Deficit Hawks

Robust New Physics Searches with Unknown Backgrounds

arxiv.org/abs/2204.03264

github.com/JelleAalbers/deficithawks

KIPAC KAVLI INSTITUTE FOR PARTICLE ASTROPHYSICS & COSMOLOGY

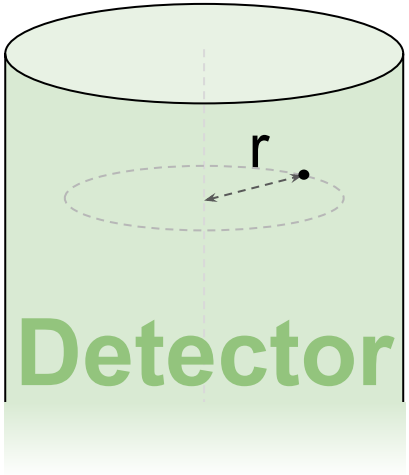
Jelle Aalbers
jaalbers@stanford.edu
21 July 2022

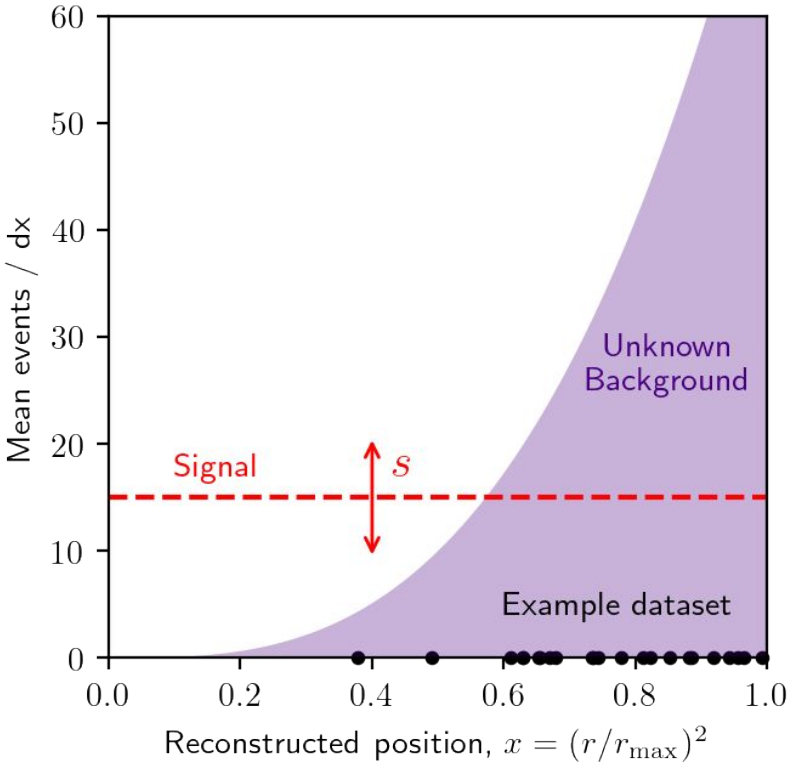
Why listen to me?

- **Stronger limits** in analyses with unknown backgrounds
- **Avoid hard choices** like fiducial volumes or energy thresholds
- **Not a whole new method**, small trick to add to existing analyses

“But we know all our backgrounds” – More likely you discard data with unknown backgrounds

