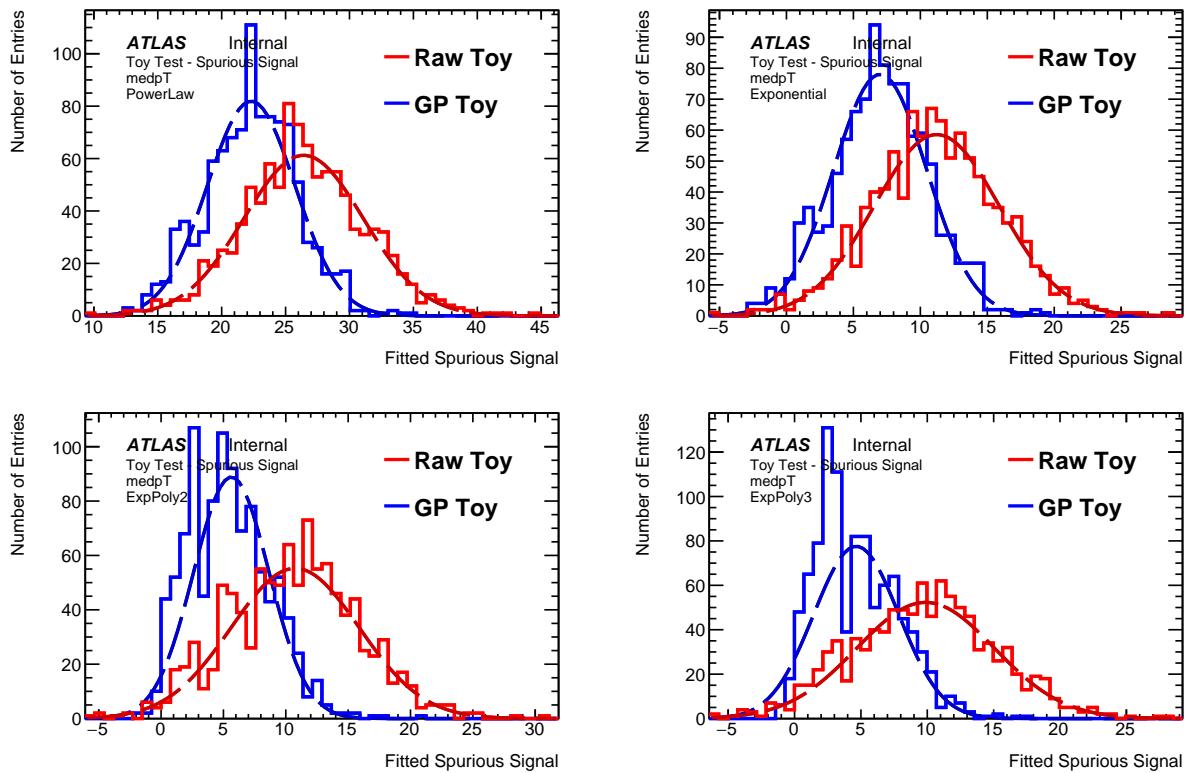


Figure A.221: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100k events.

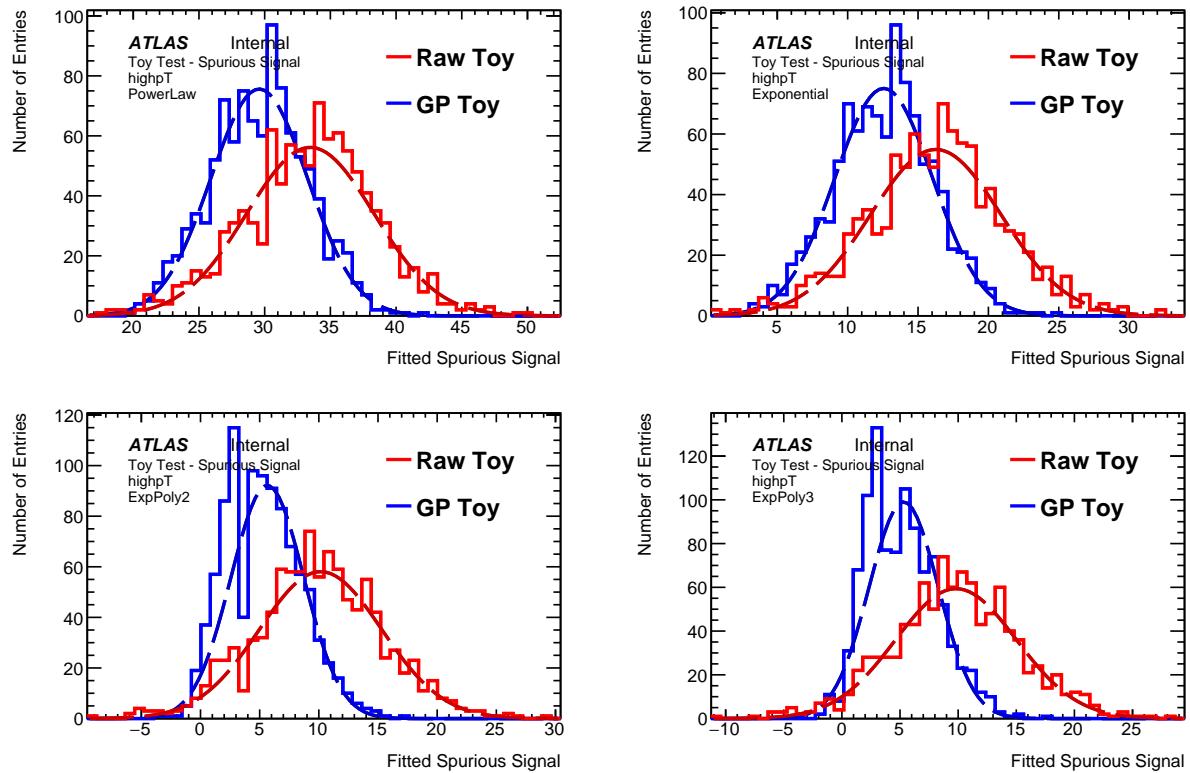


Figure A.222: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100k events.

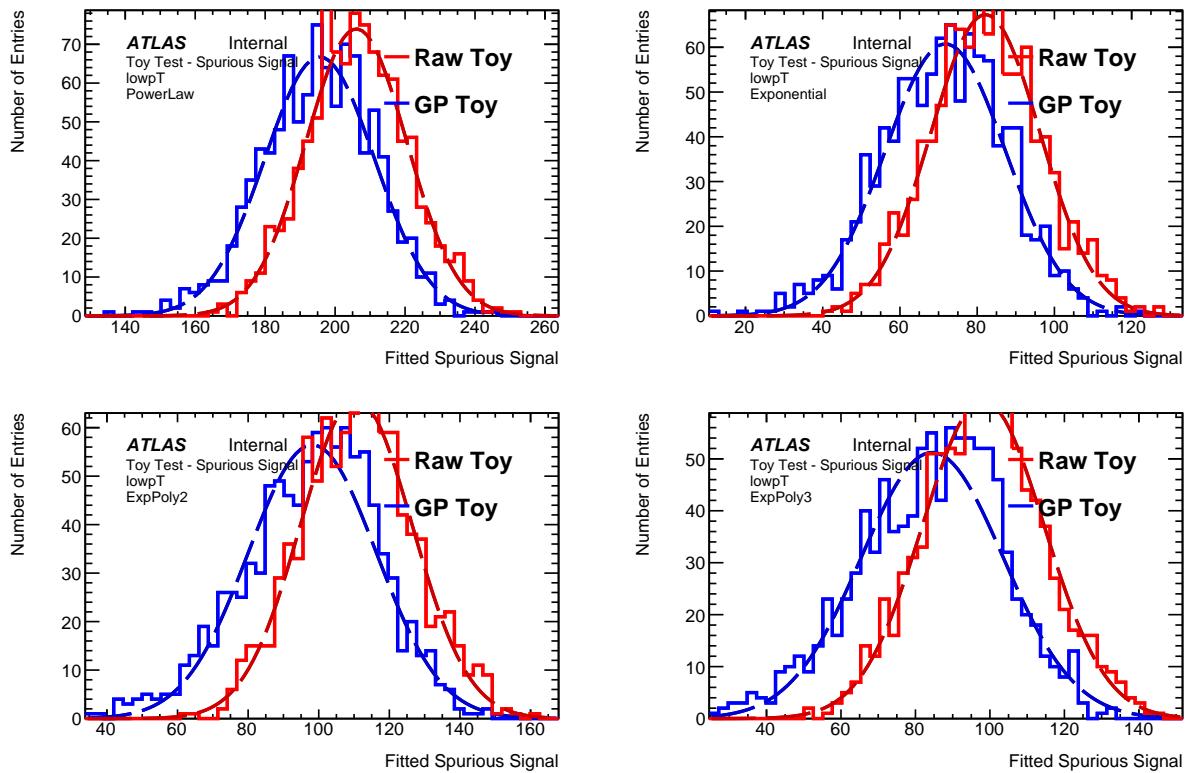


Figure A.223: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

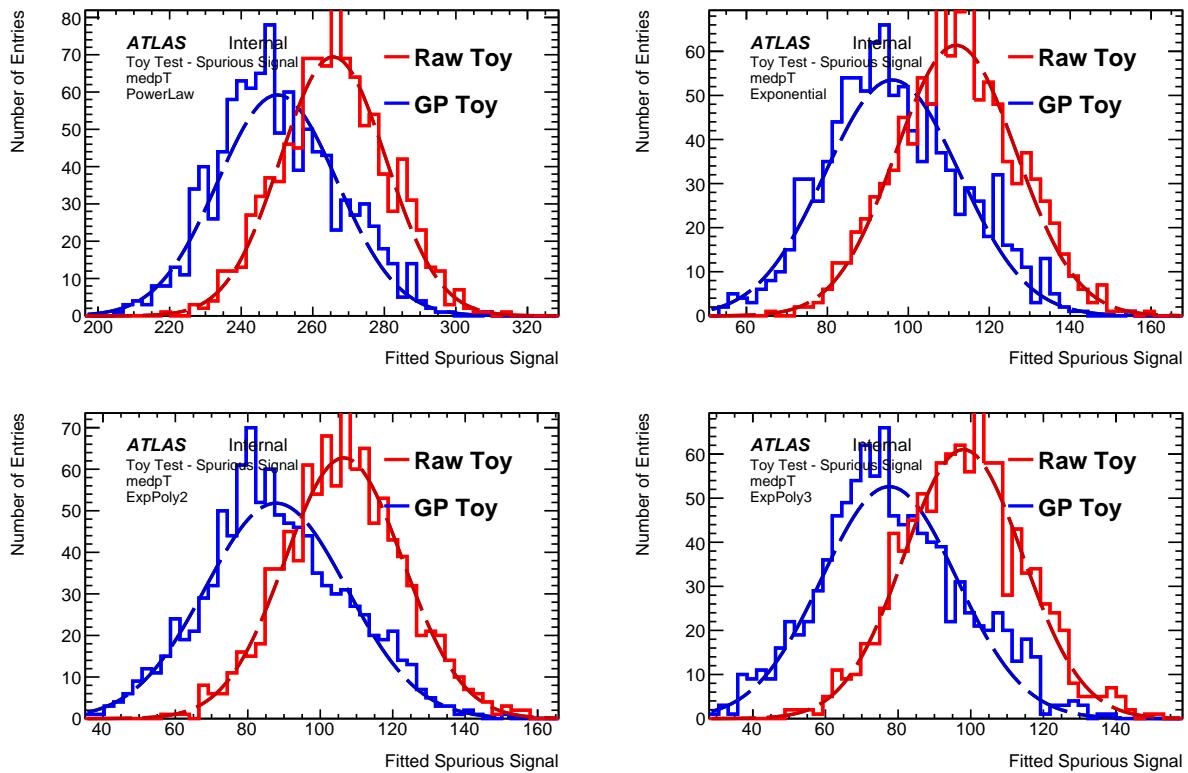


Figure A.224: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

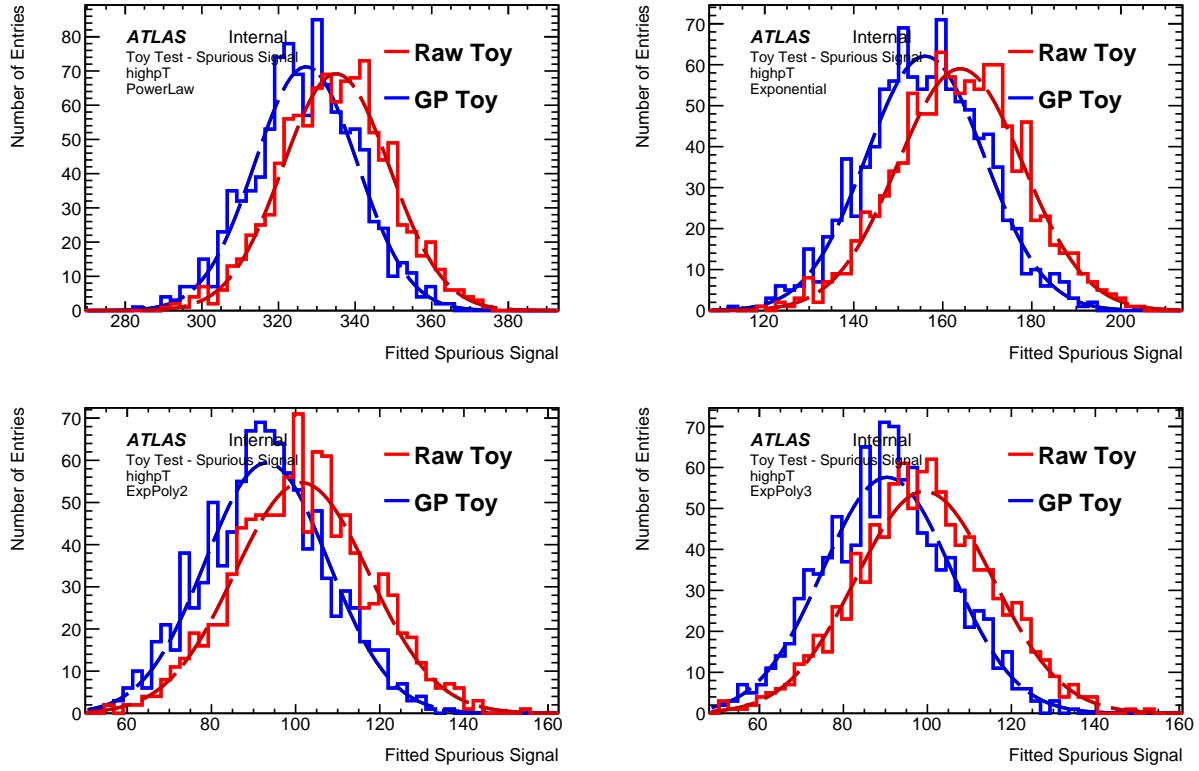


Figure A.225: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

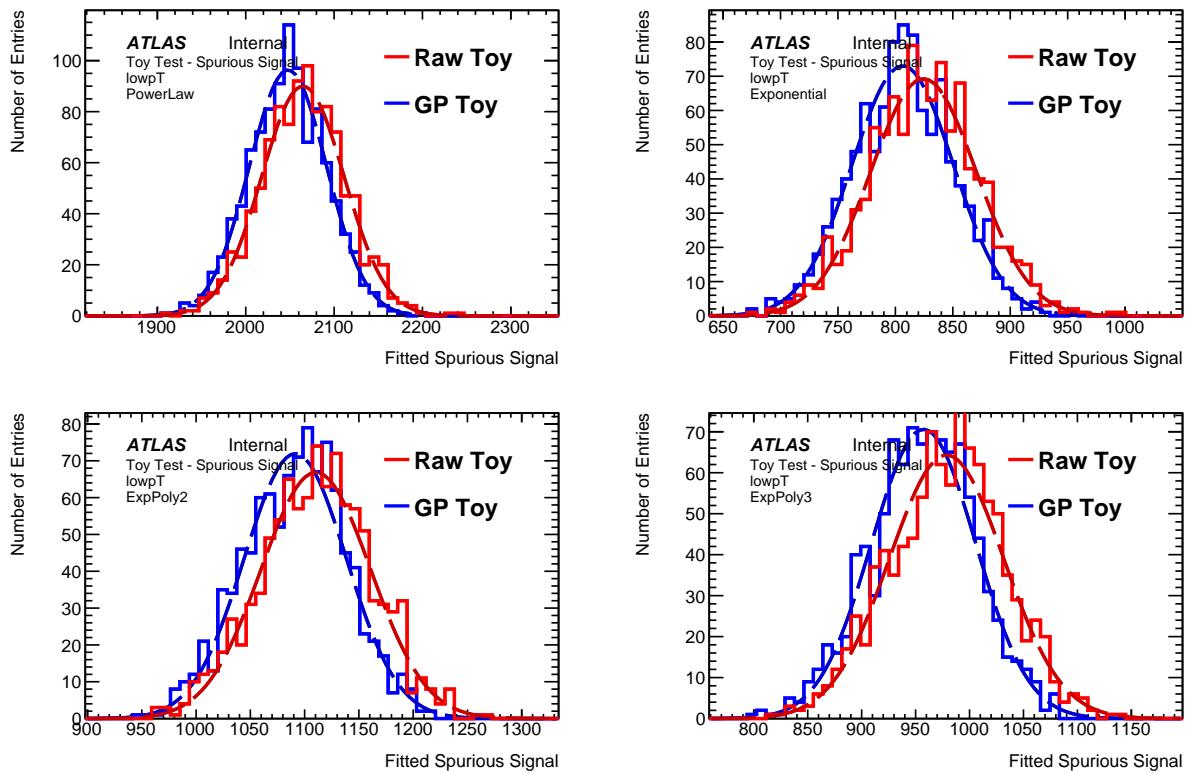


Figure A.226: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10,000,000 events.

