simplify https://github.com/eschanet/simplify

ATL-PHYS-PUB-2021-038

In the simplified likelihood [..], the background model is approximated with a single background sample, representing the total SM background rate in the different analysis channels.

ok

The pre-fit sample rate of the total background sample is set to the total post-fit background rate obtained in the background-only fit in the full likelihood.

Why? To our mind creates a bias

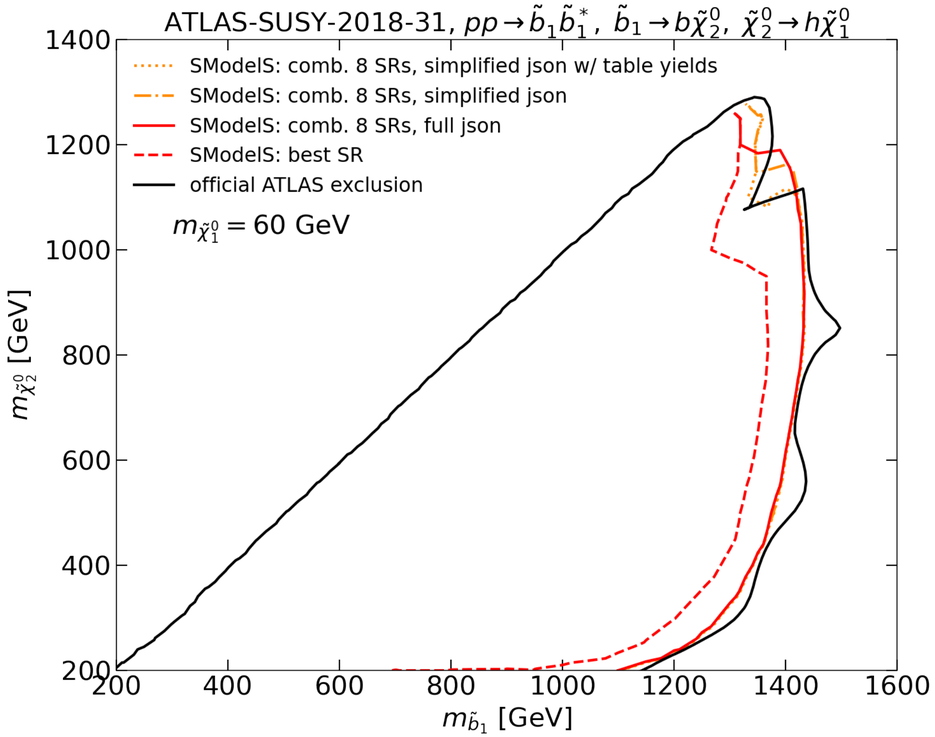
Furthermore, the complete set of nuisance parameters in the original full likelihood is reduced to a single constrained parameter α, representing the post-fit uncertainty on the total SM background estimate in each bin. It is constrained by a Gaussian $G(a = 0|\alpha, \sigma = 1)$ and is correlated over all bins in each channel. Although the final uncertainty is thus constrained by a simple Gaussian, the use of the full likelihood in the background-only fit used to calculate the pre-fit uncertainty ensures that non-Gaussian effects are included.

??

We'd like to understand how this describes correlations between SRs ?

Section 2.2: Building simplified likelihoods

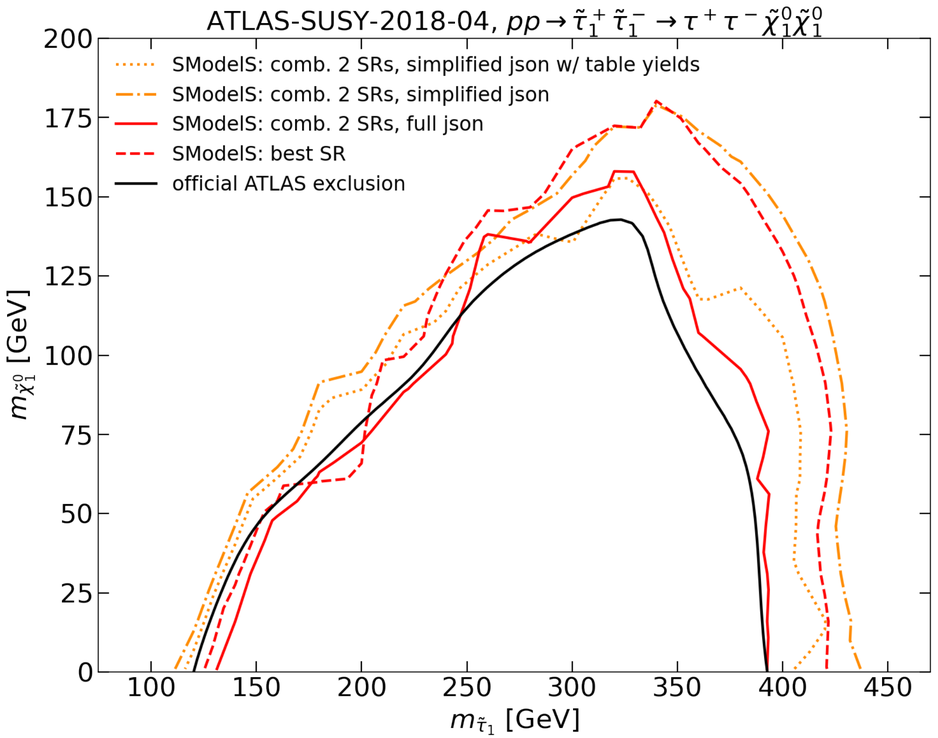
SModelS validation with full and simplified json llhds