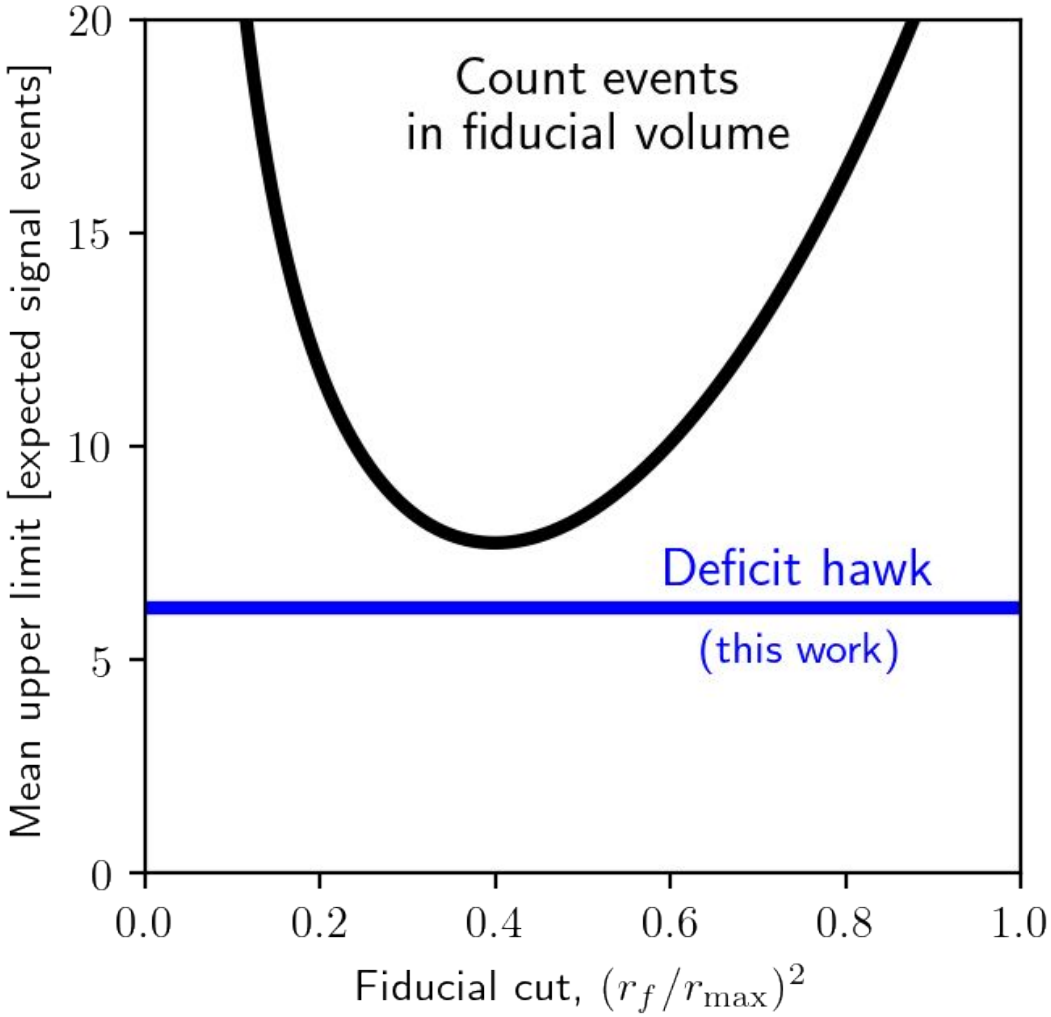
“We are so awesome we don’t need statistics” – No, you want to fight backgrounds in every way

