Imperial College London

Some statistical issues in direct WIMP searches

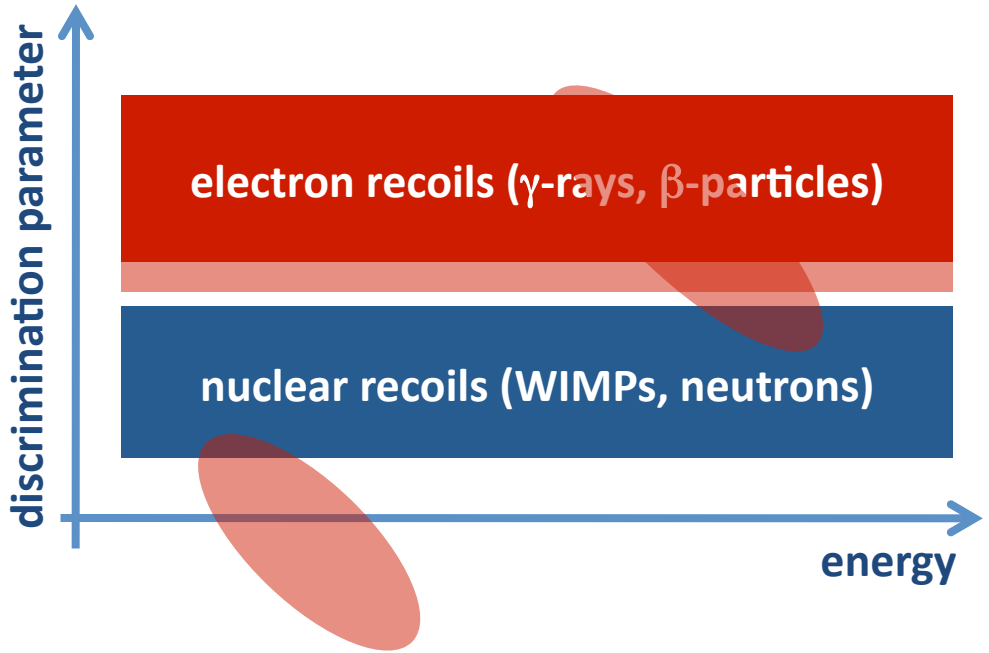
With a cautionary tale from ZEPLIN-III

Henrique Araújo & Alastair Currie

IMPERIAL COLLEGE LONDON

WIMP searches

- Looking for a handful of low-energy (few keV) nuclear recoils on the poorly understood tails of unexpected backgrounds
- Mitigation of the neutron background is the main experimental design driver, but this rarely materialises!
- 'One-bin particle physics' (most signal models show up in one bin)



WIMP searches (warts and all)

- The ideal WIMP search
  - Experiment runs for design period (background expectation ~1 event)
  - Signal region defined from signal calibration and data kept blind inside
  - Analysis optimised on signal & background calibrations and sidebands
  - Two-sided statistical procedure defined a-priori
- A typical WIMP search
  - Experiment runs only long enough to beat rivals (strictly blind?)
  - Not enough background calibration (ideally need 10x data)
  - Calibration does not model background (e.g. spatial distribution)
  - Cannot predict background reliably, blind analysis compromised
  - Open box to find rare topologies spoil sensitivity (blind analysis dead)
  - Upper limit produced using a-posteriori analysis (damage limitation)
  - Upgraded detector runs again, hopefully now better understood…
  - Result combined with previous