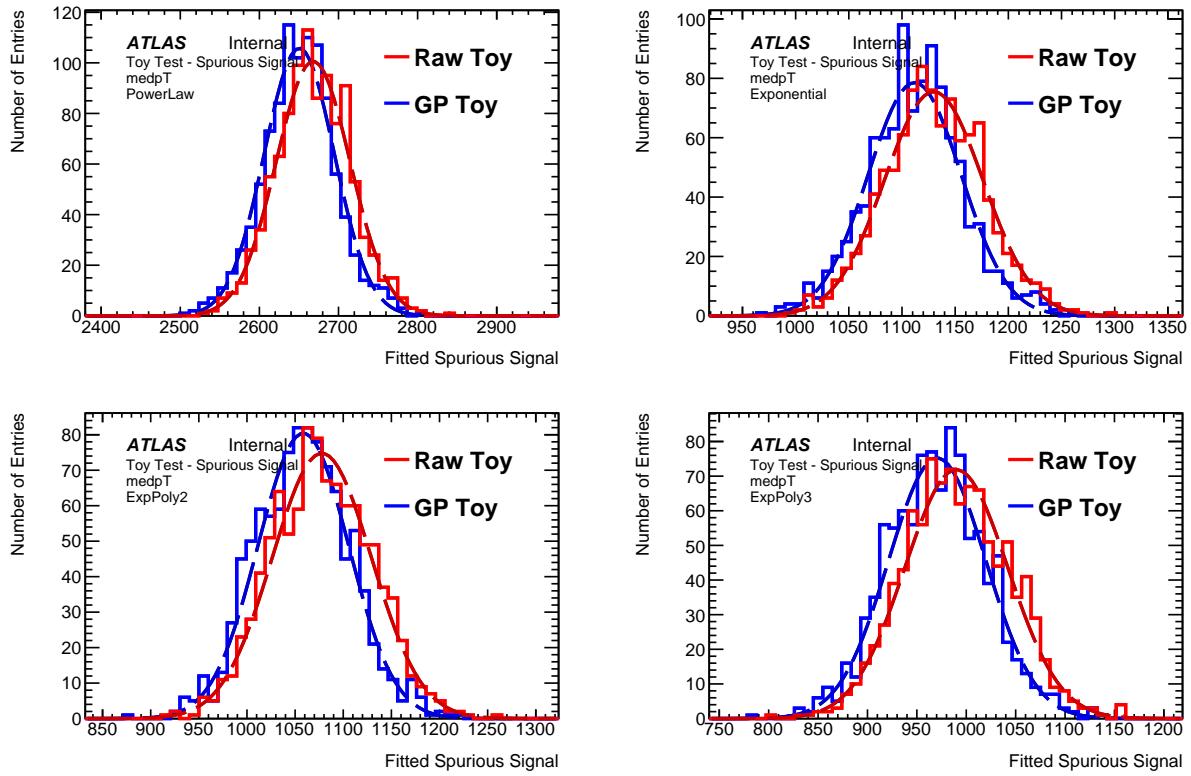


Figure A.227: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10,000,000 events.

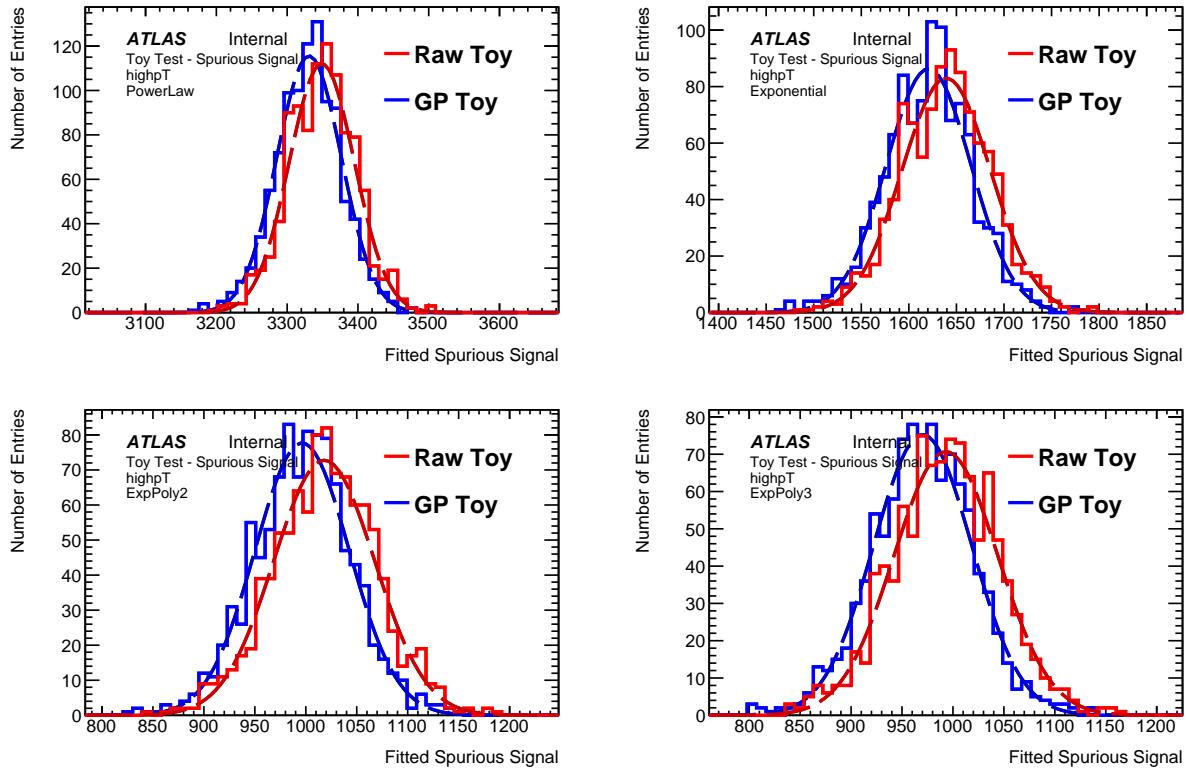


Figure A.228: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 3 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10,000,000 events.

Nominal	N_sig	Unweighted Events	Bkg events weighted	Mean SS_raw	Mean SS_GPR	GP-Raw: Bias, Mean	Sigma Toy	Sigma GPR	Bias/sigma(SS_raw)	Bias/Feature Size
Fit: Exp										
Generating: Exp2										
low10	0.04	10^1	0.4	-0.08	0.02	-0.07	0.17	0.05	-0.41	-1.69
low100	0.04	10^2	4	-0.04	0.01	-0.00	0.15	0.15	-0.01	-0.02
low1k	0.4	10^3	4×10^1	-0.33	0.24	-0.17	0.46	1.35	-0.38	-0.44
low10k	4	10^4	4×10^2	-3.21	0.25	0.54	1.47	0.95	0.37	0.14
low100k	40	10^5	4×10^3	-31.54	3.41	5.05	4.80	3.58	1.05	0.13
low1M	400	10^6	4×10^4	-317.66	71.31	11.03	14.35	15.92	0.77	0.03
low10M	4000	10^7	4×10^5	-3175.42	806.91	17.66	47.24	43.85	0.37	0.00
Fit: Exp2										
Generating: Exp3										
med10	0.04	10^1	0.4	-0.09	0.02	-0.07	0.17	0.05	-0.43	-1.82
med100	0.04	10^2	4	-0.04	-0.00	0.00	0.16	0.16	0.00	0.01
med1k	0.4	10^3	4×10^1	-0.29	0.11	0.00	0.49	0.40	0.01	0.01
med10k	4	10^4	4×10^2	-2.98	0.34	0.69	1.62	0.72	0.42	0.17
med100k	40	10^5	4×10^3	-29.41	5.43	5.16	5.22	3.26	0.99	0.13
med1M	400	10^6	4×10^4	-293.40	87.99	18.61	16.39	19.35	1.14	0.05
med10M	4000	10^7	4×10^5	-2922.56	1058.21	19.23	51.64	49.11	0.37	0.00
Fit: Exp2										
Generating: Exp3										
high10	0.04	10^1	0.4	-0.10	0.02	-0.08	0.17	0.05	-0.44	-1.90
high100	0.04	10^2	4	-0.03	0.01	-0.00	0.16	0.16	-0.01	-0.02
high1k	0.4	10^3	4×10^1	-0.33	0.10	-0.02	0.49	0.42	-0.04	-0.05
high10k	4	10^4	4×10^2	-3.08	0.29	0.63	1.70	0.78	0.37	0.16
high100k	40	10^5	4×10^3	-30.06	5.30	4.63	5.41	3.20	0.86	0.12
high1M	400	10^6	4×10^4	-298.37	92.98	8.65	15.86	14.54	0.55	0.02
high10M	4000	10^7	4×10^5	-2981.42	996.50	22.08	51.14	48.57	0.43	0.01

Table A.17: Spurious signal means and widths for the three test functional-form distributions for a range of different template statistics, with a signal feature injection that is approximately 3 GeV wide and 1% of the template integral.

The feature-injection bias for a three-sigma feature does not change appreciably with template shape (that is, for a given statistics level, the bias is approximately the same for all three templates). However, at high stats, the bias / feature size drops off as a function of template statistics. This makes sense, as the presence of true underlying features in high statistics templates is not compatible with the assumption that our true templates are smoothly falling functions. However, for templates containing greater than 20 effective background MC events per bin prior to feature injection (that is, those in the statistics range we conclude that it is safe to use GPR in), the measured bias is less than 18% of the injected feature size.

As a further check, we investigate the bias when the number of background events is fixed, but the size of the injected 3 GeV wide feature is allowed to vary. We see that, if the feature is very small or very large, the bias is small- the feature is either completely smoothed out (if it is small) or completely preserved (if it is large). In the intermediate range (feature integral $\zeta = 0.1\%$ of the total template integral), the bias varies, but is consistently less than 18% of the expected integral of the feature. We fix the toy template size at 1,000,000 events.

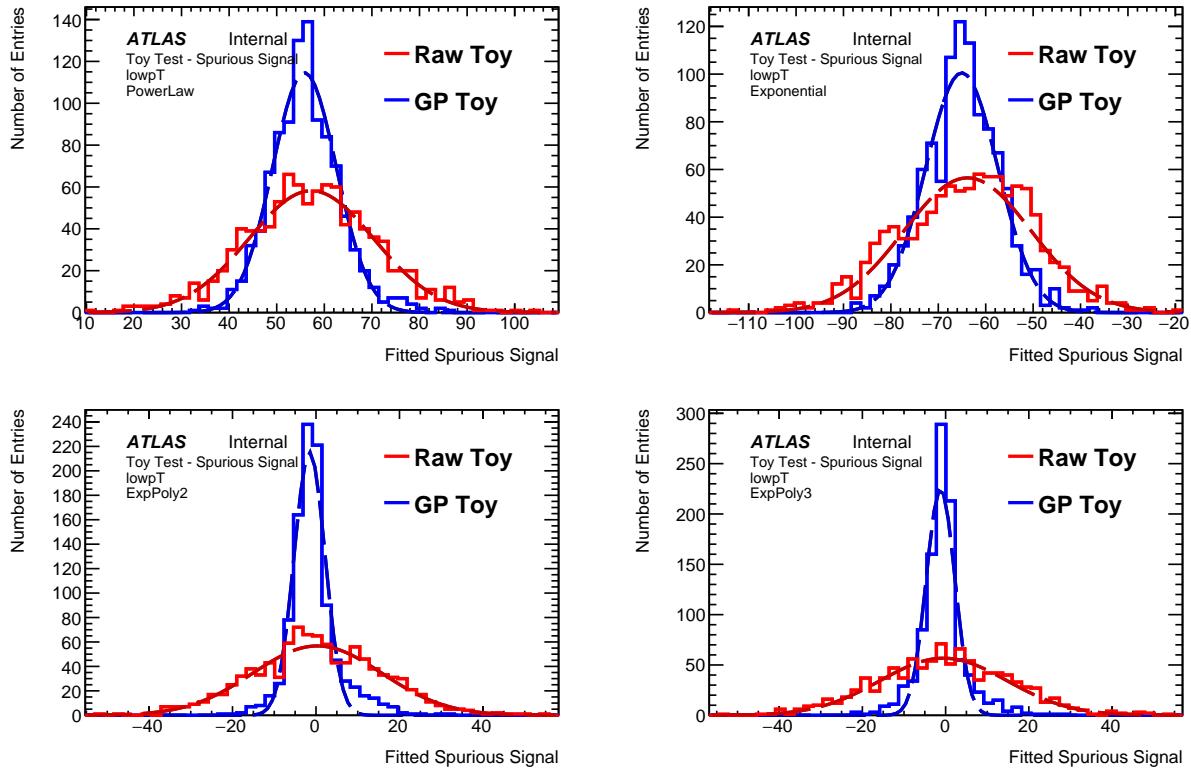


Figure A.229: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 3 GeV wide feature that is 0.01% of the template integral injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

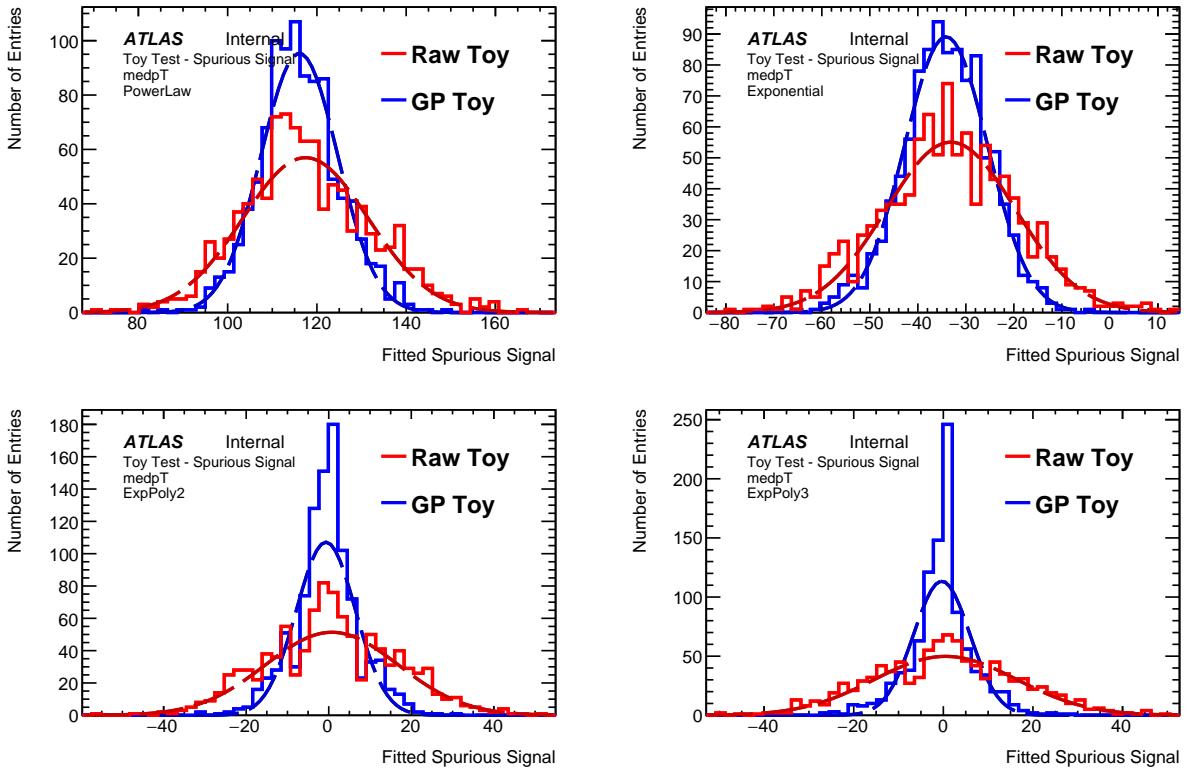


Figure A.230: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 3 GeV wide feature that is 0.01% of the template integral injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