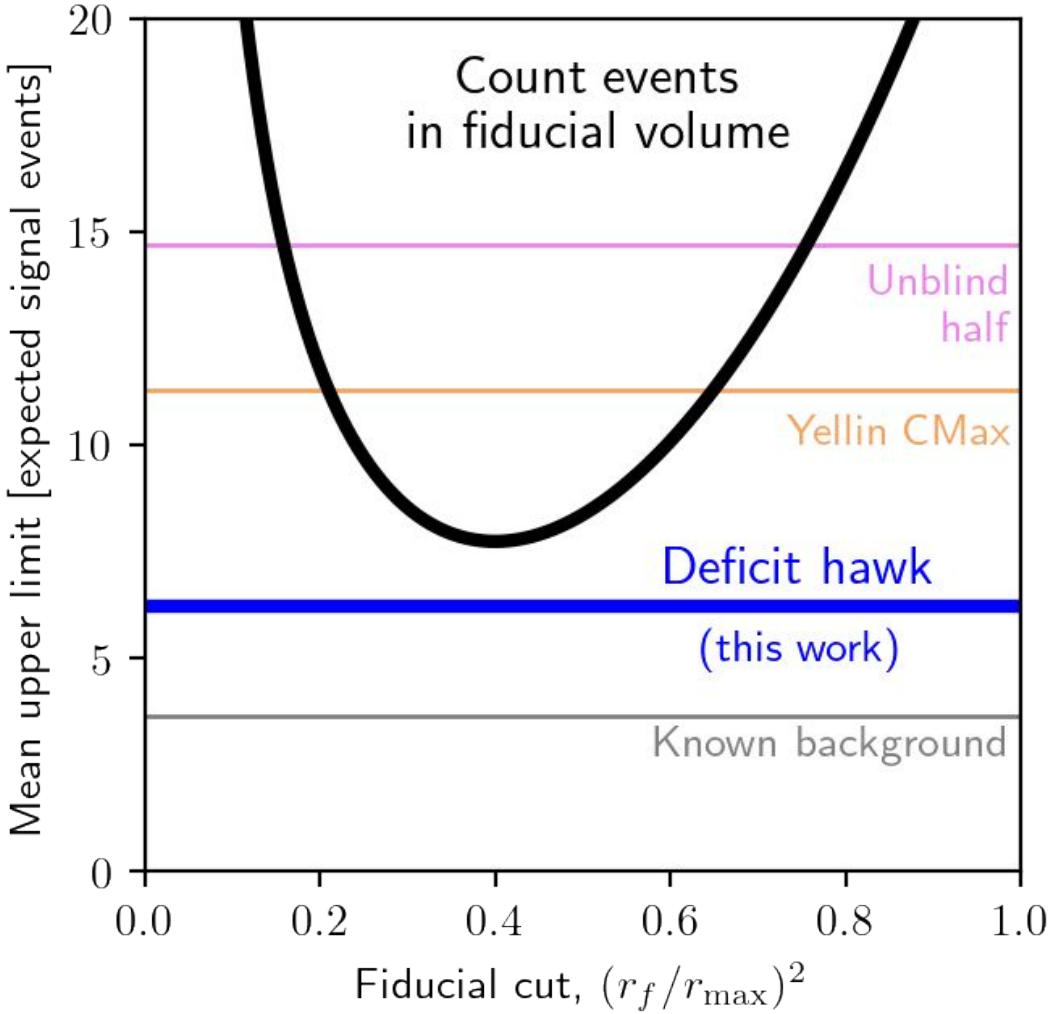




How to choose a fiducial volume?

(without peeking at the data!)


Count events
in fiducial volume
Deficit hawk
(this work)
Mean upper limit [expected signal events]
Fiducial cut, $(r_f/r_{\max})^2$
0
5
10
15
20
0.0
0.2
0.4
0.6
0.8
1.0

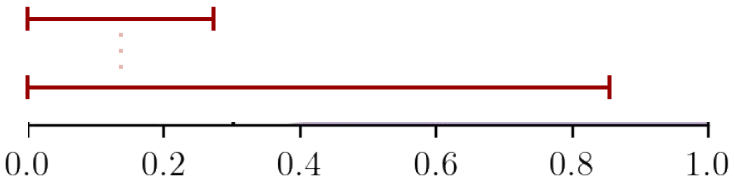


Sacrifice half your data, estimate best cut on that, then set limit on other half.

Optimum interval method
PRD 66:032005

Unbinned likelihood ratio test with all background known

1. **Choose several cuts**, and a **test statistic**



Observed events

$$\chi = \frac{N - \mu}{\sqrt{\mu}}$$

Expected signal events

2. Compute the statistic on the different cuts,
   keep the strongest deficit

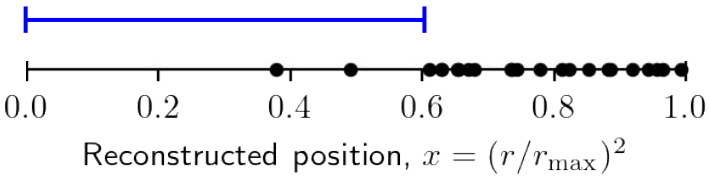
3. **Compare to simulations with true signal**
   to see if the deficit is too unlikely
   (p < 0.1 for 90% confidence limits)