Some common statistical issues

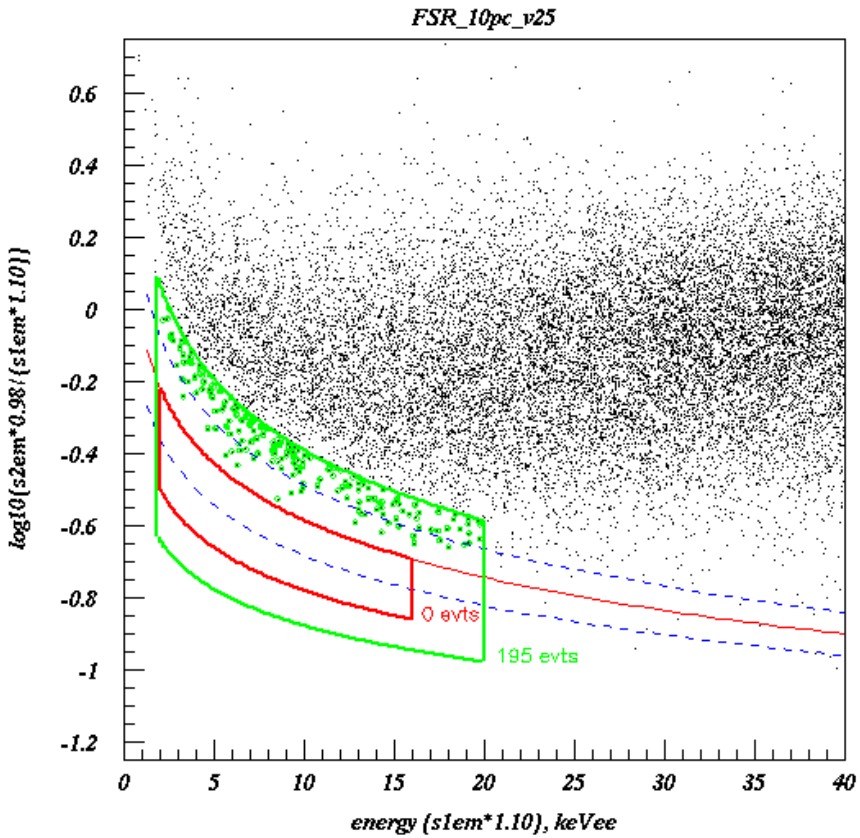
- Not “on/off” problem
  - No control of luminosity and calibration of background has systematics
  - Single signal bin, but variable background(s); one sideband
- Yet we move forward (albeit slowly)
  - Second/upgraded runs are less risky
  - Backgrounds in larger detectors are better behaved
    (but representative calibration at low energy gets even harder)
- Uncertainties in signal model – to keep the Bayesians happy
  - Nuclear physics, particle physics, but mainly astrophysics
  - Energy calibration and detection efficiency (e.g. variable Xe L_{eff})
- Blind analyses
  - ‘blind’ but not ‘fool’: a rare event search will reveal rare backgrounds
  - How to tune/debug data chain on an open sample << blind sample?
  - Really needed? Are we more reluctant to trust each other these days?

Discovery and limit-setting issues

- Single-sided or double-sided?
  - Small sensitivity increments: temptation for single-sided
  - Especially if you're not leading the pack!
- 3σ for discovery (consensus, or maybe not)
  - Little LEE, alternative hypothesis not (very) particle type/mass specific
- Feldman-Cousins
  - Addressing uncertainty in background: PLR extensions of original FC method
  - Multi-bin FC with non-uniform background using fewest assumptions
- Yellin single-sided techniques
  - Maximum gap/optimum interval/p-max: statistic of empty patches in data
  - Should not be applied a posteriori when gaps in data are obvious
  - Should a p-value be offered a posteriori? Is this flip-flopping?
- Likelihood analysis
  - Still searching near-Gaussian family of distributions that fit the main background
  - Only ZEPLIN-III attempted this, as far as I can tell.