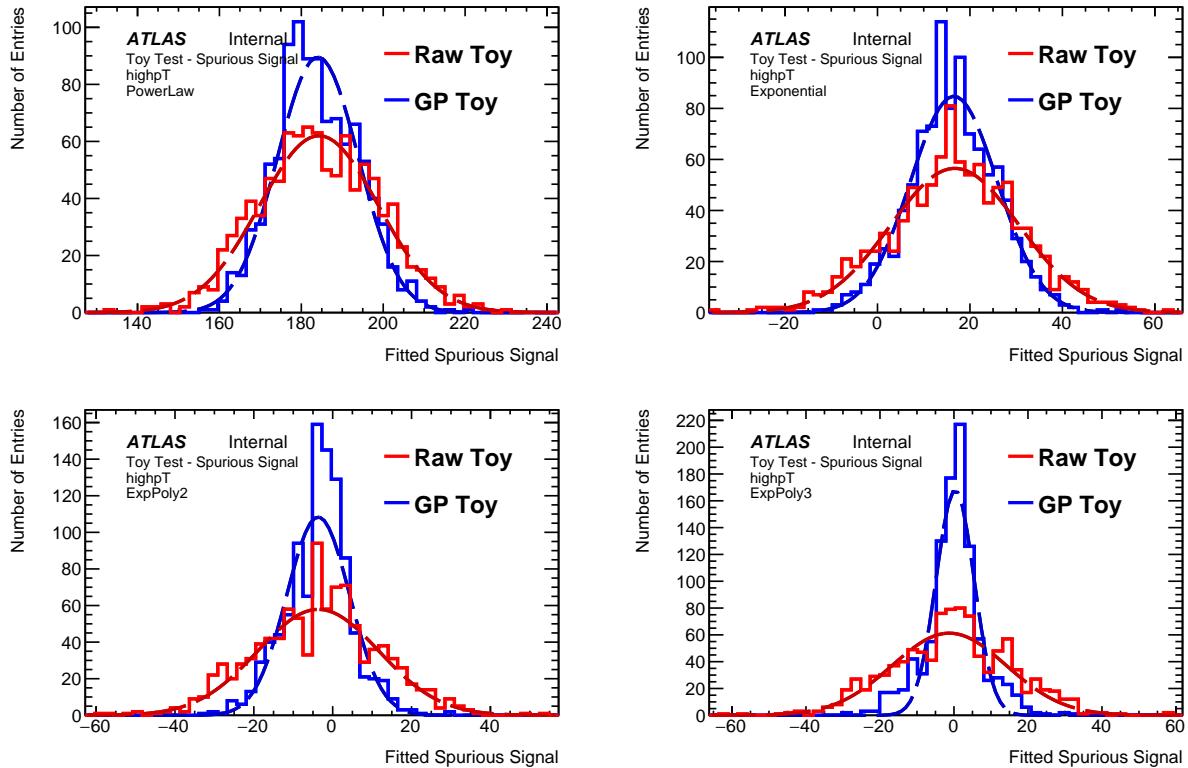


Figure A.231: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 3 GeV wide feature that is 0.01% of the template integral injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

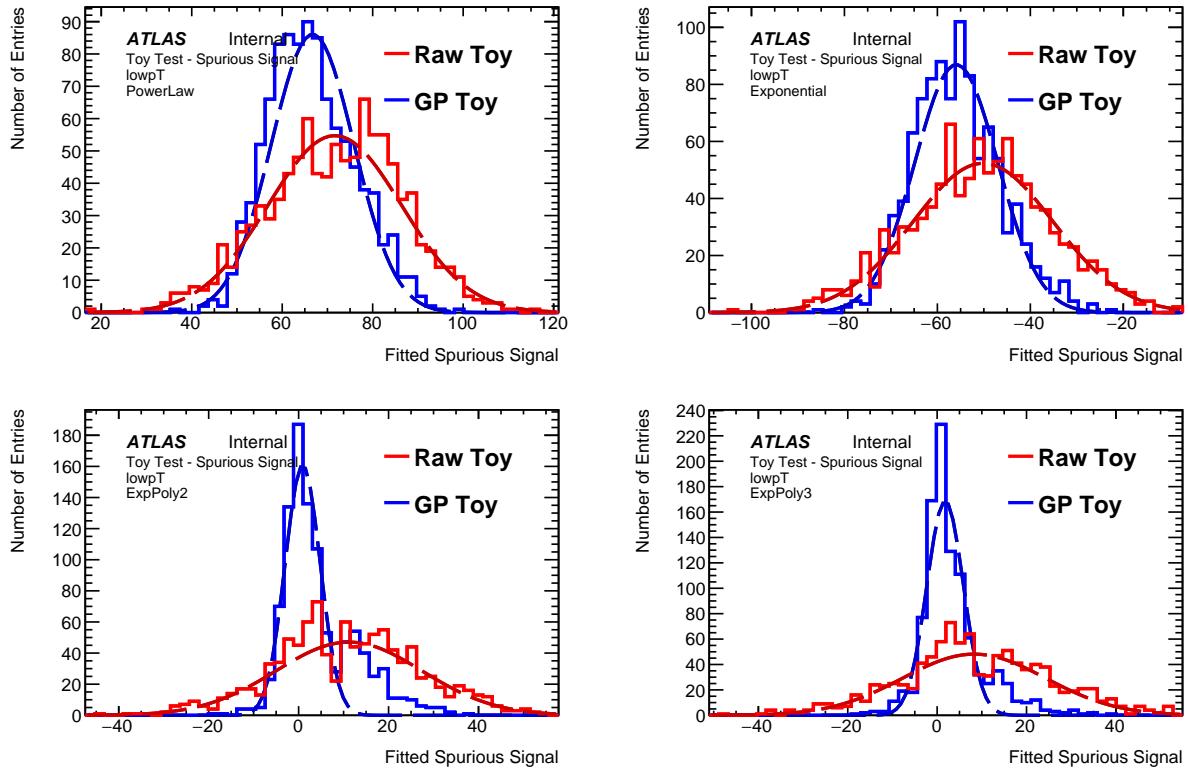


Figure A.232: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 3 GeV wide feature that is 0.1% of the template integral injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

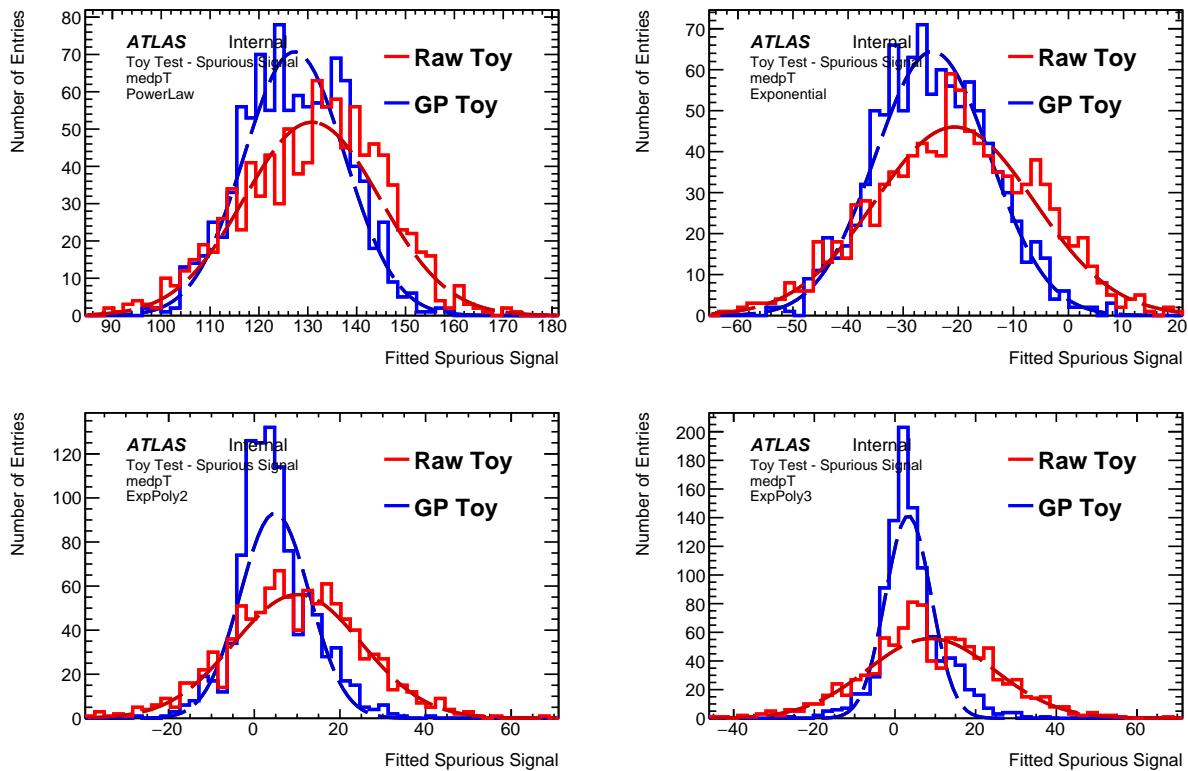


Figure A.233: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 3 GeV wide feature that is 0.1% of the template integral injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

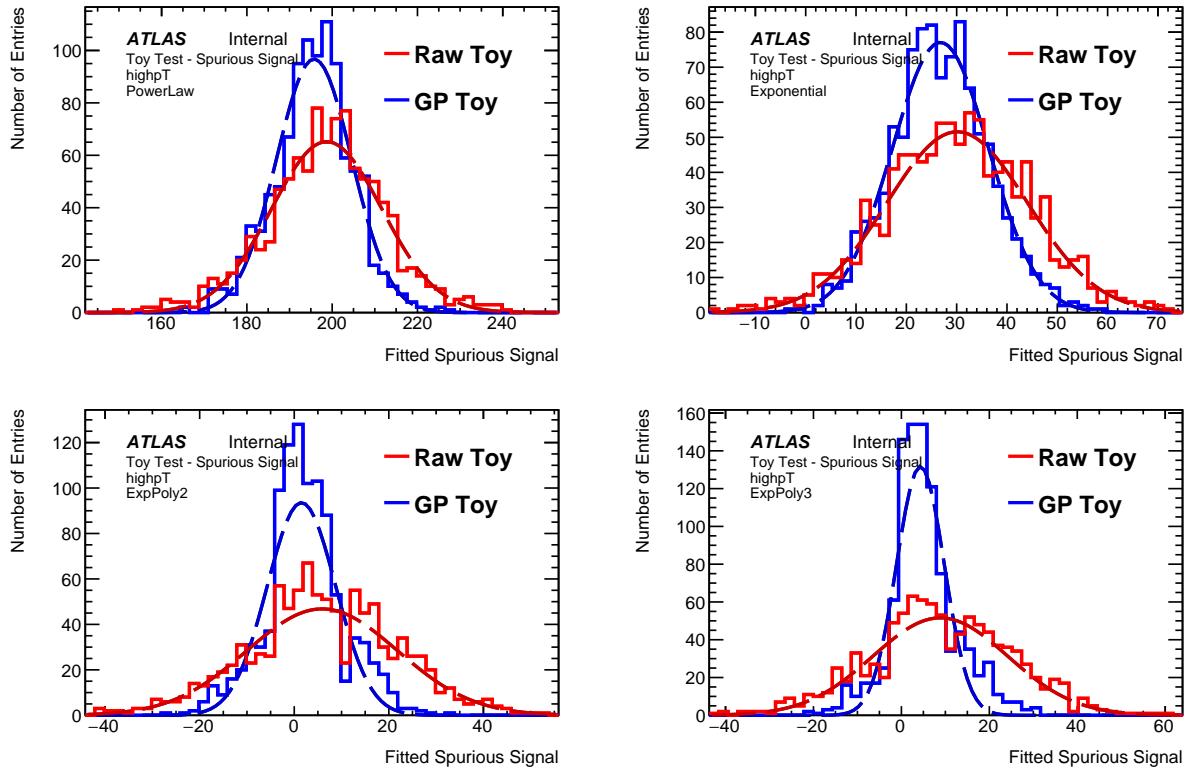


Figure A.234: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 3 GeV wide feature that is 0.1% of the template integral injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

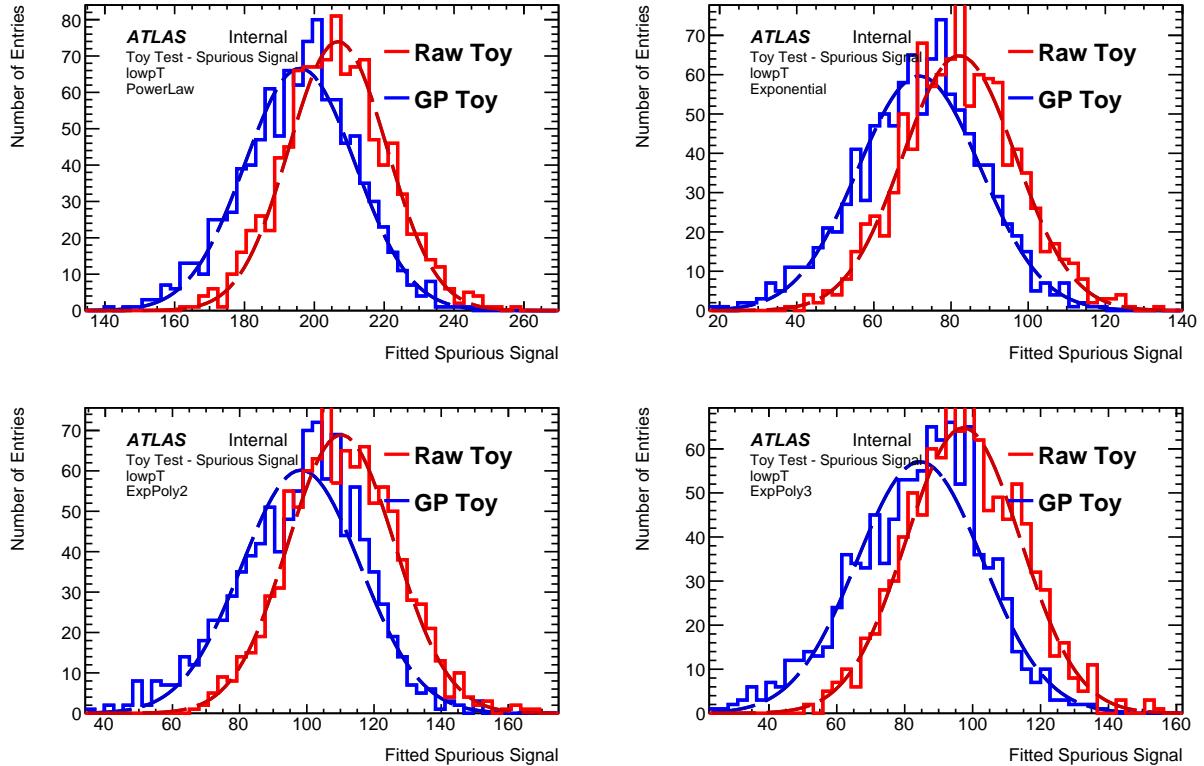


Figure A.235: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 3 GeV wide feature that is 1% of the template integral injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