Not original! Generalization of Yellin's optimum interval method (~20 years old)

Let's test: **16 signal events** / full dataset

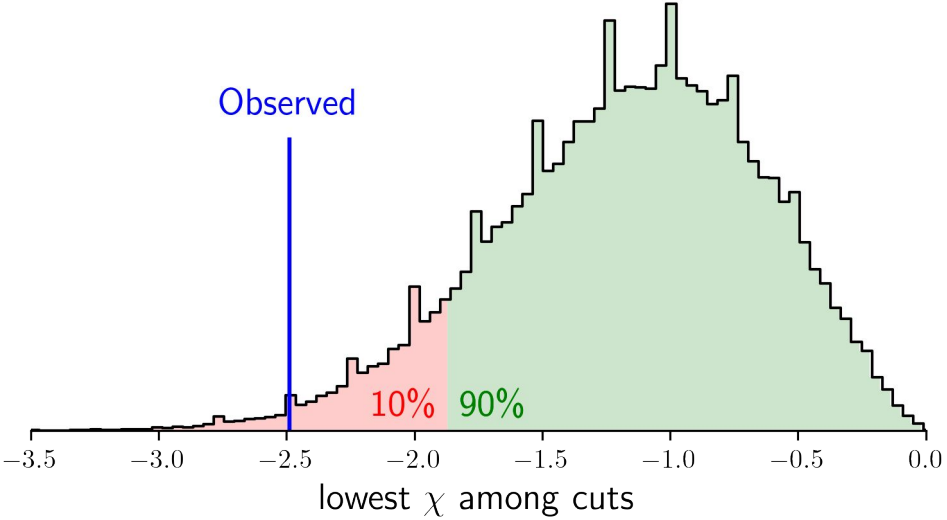
$$\chi = \frac{N - \mu}{\sqrt{\mu}}$$

2. Compute, **keep the strongest deficit**



$$\chi \approx \frac{2 - 10}{\sqrt{10}} \approx -2.5$$

3. Compare to simulations with true signal
to see if the deficit is too unlikely
(p < 0.1)

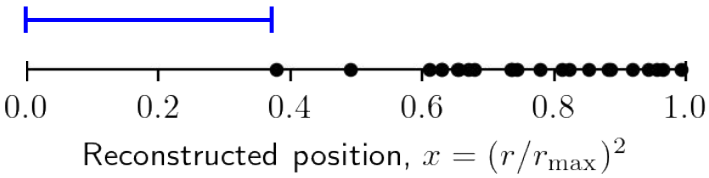


'16 signal events' is excluded!