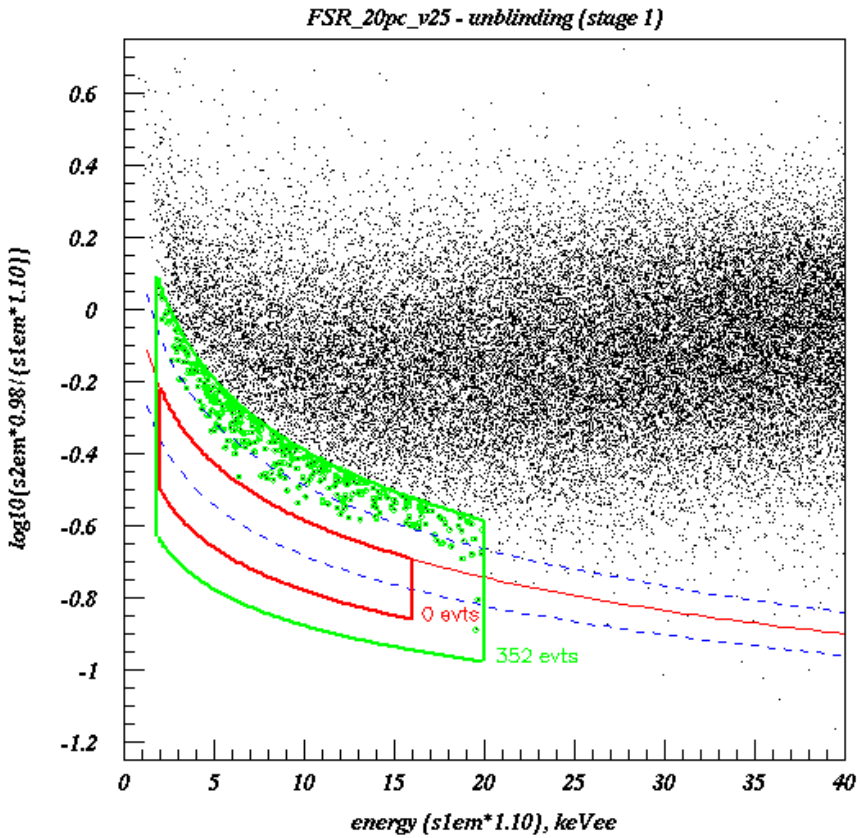
A blind analysis – 10% sacrificial data

- First science run of ZEPLIN-III : 11,000 data files over 83 days
- Analysis optimised on sacrificial 10% (files ending '1')
- Red region kept blind for remaining data (~50% signal acceptance)



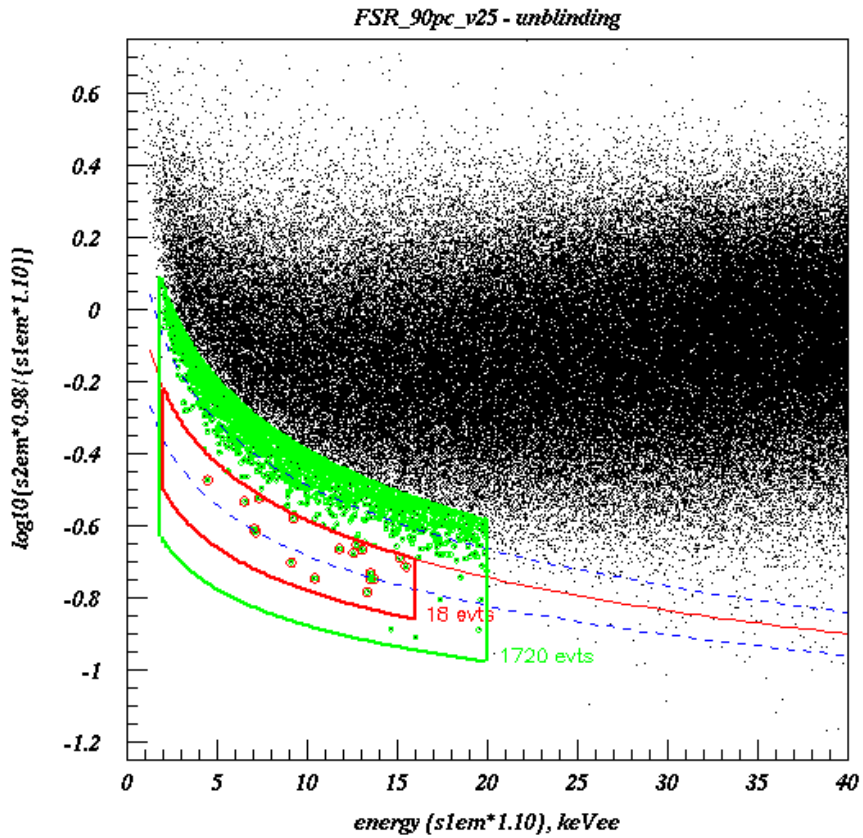
A blind analysis – 20% step 'just in case'

- Further 20% unblinded next (files ending '4' and '7')
- Could be sacrificed if something very wrong found
- Found to be empty!
- Expect ≤8 events (90% CL)



A blind analysis – opening remaining data

- Full opening of the box (remaining 7 file endings): 18 events!
- Distribution not compatible with signal



A blind analysis – is no longer

- Fault in vertex reconstruction found upon inspection:
  systematic mis-reconstruction of energy was origin of most events
- Blind analysis prevented us from spotting the problem early enough:
  - Not surprising that 10% data were clean (tuning the cuts tends to do this!)
  - But very unlucky with the next 20%!
- Re-analysis gave 7 events (top of box)
- μ<3.0 events (90% CL)

  Lebedenko et al 2009, PRD 90: 052010

- In general, failed blind analyses
  can lead to
  - Redefinition of the signal box
  - Recalculation of background estimates
  - Change of statistical analysis
  - (Pain)