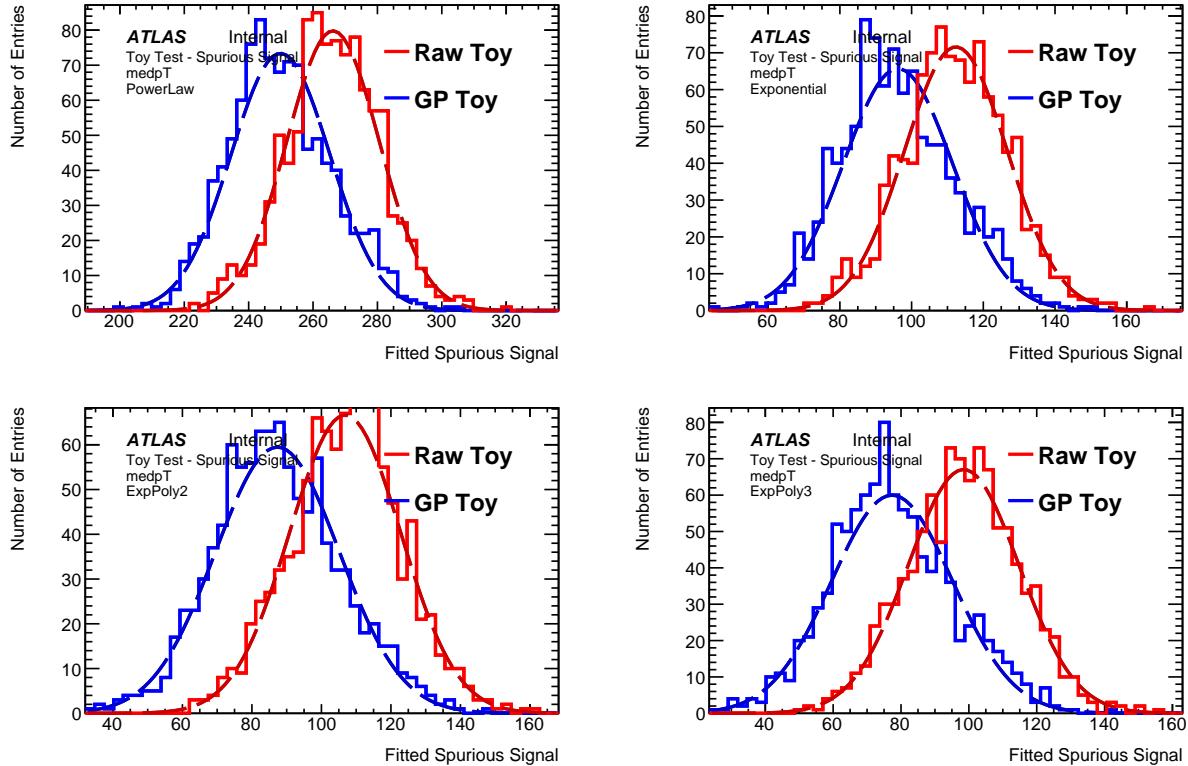


Figure A.236: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 3 GeV wide feature that is 1% of the template integral injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

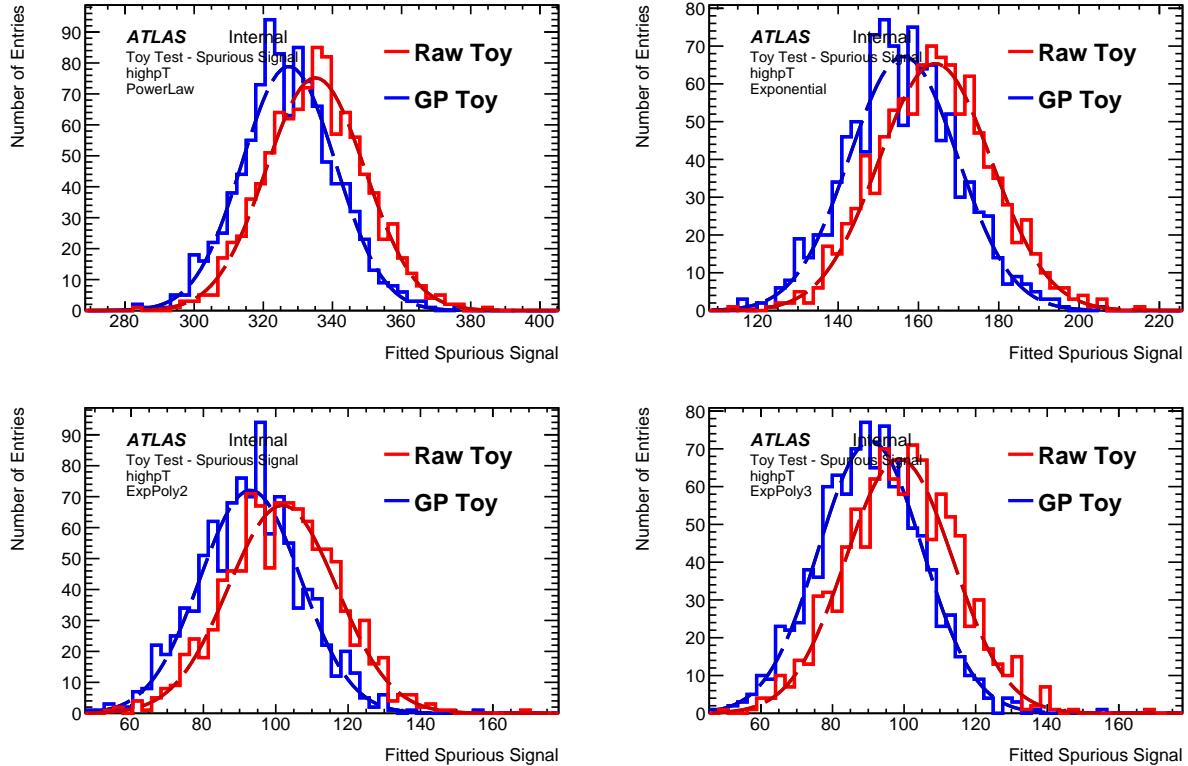


Figure A.237: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 3 GeV wide feature that is 1% of the template integral injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

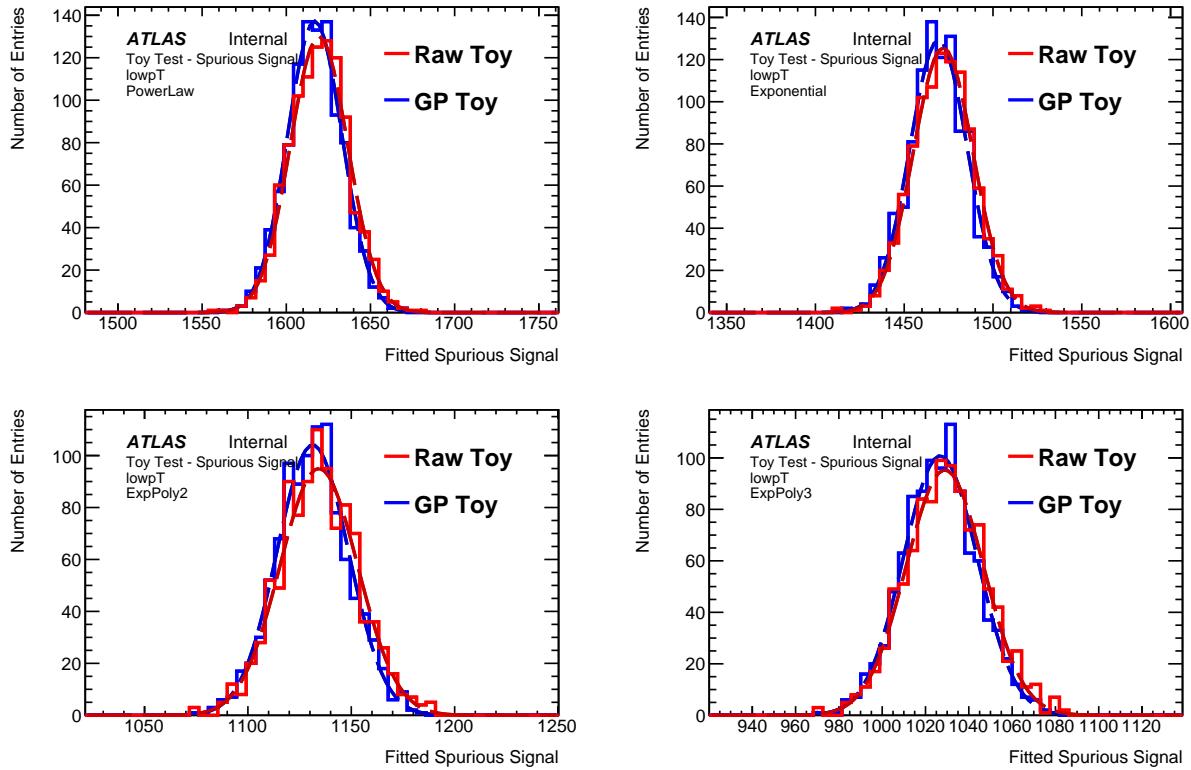


Figure A.238: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 3 GeV wide feature that is 10% of the template integral injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

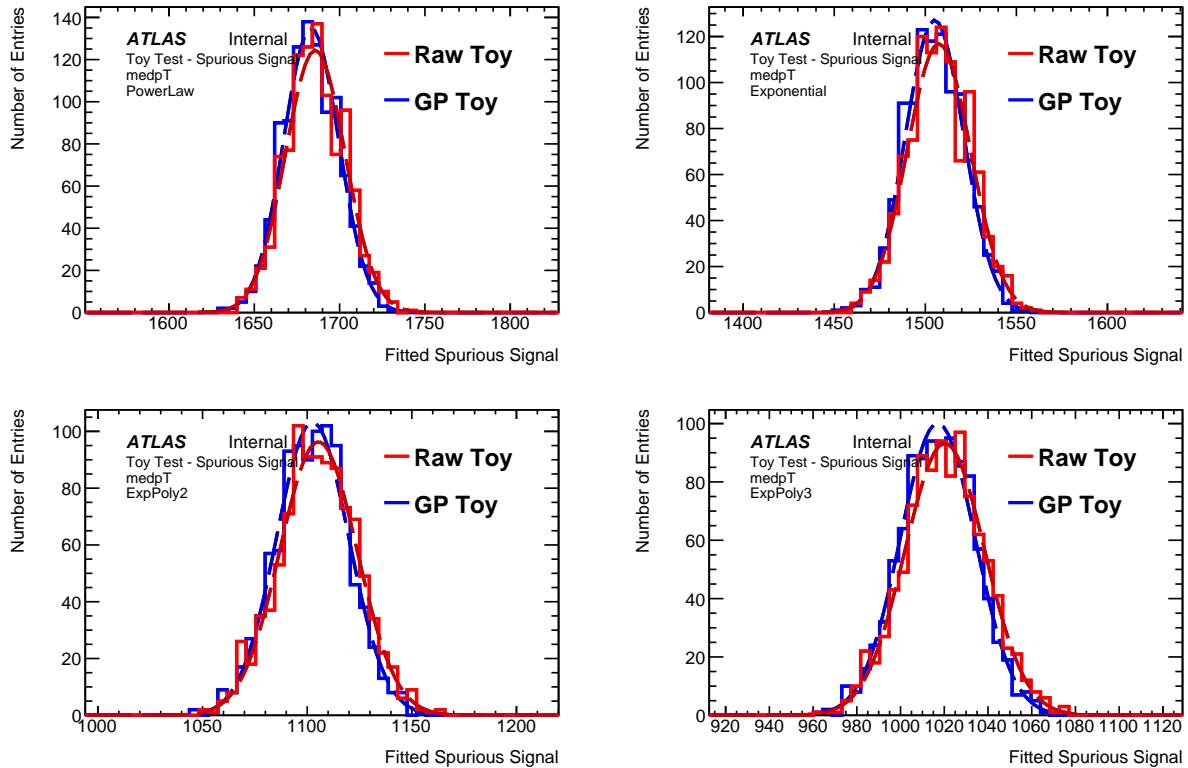


Figure A.239: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 3 GeV wide feature that is 10% of the template integral injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

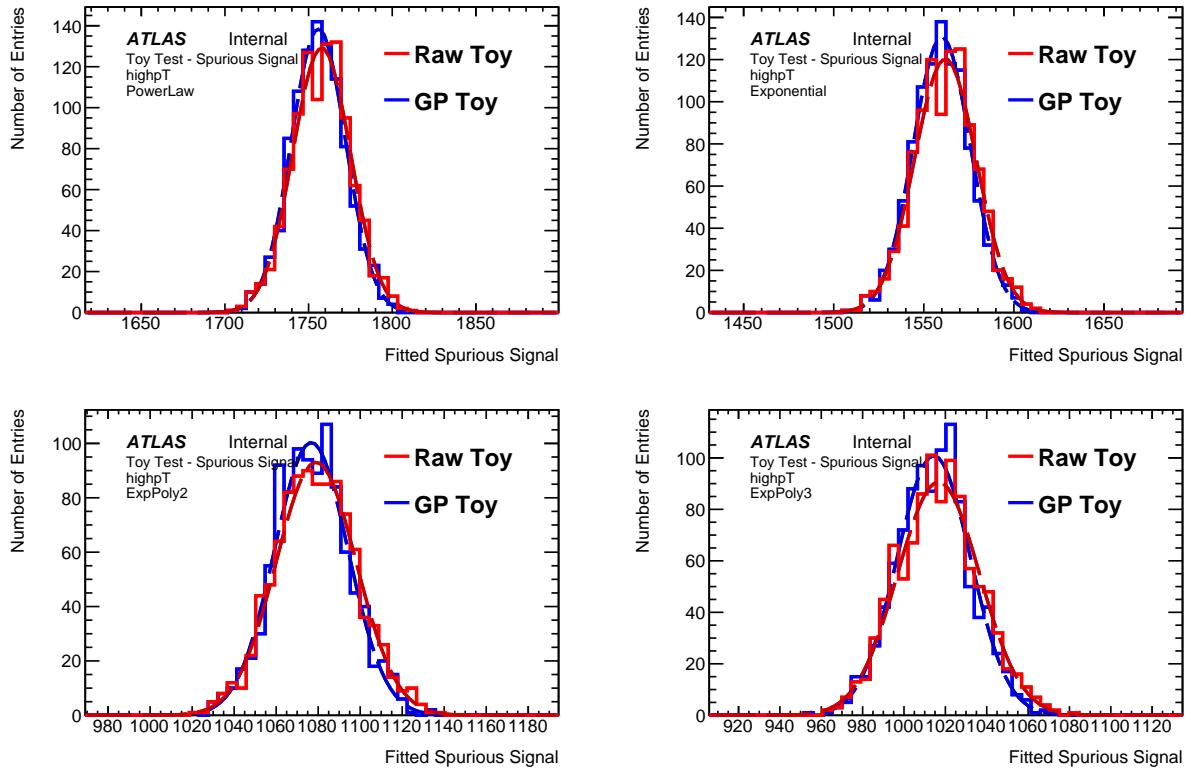


Figure A.240: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 3 GeV wide feature that is 10% of the template integral injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

Nominal	N_sig	Unweighted Events	Bkg events weighted	Mean SS_raw	Mean SS_GPR	GP-Raw: Bias, Mean	Sigma Toy	Sigma GPR	Bias/sigma(SS_raw)	Bias/Feature Size
Fit Function: Exp										
GeneratingFunction: Exp2										
low100	4	1×10^6	4×10^4	-64.14	-64.86	0.72	13.95	7.83	0.05	0.18
low1k	40	1×10^6	4×10^4	-50.71	-55.73	5.02	14.97	9.29	0.33	0.13
low10k	400	1×10^6	4×10^4	110.66	98.26	12.39	16.24	18.61	0.76	0.03
low100k	4000	1×10^6	4×10^4	1471.88	1469.30	2.58	17.08	16.18	0.15	0.00
Fit Function: Exp2										
Generating Function: Exp3										
med100	4	1×10^6	4×10^4	0.67	-0.52	1.20	15.83	7.77	0.08	0.30
med1k	40	1×10^6	4×10^4	9.68	4.76	4.92	15.38	8.66	0.32	0.12
med10k	400	1×10^6	4×10^4	107.02	87.33	19.70	16.01	17.99	1.23	0.05
med100k	4000	1×10^6	4×10^4	1105.21	1102.32	2.89	18.69	17.31	0.15	0.00
Fit Function: Exp2										
Generating Function: Exp3										
high100	4	1×10^6	4×10^4	-3.86	-3.58	-0.28	15.62	8.18	-0.02	-0.07
high1k	40	1×10^6	4×10^4	5.91	2.15	3.76	15.84	8.19	0.24	0.09
high10k	400	1×10^6	4×10^4	101.86	93.13	8.73	15.46	14.34	0.56	0.02
high100k	4000	1×10^6	4×10^4	1078.98	1076.94	2.04	19.15	17.59	0.11	0.00

Table A.18: Spurious signal means and widths for the three test functional-form distributions for a range of different template statistics, with a signal feature injection that is approximately 3 GeV wide. The template statistics are fixed at one million events, and the feature size is varied.

As a final check, we investigate the bias when a standard-model-signal like feature is injected (1 GeV wide). As expected, we see that the bias is larger- narrow features are more smoothed by the Gaussian Process fit, but are still present in the templates.