Let's test: **8 signal events** / full dataset

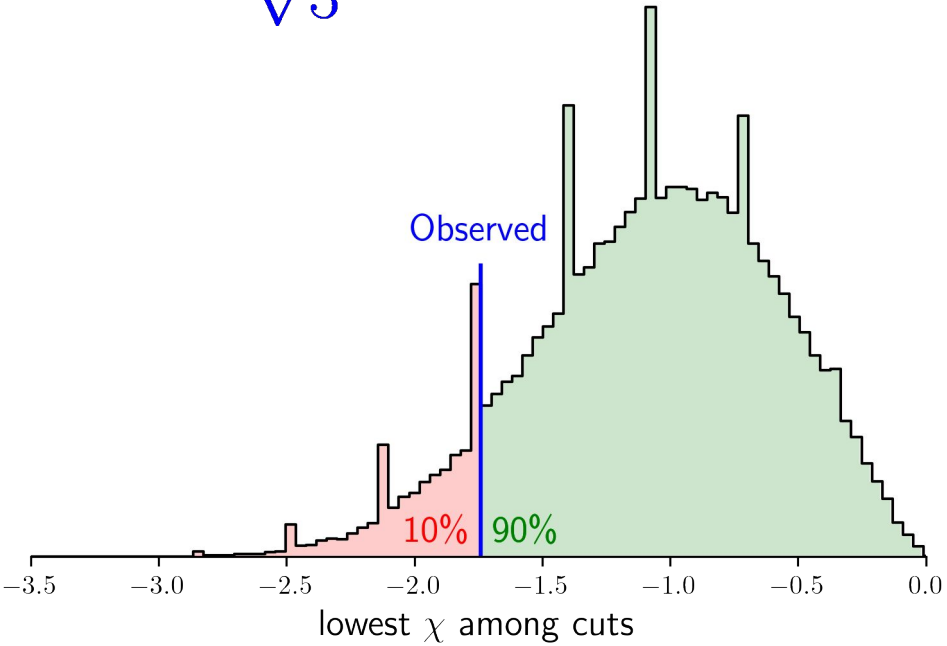
$$\chi = \frac{N - \mu}{\sqrt{\mu}}$$

2. Keep the strongest deficit

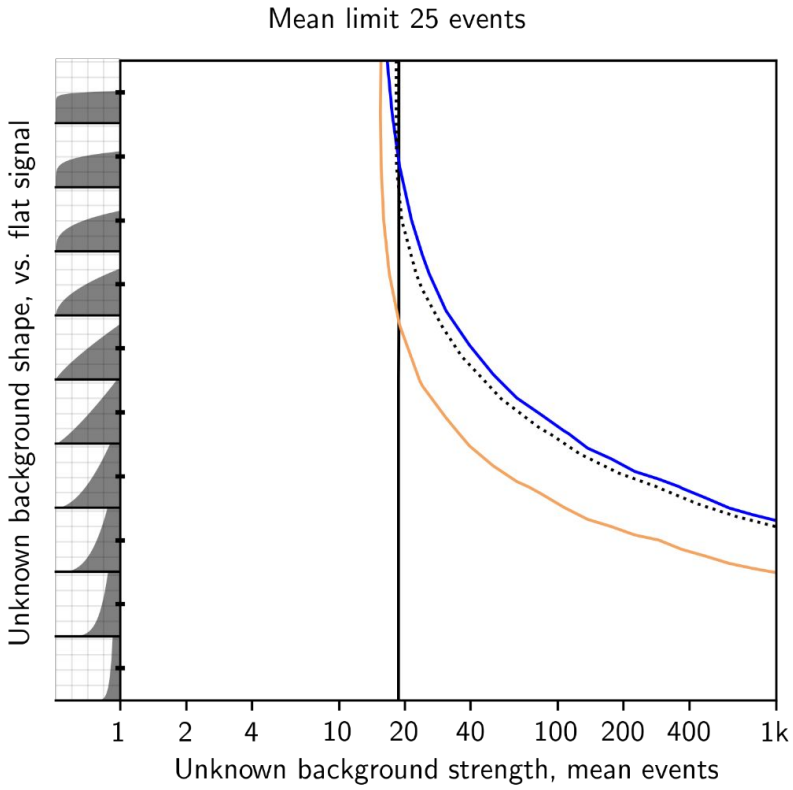


$$\chi \approx \frac{0 - 3}{\sqrt{3}} \approx -1.7$$

3. Compare to simulations with true signal
to see if the deficit is too unlikely
(p < 0.1)



'8 signal events' is *just* allowed → **< 8 signal events / full dataset**



Other background sizes and shapes

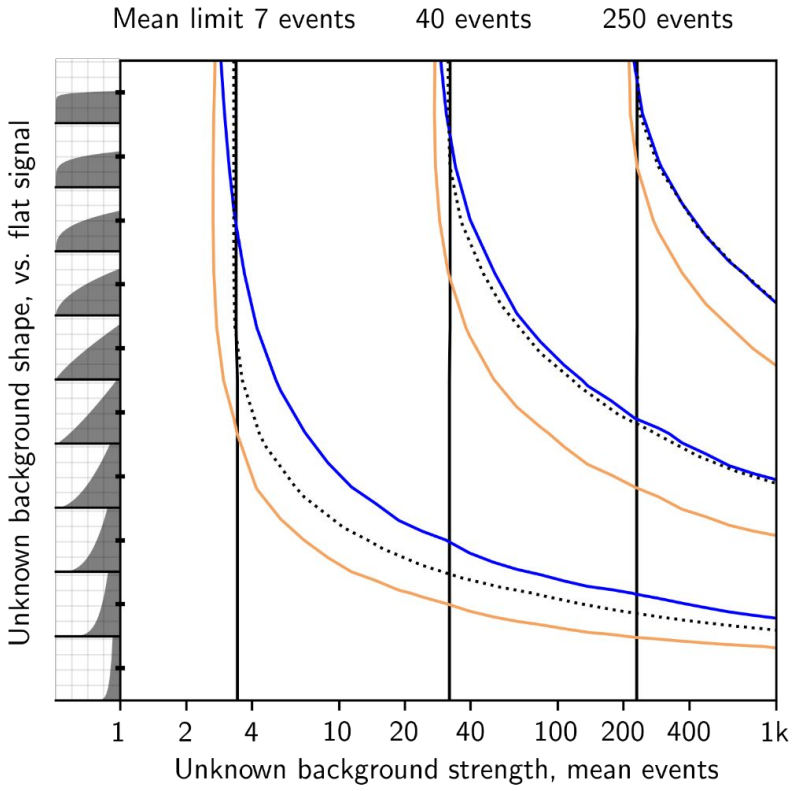
Curve further to the right
⇒ method tolerates more background

- Full detector