plot by Timothée Pascal

Seems to work well

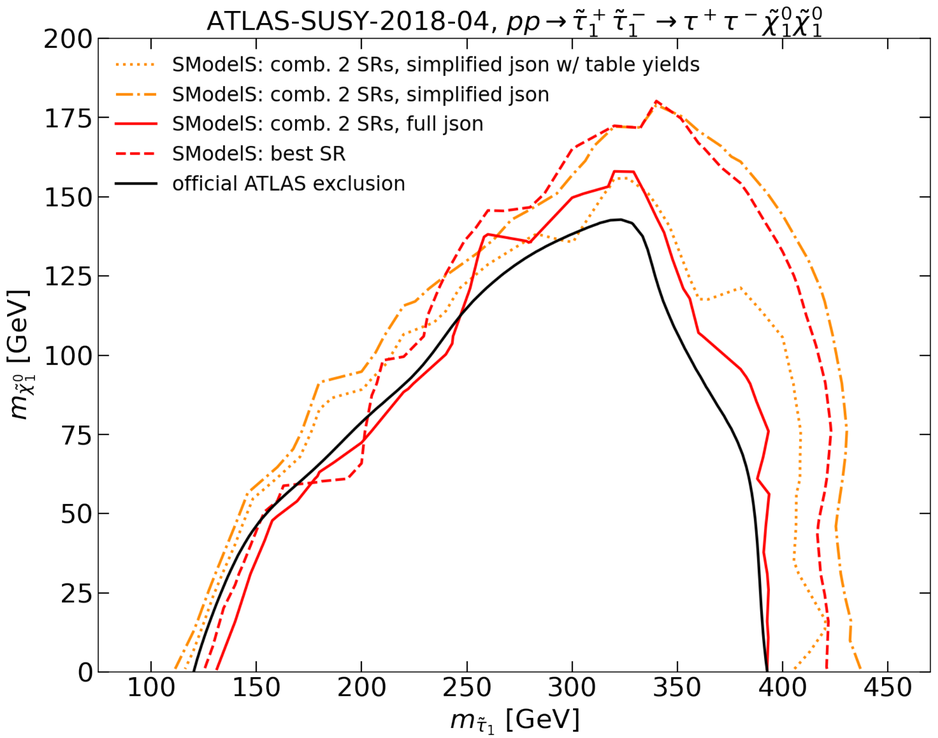
SModelS validation with full and simplified json llhds



plot by Timothée Pascal

SR	full json	table in pub.	simplified json
SRlow	6.06	6.0 ± 1.7	7.3 ± 2.8
SRhig	10.34	10.2 ± 3.3	8.7 ± 2.0

SModelS validation with full and simplified json llhds



plot by Timothée Pascal

SM process	Multi-jet CR-A -lowMass	Multi-jet CR-A -highMass	WCR	SR -lowMass	SR -highMass
Diboson	1.4 ± 0.6	1.9 ± 1.0	63 ± 21	1.4 ± 0.8	2.6 ± 1.4
W+jets	13 ± 4	4^{+7}_{-4}	850 ± 70	1.5 ± 0.7	2.5 ± 1.8
Top quark	2.7 ± 0.9	3.3 ± 1.6	170 ± 40	$0.04^{+0.80}_{-0.04}$	2.0 ± 0.6
Z+jets	$0.25^{+1.43}_{-0.25}$	1.5 ± 0.8	13 ± 7	$0.4^{+0.5}_{-0.4}$	$0.05^{+0.13}_{-0.05}$
Multi-jet	55 ± 10	16 ± 6	–	2.6 ± 0.7	3.1 ± 1.4
SM total	72 ± 8	27 ± 5	1099 ± 33	6.0 ± 1.7	10.2 ± 3.3
Observed	72	27	1099	10	7

```
"name": "SR1cut_cuts",
"data": 7.32849661754466
"hi_data": 10.082253560804581
"lo_data": 4.574739674284739
```