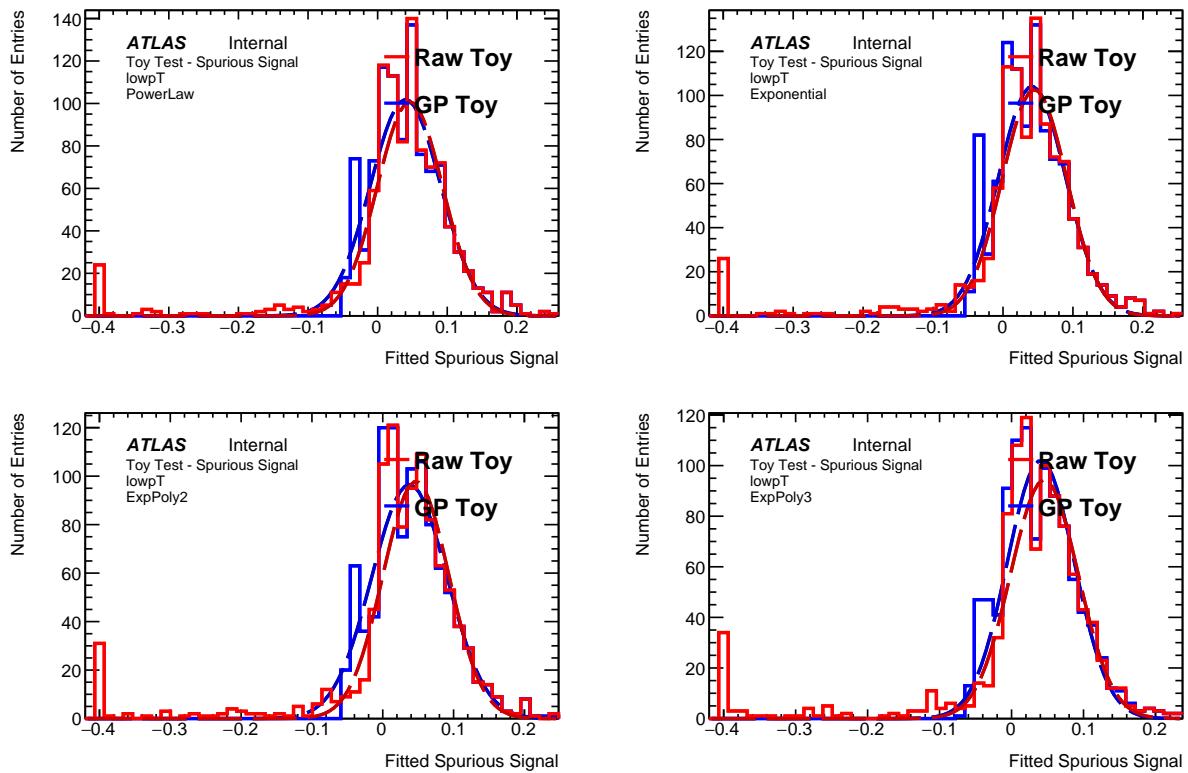


Figure A.241: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10 events.

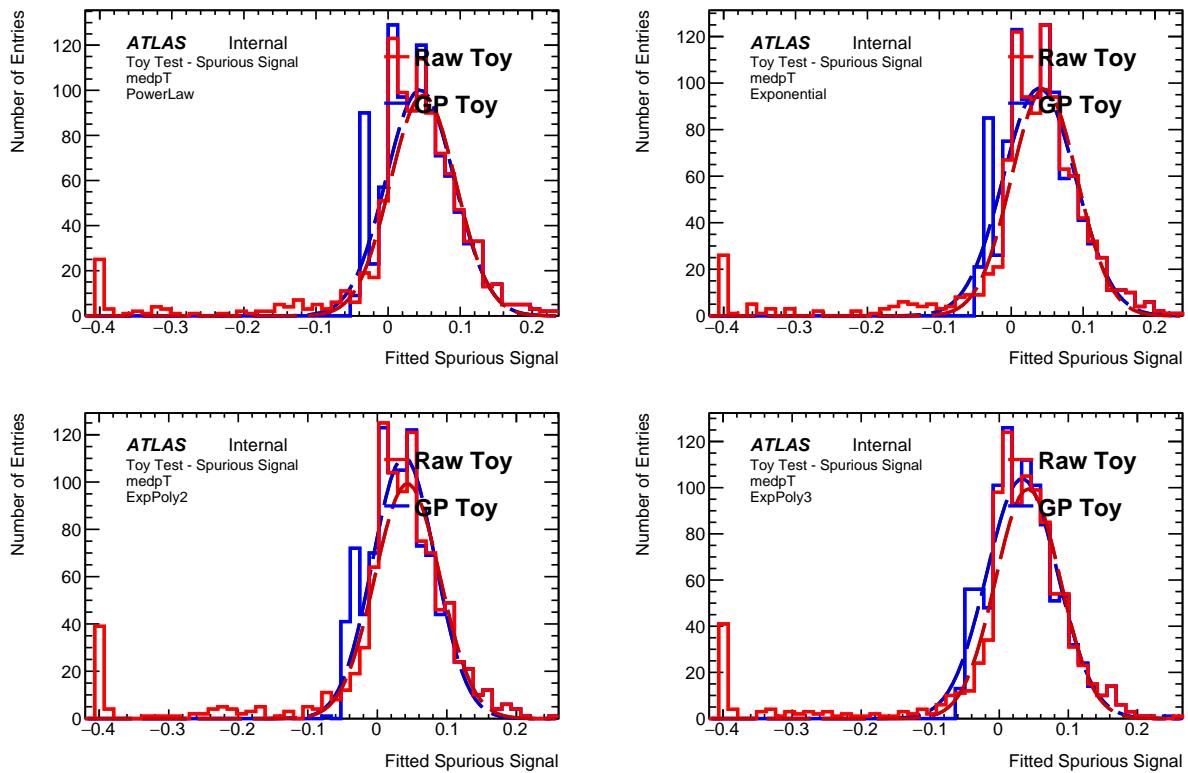


Figure A.242: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10 events.

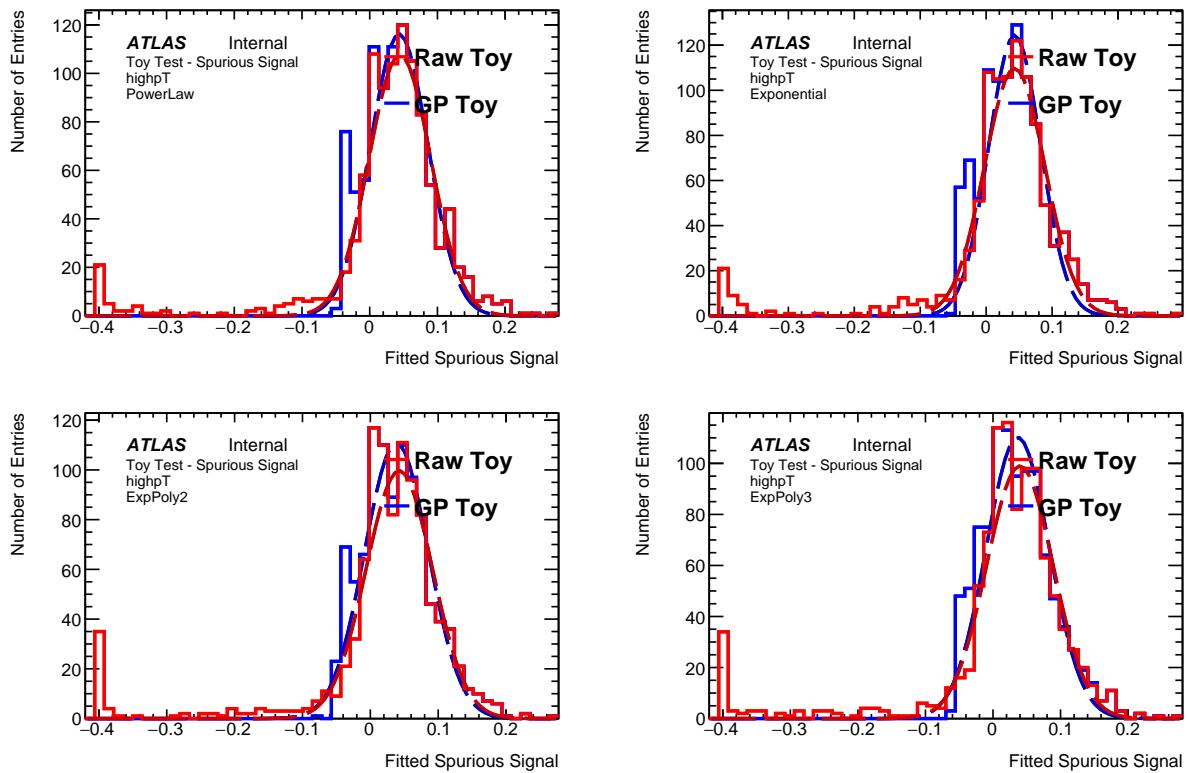


Figure A.243: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10 events.

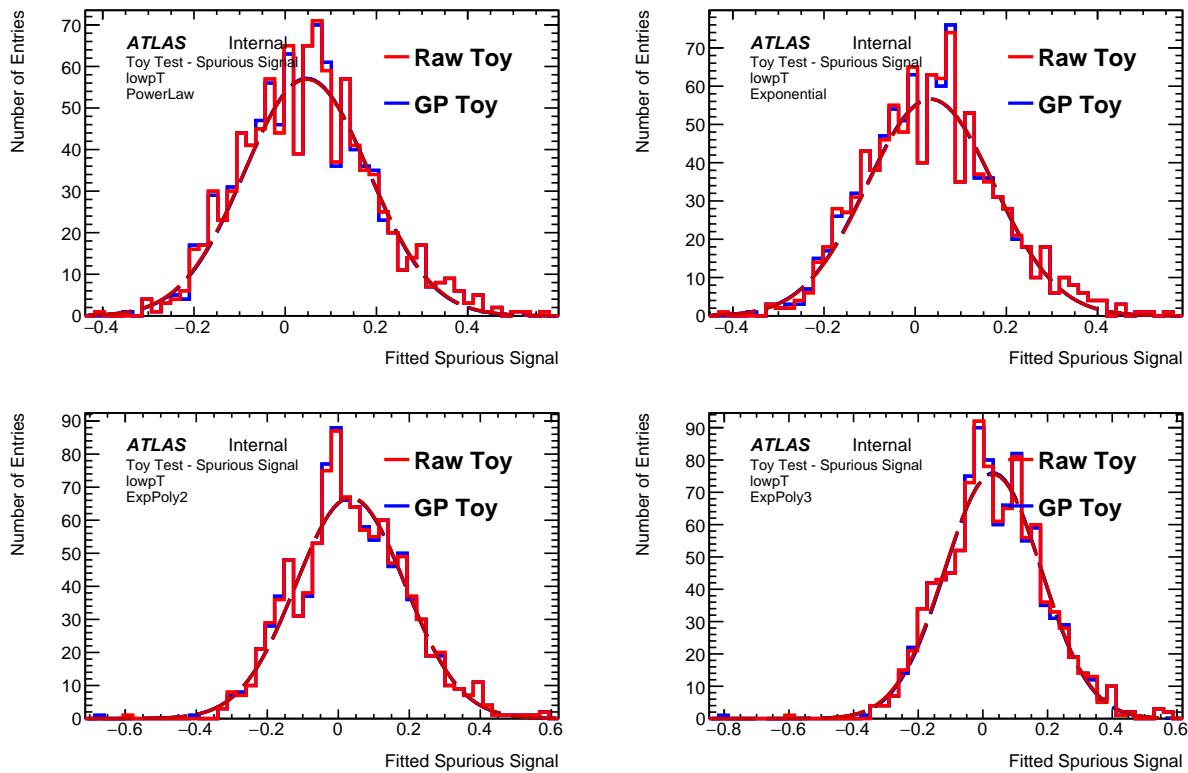


Figure A.244: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100 events.

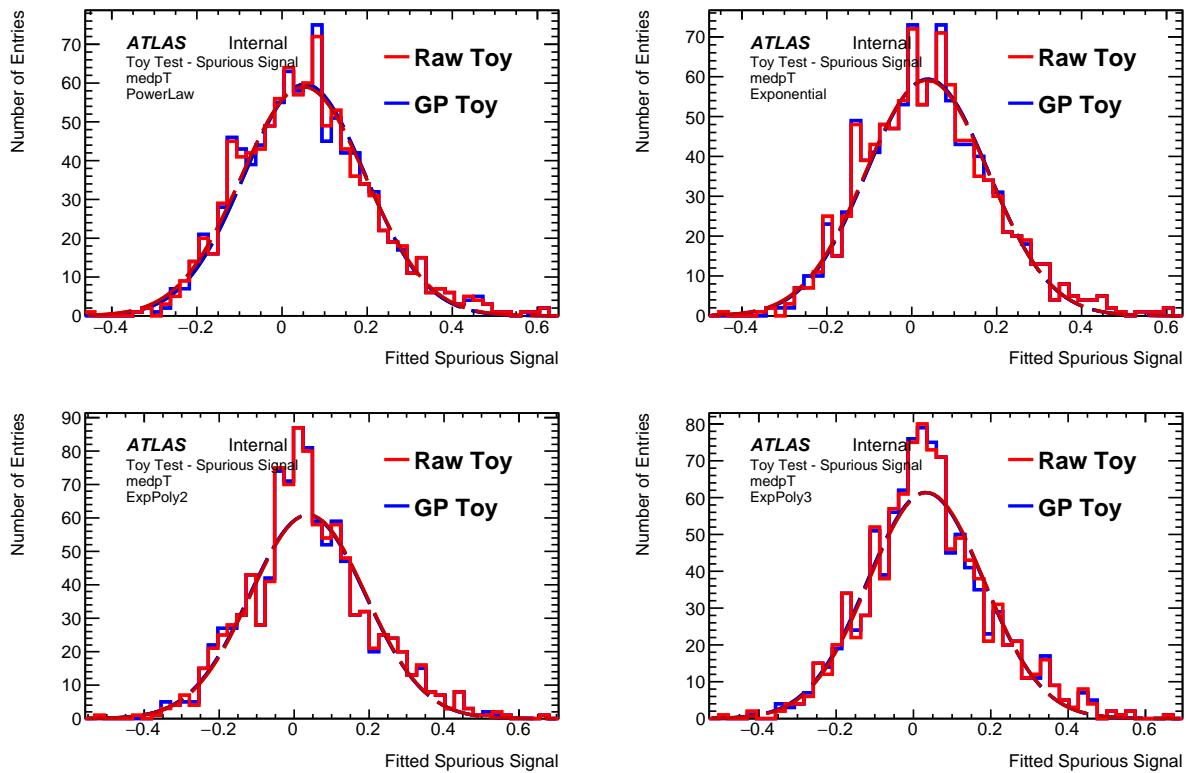


Figure A.245: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100 events.

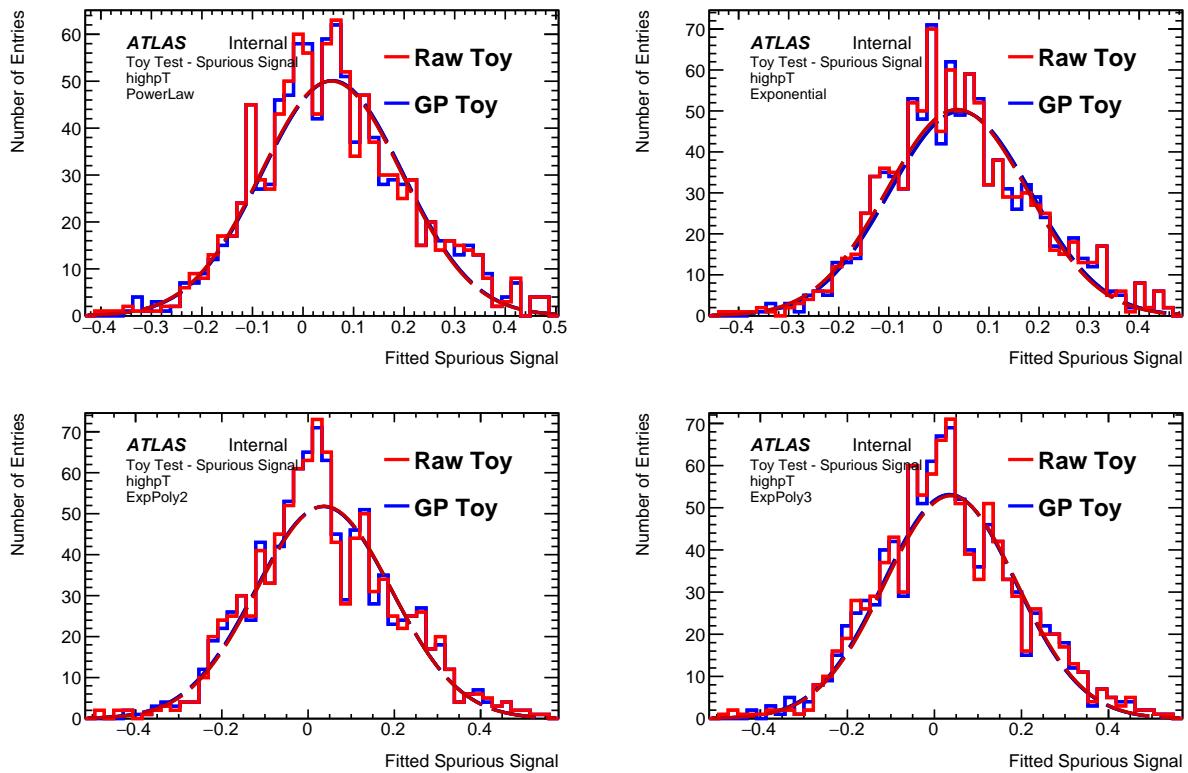


Figure A.246: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100 events.