- All intervals in x
- Central cylinders (as before)
- Optimal fiducial volume for background

Plots in the paper use an 'augmented Poisson' for a fairer comparison – I didn't have time to introduce that here

Only test cuts that might help you!

Testing superfluous cuts weakens the result



Other background sizes and shapes

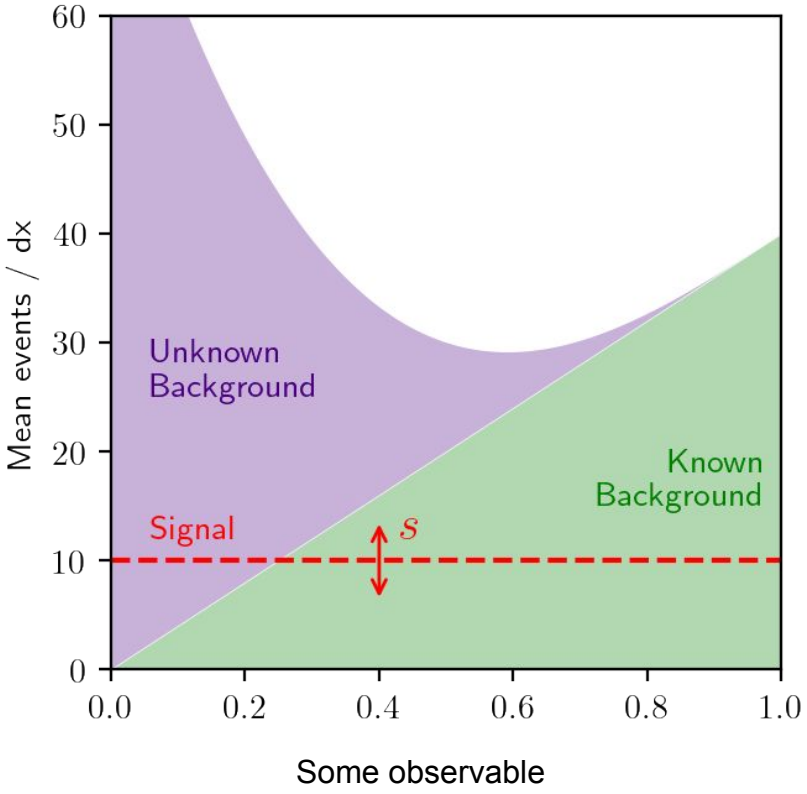
Curve further to the right
⇒ method tolerates more background

- Full detector
- All intervals in x
- Central cylinders (as before)
- Optimal fiducial volume for background

Only test cuts that might help you!