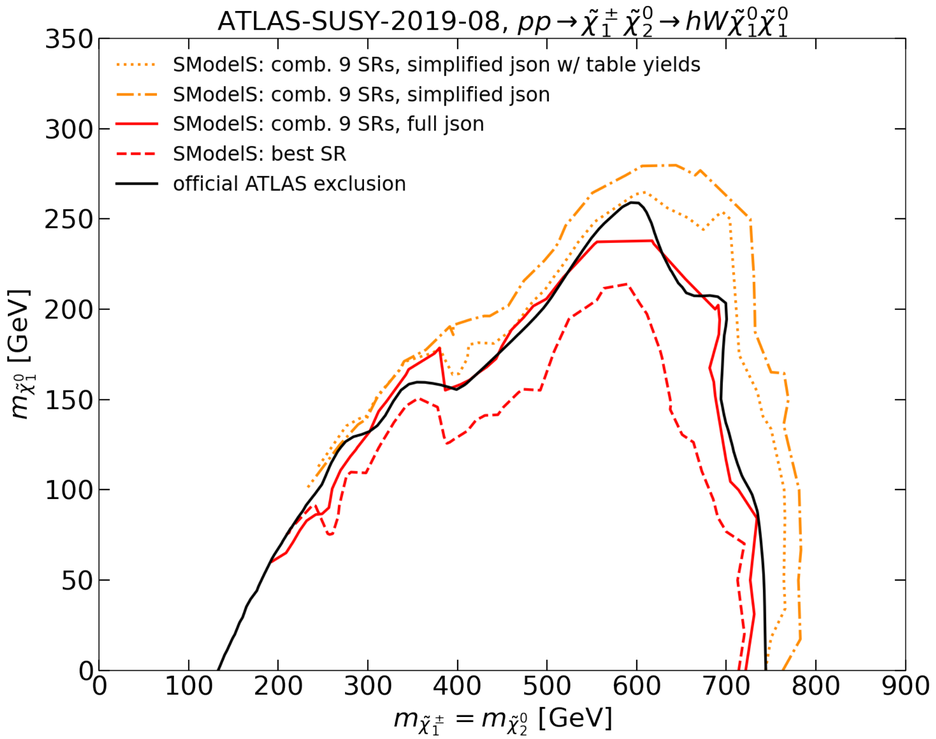
7.33 ± 2.75

```
"name": "SR2cut_cuts",
"data": 8.672689439530048
"hi_data":  10.684310727145936
"lo_data": 6.66106815191416
```

8.67 ± 2.01

How are SR correlations encoded in this?

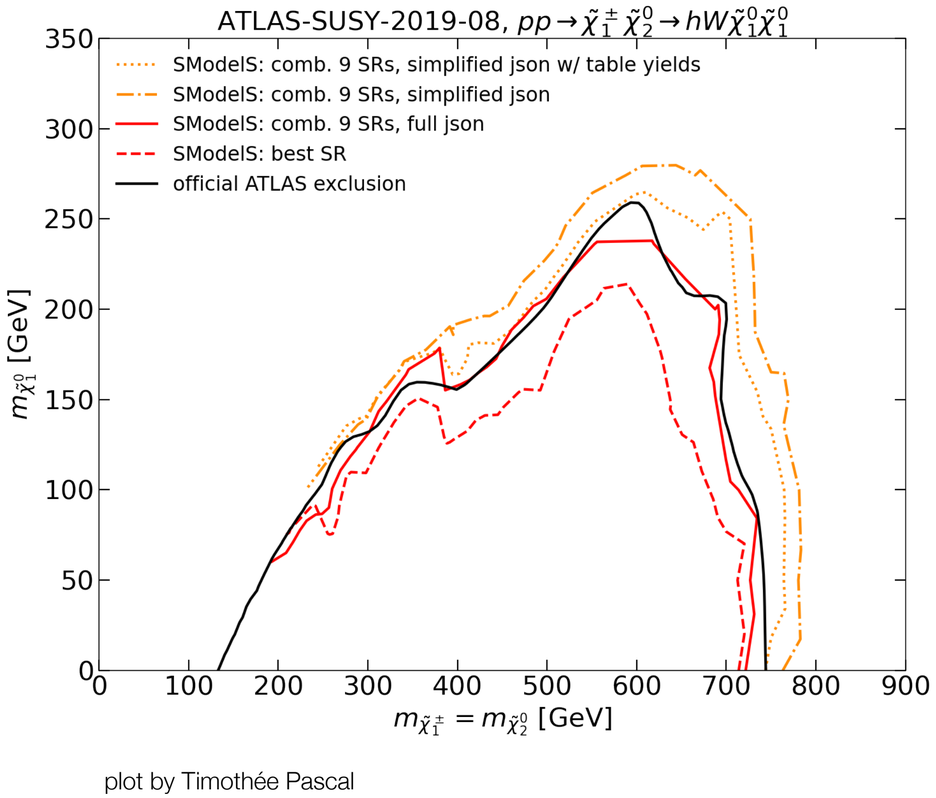
SModelS validation with full and simplified json llhds