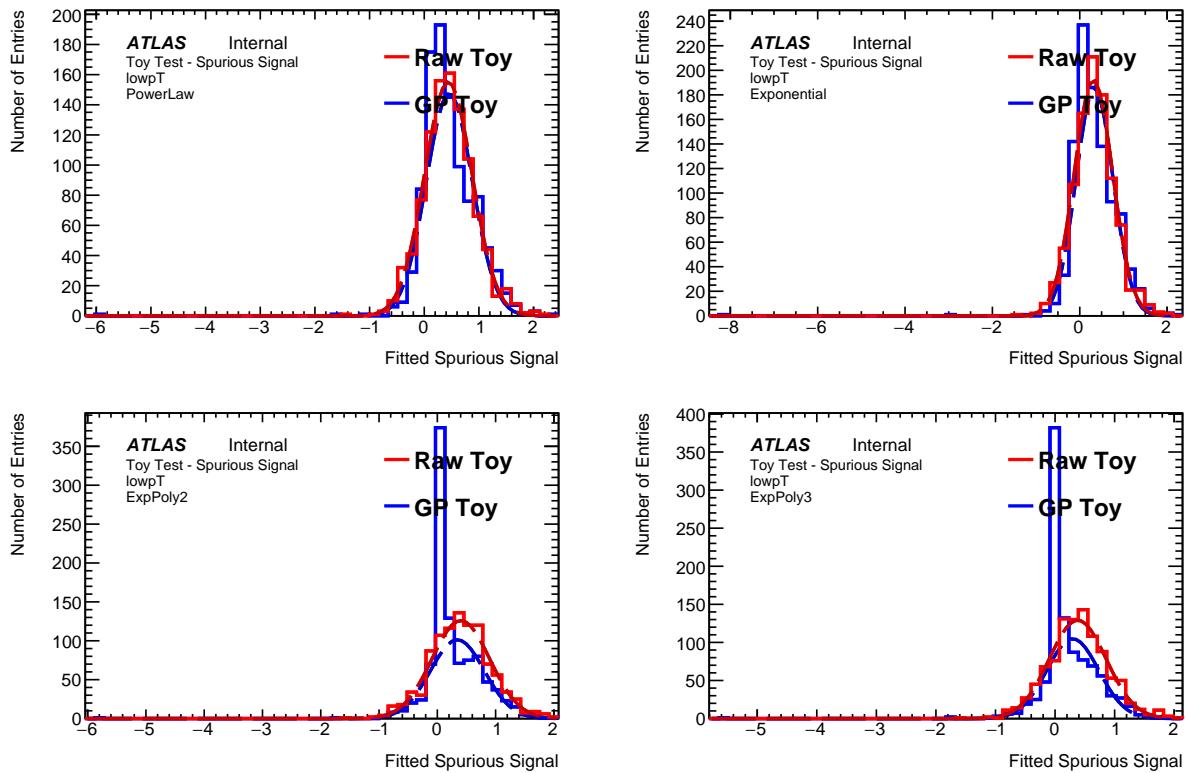


Figure A.247: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1000 events.

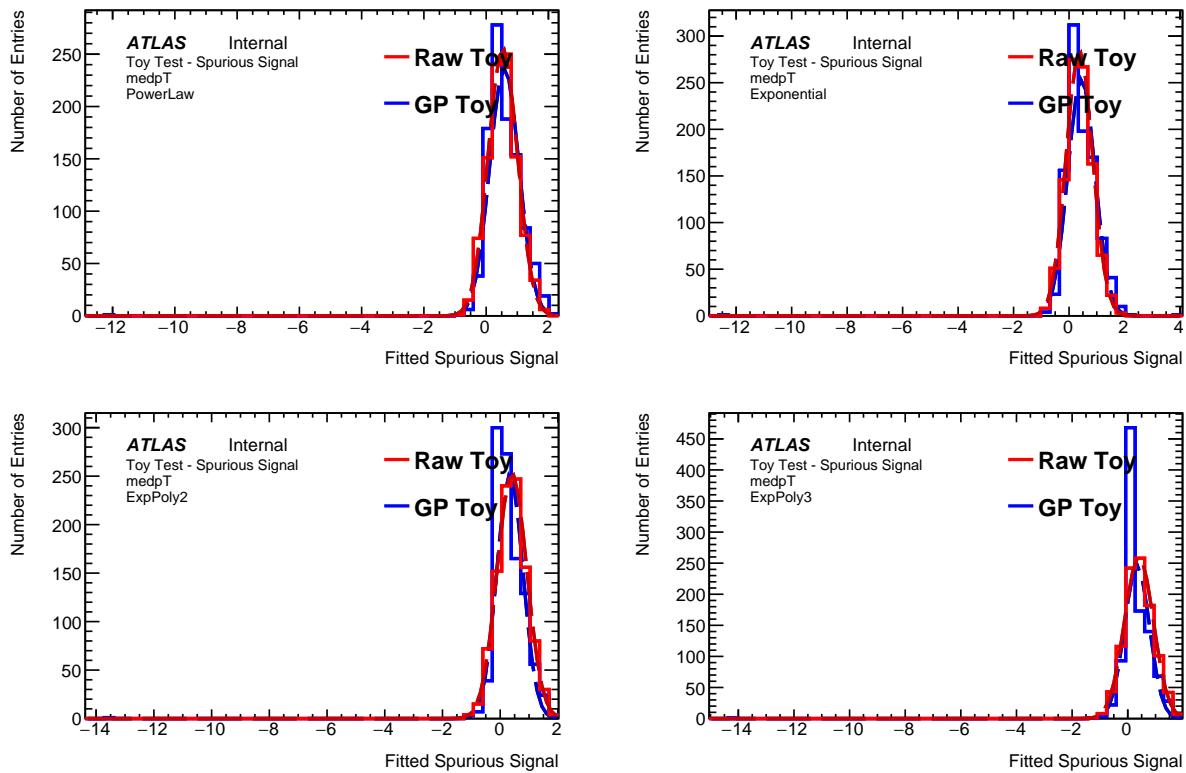


Figure A.248: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1000 events.

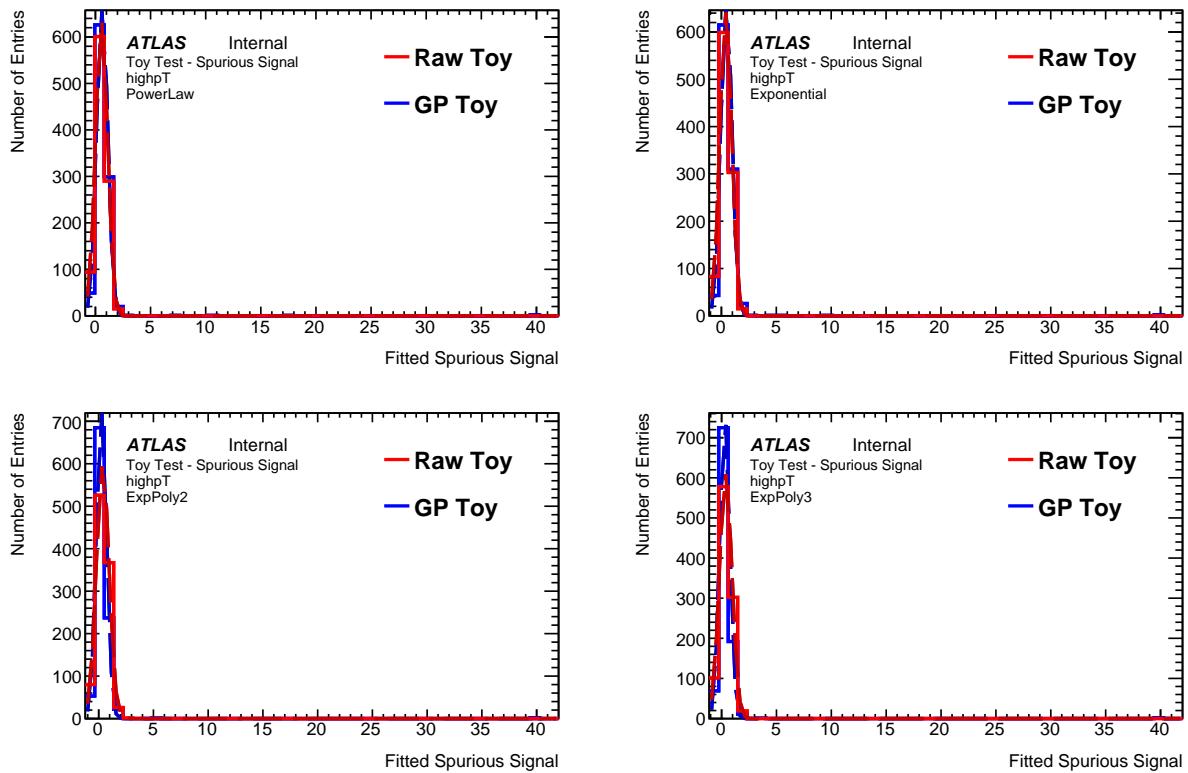


Figure A.249: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1000 events.

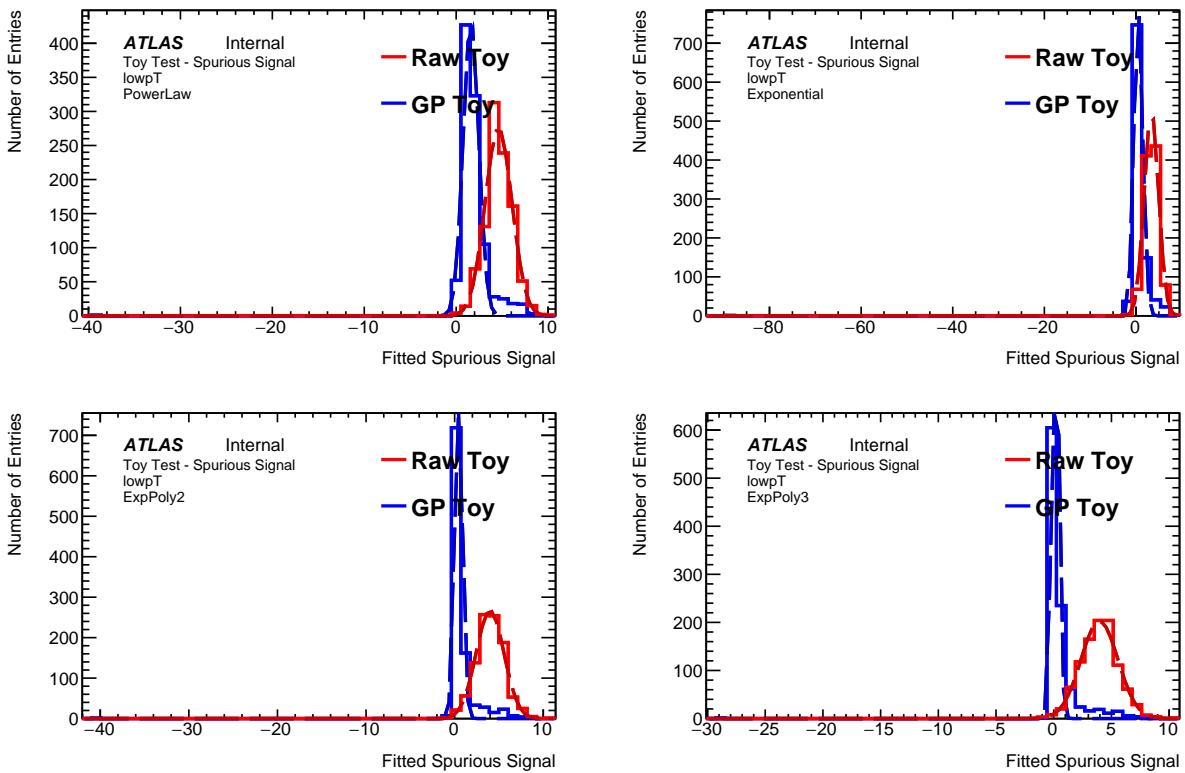


Figure A.250: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10k events.

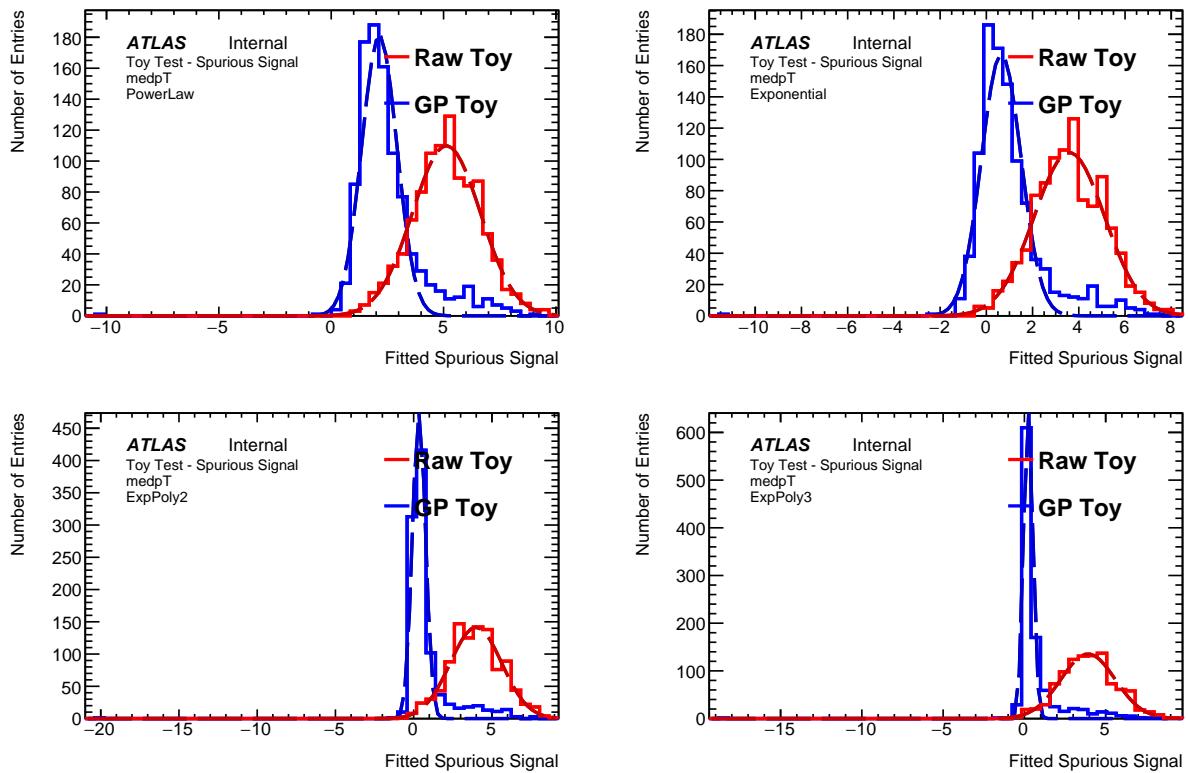


Figure A.251: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10k events.