Testing superfluous cuts weakens the result

More complex example

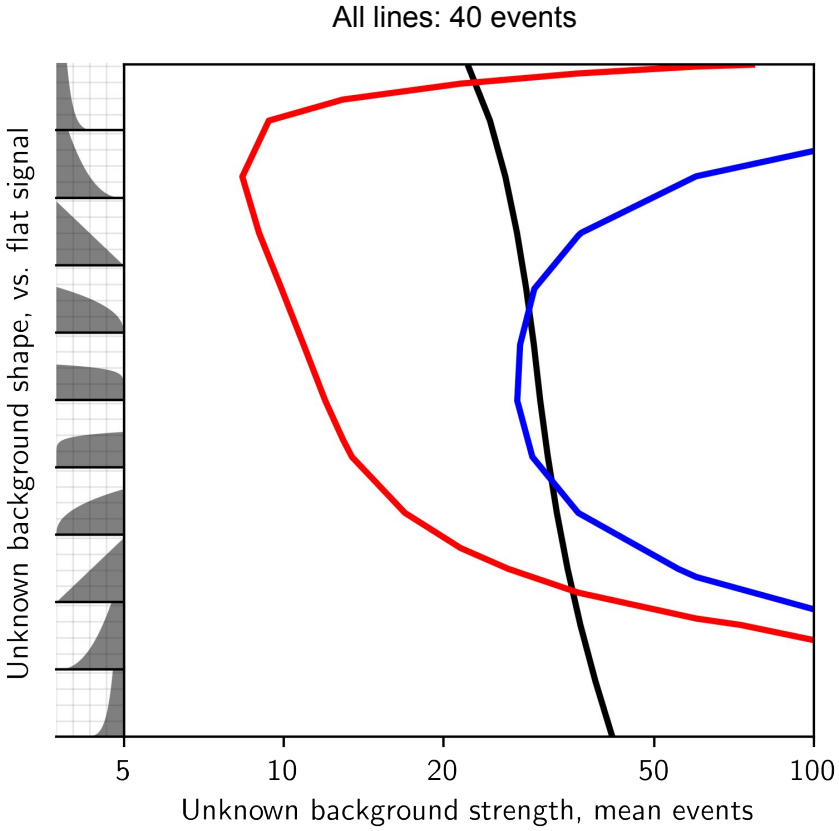


You cannot always guess good cuts

⇒ **Fine to test 'all intervals'** (or a subset)

With known backgrounds, cut and count is inferior to likelihood ratio tests

⇒ **Replace χ with a likelihood ratio**



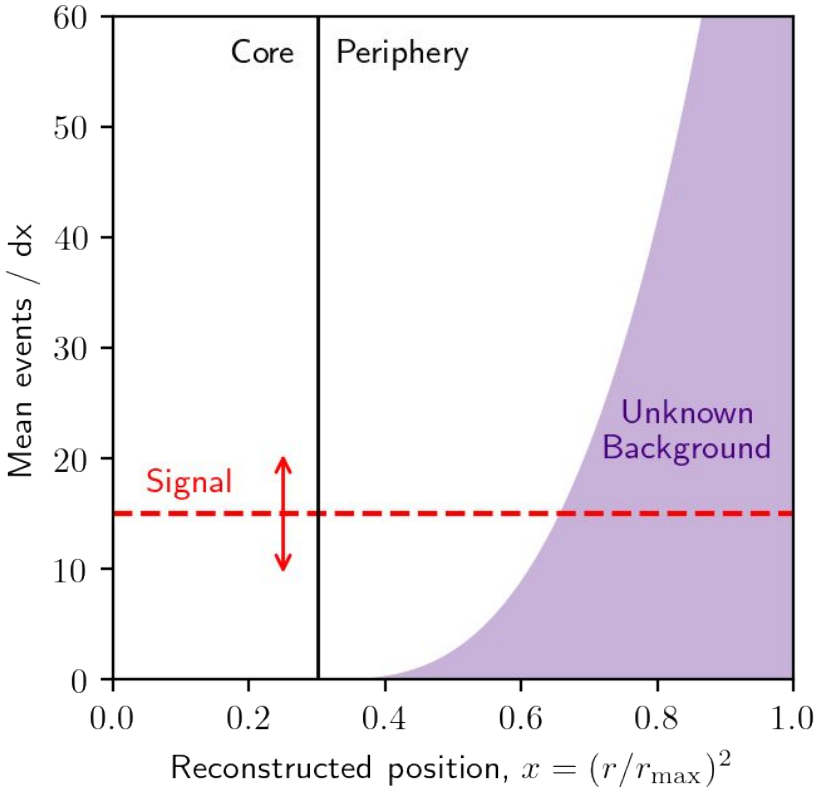
Curves further to the right
⇒ method tolerates more background