plot by Timothée Pascal

SR	full json	table	simplified json
SR_HM_Low_MCT	4.19	4.1 ± 1.9	6.4 ± 2.0
SR_HM_Med_MCT	2.90	2.9 ± 1.3	4.6 ± 1.4
SR_HM_High_MCT	1.01	1.1 ± 0.5	1.6 ± 0.6
SR_MM_Low_MCT	5.38	4.6 ± 1.7	3.8 ± 1.6
SR_MM_Med_MCT	2.81	2.6 ± 1.3	5.9 ± 2.0
SR_MM_High_MCT	1.39	1.4 ± 0.6	1.8 ± 0.6
SR_LM_Low_MCT	10.96	8.8 ± 2.8	13.0 ± 2.6
SR_LM_Med_MCT	10.32	11.3 ± 3.1	9.8 ± 2.3
SR_LM_High_MCT	6.93	7.3 ± 1.5	6.7 ± 1.5

SModelS validation with full and simplified json llhds



plot by Timothée Pascal

SR-LM	All m_{CT} bins	Low m_{CT}	Medium m_{CT}	High m_{CT}
Observed	34	16 13.01 ± 2.63	11 9.84 ± 2.3	7 6.67 ± 1.45
Expected	27 ± 4	8.8 ± 2.8 ↗	11.3 ± 3.1 ↙	7.3 ± 1.5 ↙
$t\bar{t}$	16.2 ± 3.4	4.4 ± 2.2	7.3 ± 2.5	4.6 ± 1.2
Single top	2.7 ± 1.8	1.3 ± 1.1	$0.9^{+1.0}_{-0.9}$	0.6 ± 0.6
W+jets	5.5 ± 2.0	2.0 ± 0.9	2.4 ± 1.3	1.1 ± 0.5
Di-/Multiboson	0.67 ± 0.19	0.39 ± 0.13	$0.09^{+0.11}_{-0.09}$	0.18 ± 0.04
Others	2.23 ± 0.29	0.81 ± 0.25	0.64 ± 0.15	0.77 ± 0.12

SR-MM	All m_{CT} bins	Low m_{CT}	Medium m_{CT}	High m_{CT}
Observed	13	4 3.84 ± 1.59	7 5.93 ± 2.03	2 1.84 ± 0.63
Expected	8.6 ± 2.2	4.6 ± 1.7 ↙	2.6 ± 1.3 ↗	1.4 ± 0.6 ↗
$t\bar{t}$	2.7 ± 1.4	1.6 ± 0.9	0.8 ± 0.7	0.30 ± 0.24
Single top	2.7 ± 1.9	1.6 ± 1.5	$1.0^{+1.1}_{-1.0}$	$0.15^{+0.19}_{-0.15}$
W+jets	1.5 ± 0.7	0.6 ± 0.4	$0.3^{+0.4}_{-0.3}$	0.57 ± 0.26
Di-/Multiboson	0.29 ± 0.08	0.09 ± 0.04	0.065 ± 0.028	0.14 ± 0.06
Others	1.33 ± 0.27	0.69 ± 0.20	0.40 ± 0.13	0.24 ± 0.09

SR-HM	All m_{CT} bins	Low m_{CT}	Medium m_{CT}	High m_{CT}
Observed	14	6 6.44 ± 1.99	5 4.62 ± 1.43	3 1.60 ± 0.63
Expected	8.1 ± 2.7	4.1 ± 1.9 ↗	2.9 ± 1.3 ↗	1.1 ± 0.5 ↗
$t\bar{t}$	1.4 ± 0.5	0.8 ± 0.4	0.36 ± 0.25	0.22 ± 0.15
Single top	$2.0^{+2.4}_{-2.0}$	$0.9^{+1.5}_{-0.9}$	0.9 ± 0.9	$0.16^{+0.26}_{-0.16}$
W+jets	3.7 ± 1.0	1.9 ± 0.8	1.4 ± 0.8	0.45 ± 0.19
Di-/Multiboson	0.21 ± 0.06	0.057 ± 0.025	0.075 ± 0.027	0.08 ± 0.04
Others	0.74 ± 0.16	0.34 ± 0.09	0.19 ± 0.08	0.21 ± 0.08

Suggestion by Gaël: keep individual BG sources but with modifiers summed up