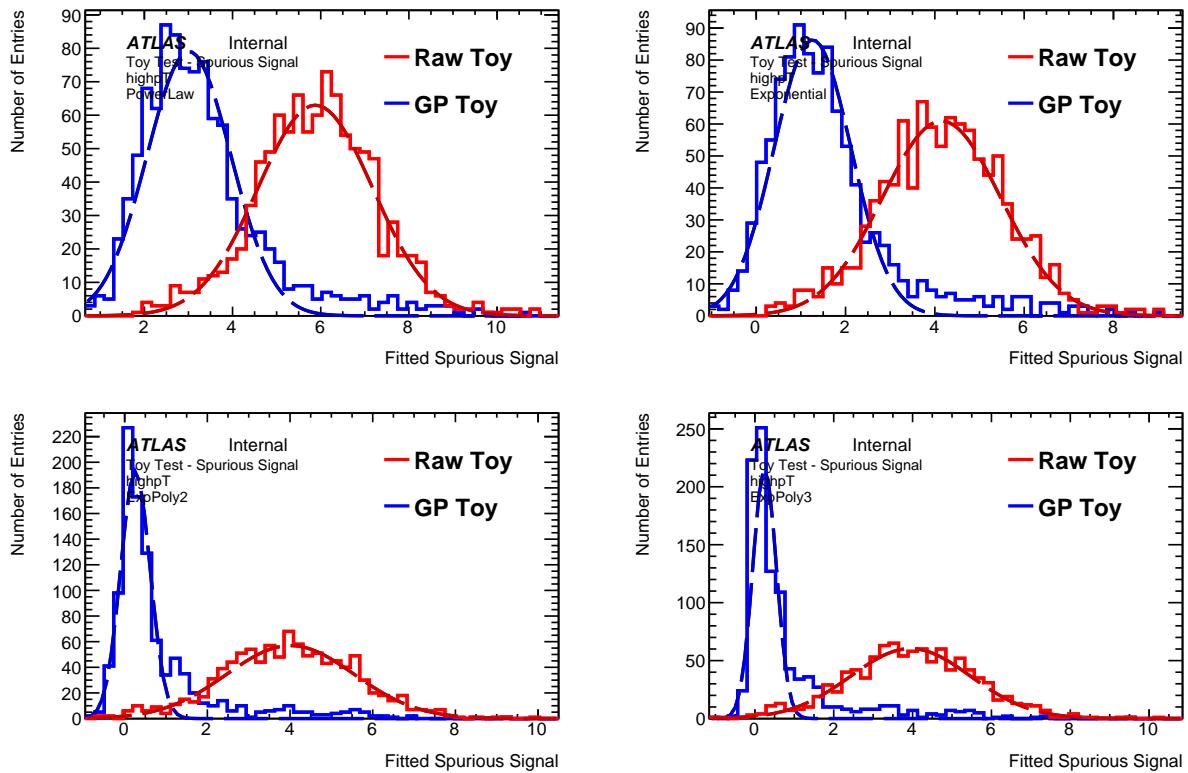


Figure A.252: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10k events.

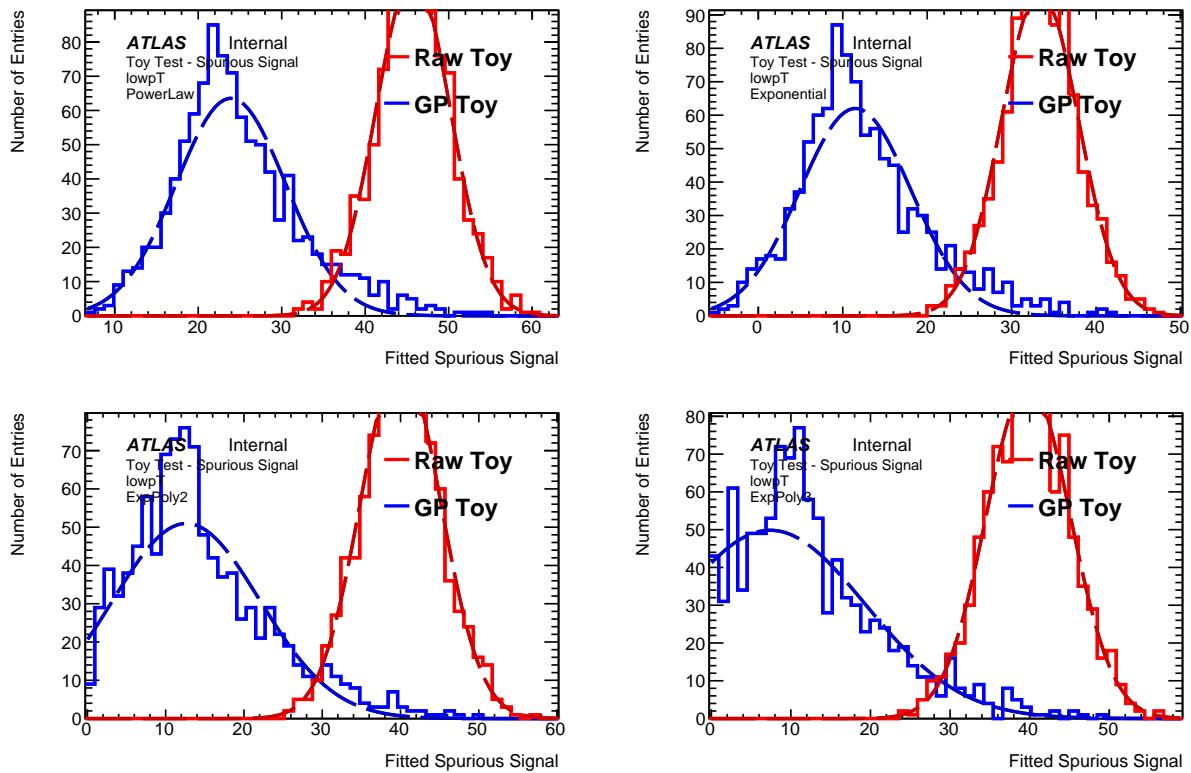


Figure A.253: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100k events.

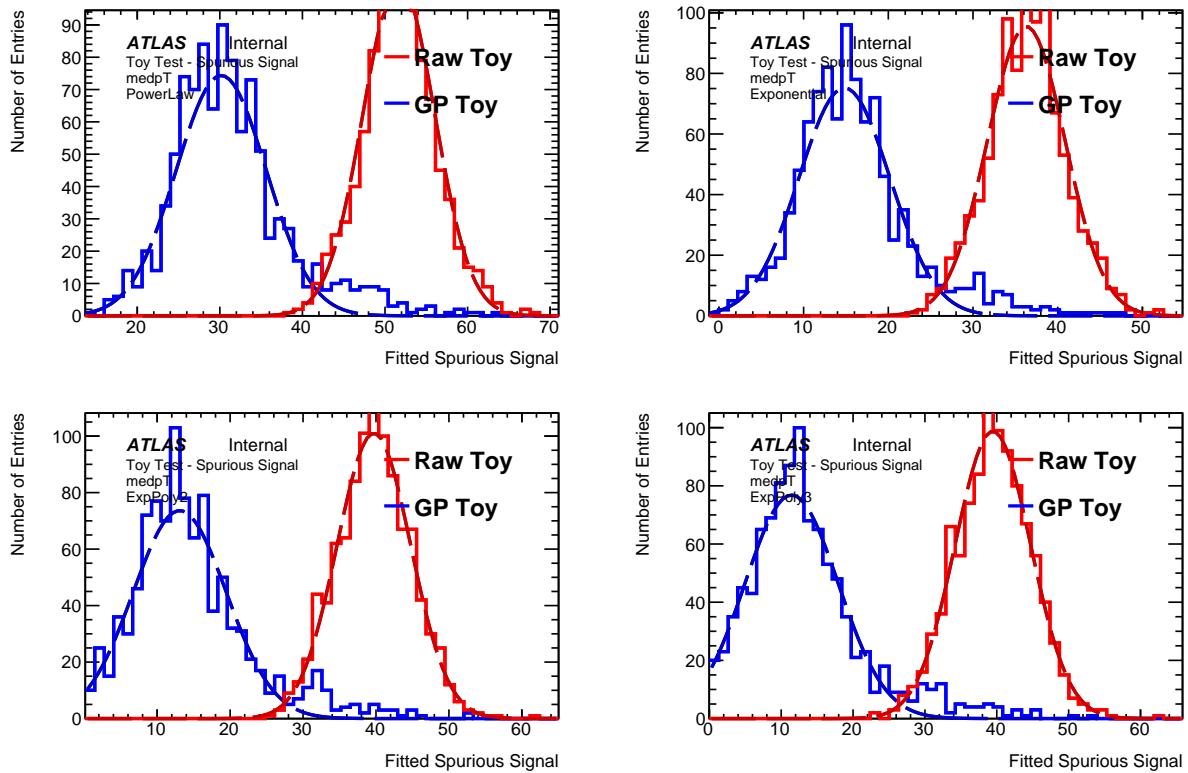


Figure A.254: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100k events.

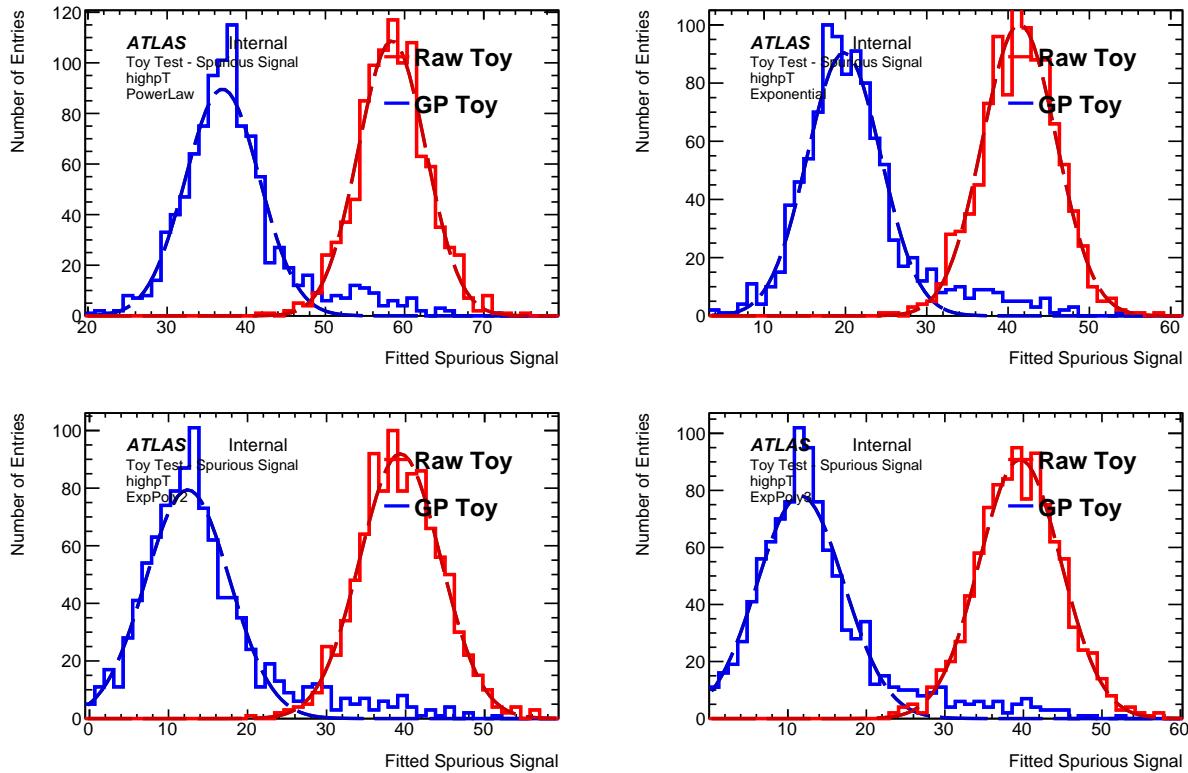


Figure A.255: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 100k events.