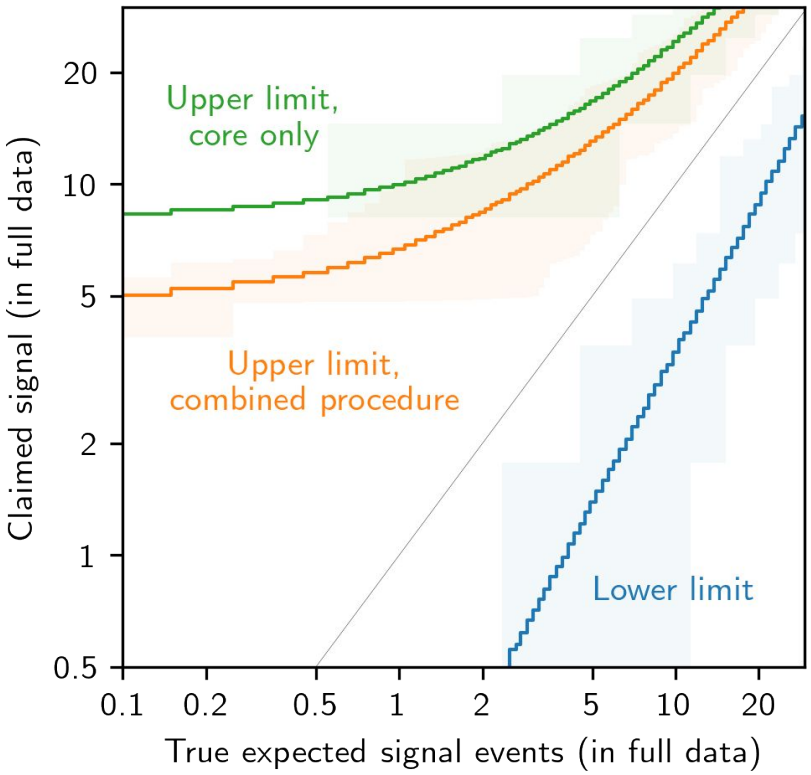
- Likelihood ratio in full detector
- All intervals, ignore known bg. (use χ)
- All intervals, likelihood ratio

Deficit hawks work great with likelihood ratios

(here the known background is fixed at 20 mean events, linearly rising)

Preserving discovery capability





Limitations

You must know the signal model

It's fine to use a conservative estimate

Overestimated efficiences create fake deficits

You cannot use nuisance parameters

… *except* in the core region of a discovery search

Instead, use a conservative (low) background model

With nuisance parameters, unknown backgrounds can cause *stronger* limits

Conclusions

- **Stop throwing away data –** test multiple cuts instead
- Try to test **useful cuts only**: use your physics knowledge
- **Use likelihood ratios** to discriminate known backgrounds
- **Not a whole new method**, small trick to add to existing analyses

In particular, you don’t have to give up discovery capability if you have it

Fully unknown Background knowledge **Fully known**

Just count events, set upper limit

Likelihood ratios

Maximum gap

Nuisance parameters

Optimum interval

Deficit hawks

"Well, it's probably larger at low energies…"

"I have a model but don't fully trust it"

"I can only model some components"

"Our model is fine, except near the walls"