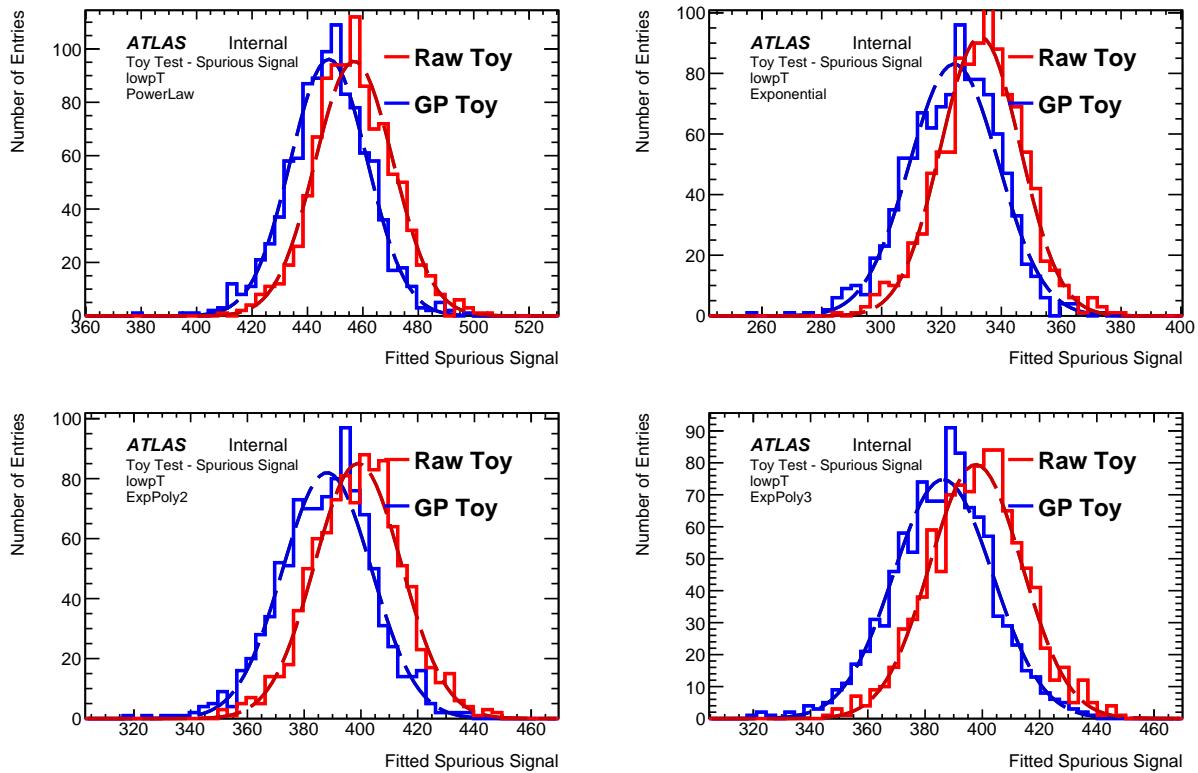


Figure A.256: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

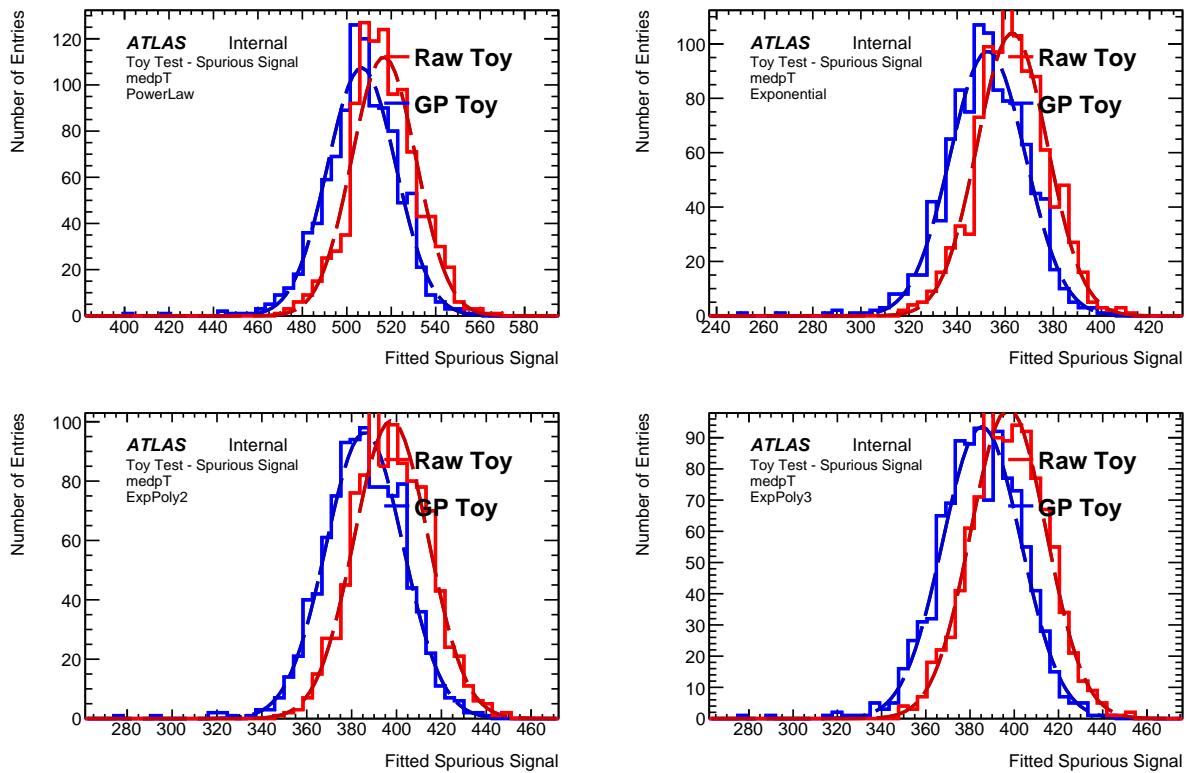


Figure A.257: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

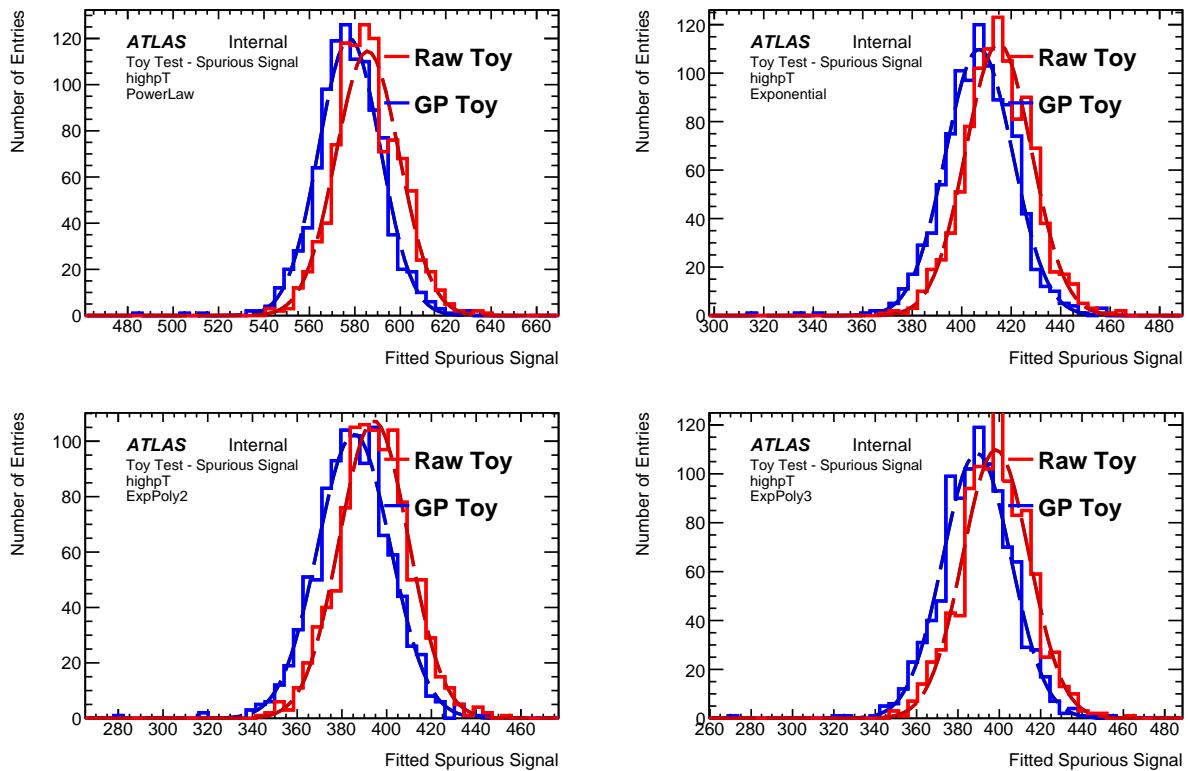


Figure A.258: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 1,000,000 events.

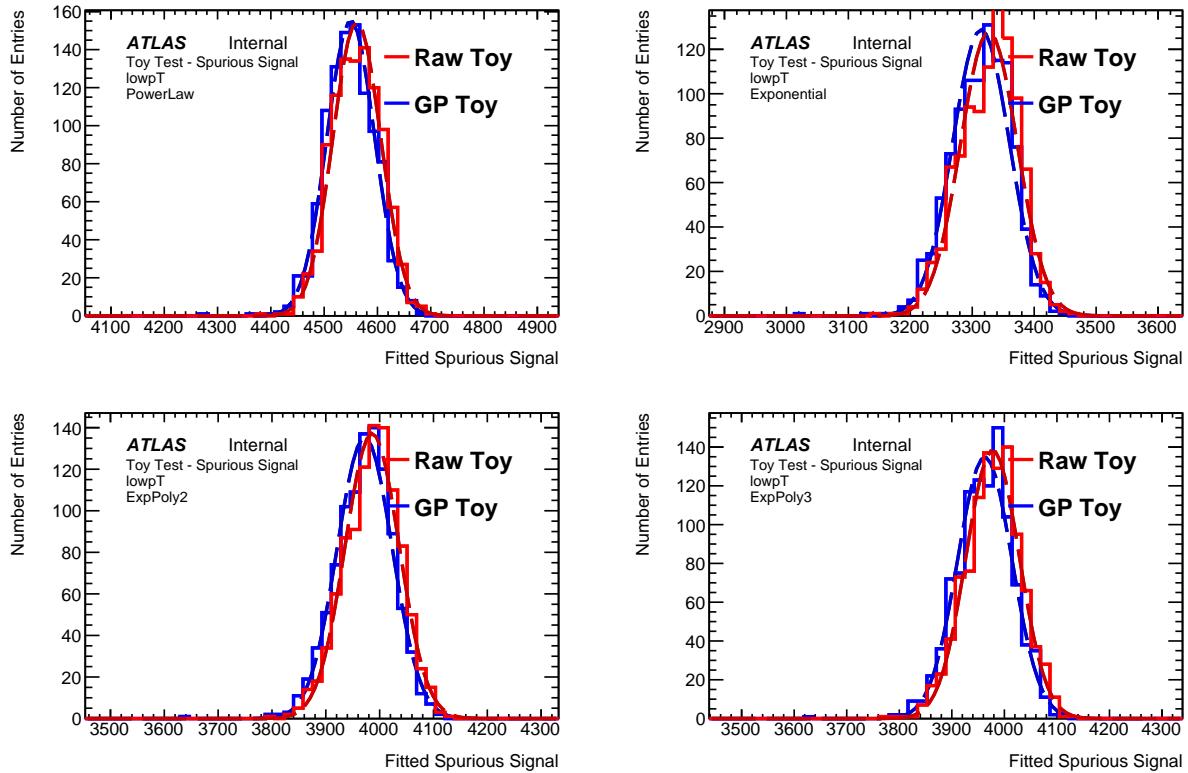


Figure A.259: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly2-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10,000,000 events.

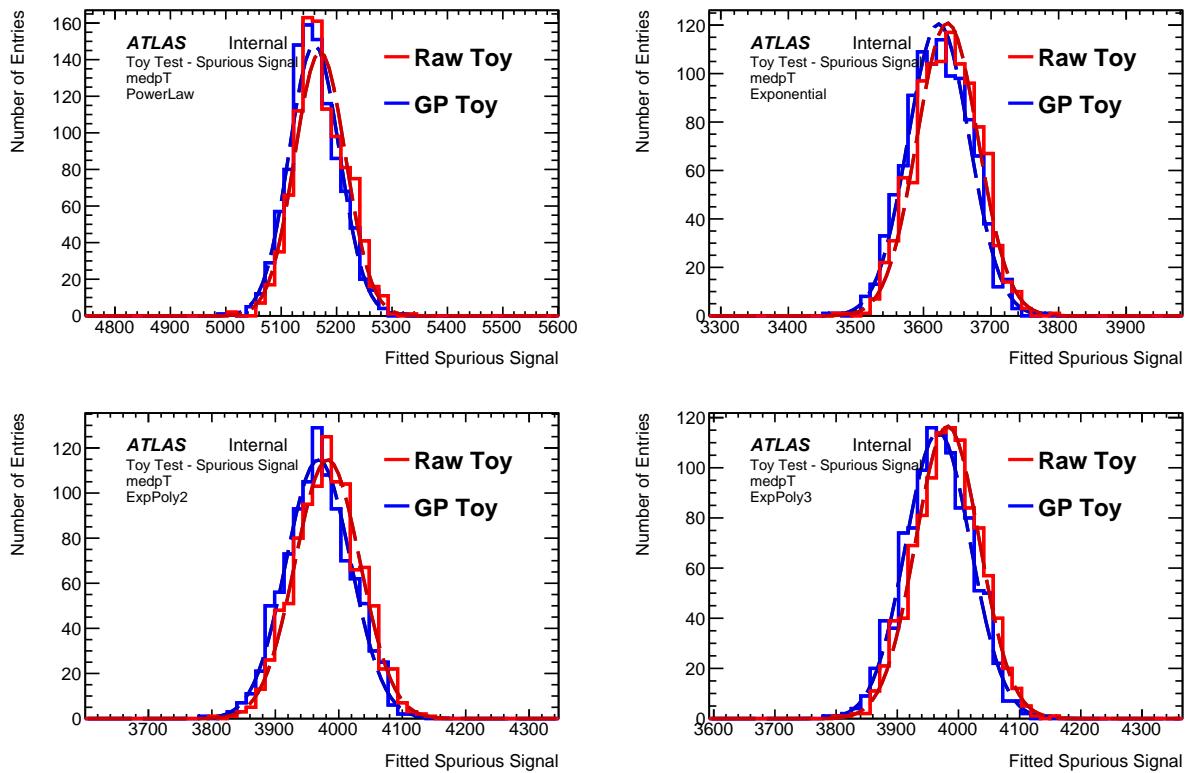


Figure A.260: The distribution of spurious signal for various functions for both the GPR and raw template, using an expPoly3-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10,000,000 events.

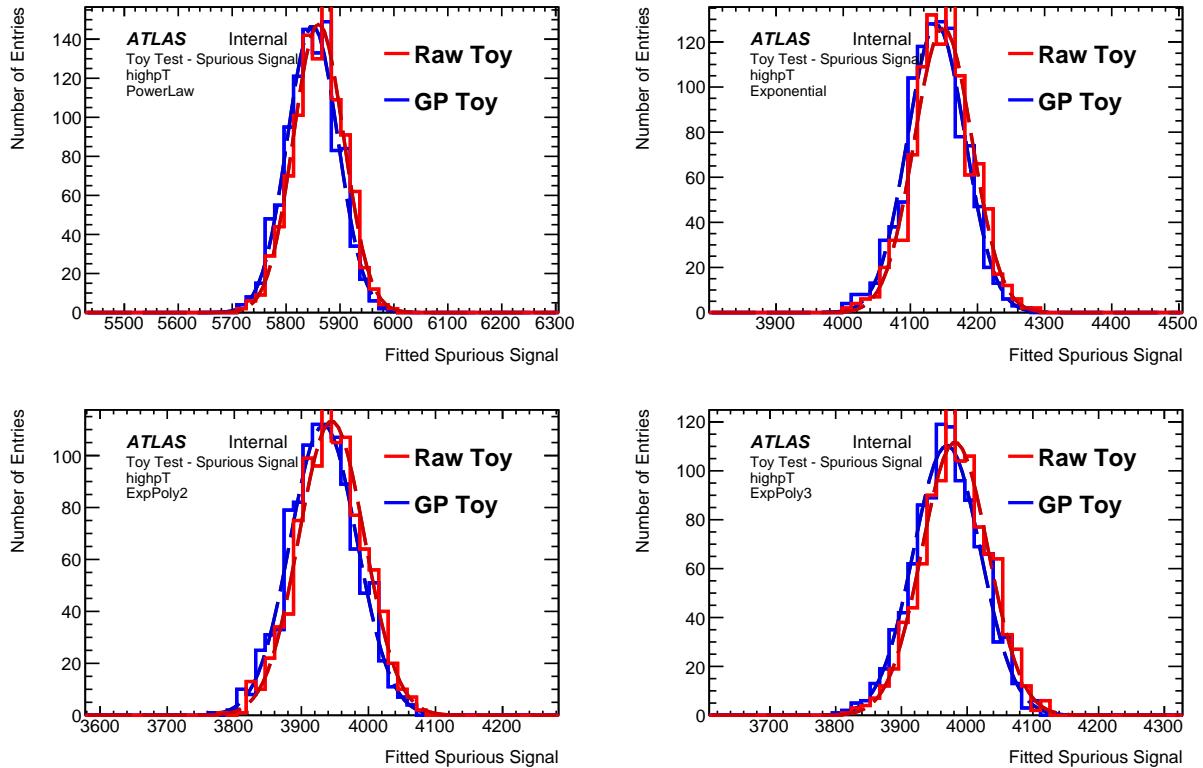


Figure A.261: The distribution of spurious signal for various functions for both the GPR and raw template, using a different expPoly3-derived template with a 1 GeV wide feature injected, extended by 5 GeV on either side, and fit using a linear error kernel. Each toy in this test has 10,000,000 events.

Nominal	N_sig	Unweighted Events	Bkg events weighted	Mean SS_raw	Mean SS_GPR	GP-Raw: Bias, Mean	Sigma Toy	Sigma GPR	Bias/sigma(SS_raw)	Bias/Feature Size
Fit Function: Exp										
GeneratingFunction: Exp2										
low10	4E-2	1	4×10^{-1}	0.03	0.04	-0.01	0.09	0.05	-0.16	-0.37
low100	4E-2	1×10^2	4	0.04	0.04	0.00	0.14	0.14	0.00	0.00
low1k	4E-1	1×10^3	4×10^1	0.33	0.33	0.00	0.46	0.53	0.00	0.00
low10k	4E+0	1×10^4	4×10^2	3.37	0.63	2.74	1.44	3.18	1.91	0.69
low100k	4E+1	1×10^5	4×10^3	33.39	12.36	21.03	4.52	7.60	4.65	0.53
low1M	4E+2	1×10^6	4×10^4	333.18	323.80	9.38	14.11	15.11	0.66	0.02
low10M	4E+3	1×10^7	4×10^5	3326.74	3314.39	12.35	46.12	47.61	0.27	0.00
Fit Function: Exp2										
Generating Function: Exp3										
med10	4E-2	1×10^1	4×10^{-1}	0.02	0.04	-0.02	0.11	0.05	-0.21	-0.59
med100	4E-2	1×10^2	4	0.04	0.04	0.00	0.16	0.16	0.00	0.01
med1k	4E-1	1×10^3	4×10^1	0.41	0.31	0.11	0.51	0.64	0.21	0.26
med10k	4E+0	1×10^4	4×10^2	3.98	0.80	3.18	1.66	1.50	1.92	0.80
med100k	4E+1	1×10^5	4×10^3	39.72	14.86	24.86	5.09	8.20	4.89	0.62
med1M	4E+2	1×10^6	4×10^4	397.57	385.41	12.16	16.56	18.26	0.73	0.03
med10M	4E+3	1×10^7	4×10^5	3982.59	3968.35	14.24	51.18	51.29	0.28	0.00
Fit Function: Exp2										
Generating Function: Exp3										
high10	4E-2	1×10^1	4×10^{-1}	0.02	0.04	-0.02	0.11	0.05	-0.19	-0.52
high100	4E-2	1×10^2	4	0.04	0.04	-0.00	0.16	0.16	-0.00	-0.01
high1k	4E-1	1×10^3	4×10^1	0.36	0.36	0.01	0.52	1.84	0.02	0.02
high10k	4E+0	1×10^4	4×10^2	3.98	0.81	3.17	1.60	1.35	1.99	0.79
high100k	4E+1	1×10^5	4×10^3	39.33	14.61	24.72	5.19	8.32	4.76	0.62
high1M	4E+2	1×10^6	4×10^4	394.50	385.36	9.13	15.97	16.78	0.57	0.02
high10M	4E+3	1×10^7	4×10^5	3944.36	3932.44	11.91	49.17	49.29	0.24	0.00

Table A.19: Spurious signal means and widths for the three test functional-form distributions for a range of different template statistics, with a signal feature injection that is approximately 3 GeV wide and 1% of the template integral.

From these studies, we conclude that, in the presence of underlying features that we wish to preserve, the bias is dependent on both the size and shape of the expected feature- features are blunted somewhat by the GP, but are still present in the smoothed template; how much they are blunted depends on their shape and size, both absolute and relative to the template as a whole.

A.5.5 Results

From these studies, we conclude that, for the purposes of the Couplings analysis, the bias depends only on statistics, and that the bias is, for templates containing greater than 20 effective Monte Carlo events per bin, less than 20% of the statistical uncertainty of the nominal spurious signal (i.e., the MC stat uncertainty on the template), with a bias of near 20% in the very low-stat regime. As this rises to a threshold of However, in this very low-stat regime, statistical uncertainty is expected to dominate over the spurious signal systematic, so the effects of GPR bias are trivial. Thus, we can safely conclude that GPR is effectively unbiased in this stat regime, so long as no features are expected in the true underlying distribution.

If we did expect true, significantly-sized underlying features in the templates (i.e., our templates were not assumed to be smoothly falling), we can also conclude that we could conservatively model the bias due to GPR as approximately 18% of the expected integral of the feature, so long as the feature was larger than .01% of the expected template integral. However, the unbinned fitting method used in the analyses discussed in this dissertation assumes the templates can be modelled by smoothly falling functions, so as a result, we can safely neglect the feature-injection bias on the spurious signal numbers.

These validation tests prove the robustness of the Gaussian Process method. In addition to the Coupling analysis, GPR has shown promising results in XXX, XXX, and XXX. By reducing problematic statistical fluctuations, the spurious signal systematic is reduced substantially, and may be able to significantly increase the precision of $H \rightarrow \gamma\gamma$ analyses in the future.