

# Data analysis challenges in the signal dominated era

Panel: Van Den Broeck (chair), Green, Sberna, Vijaykumar  
with contributions from G. Ashton and K. Chandra

IXth Physics and Astrophysics at the eXtreme (PAX-IX) Workshop, King's College London  
9am, Tuesday 23/07/2024 (90 minutes)

To the polls...

[To Slido](#)



# Overview of the session

1. Panel introductions
2. Panel-introduced challenges
3. Other challenges

# The panel

## **Chris Van Den Broeck**

Utrecht University & Nikhef

Interests: GW parameter estimation, glitch mitigation, tests of GR, dense nuclear environments

## **Laura Sberna**

University of Nottingham

Interests:  
black hole perturbation theory, ringdown, environmental effects and LISA science

## **Stephen Green**

University of Nottingham

Interests: GW parameter estimation, machine learning, black hole perturbation theory, ringdown

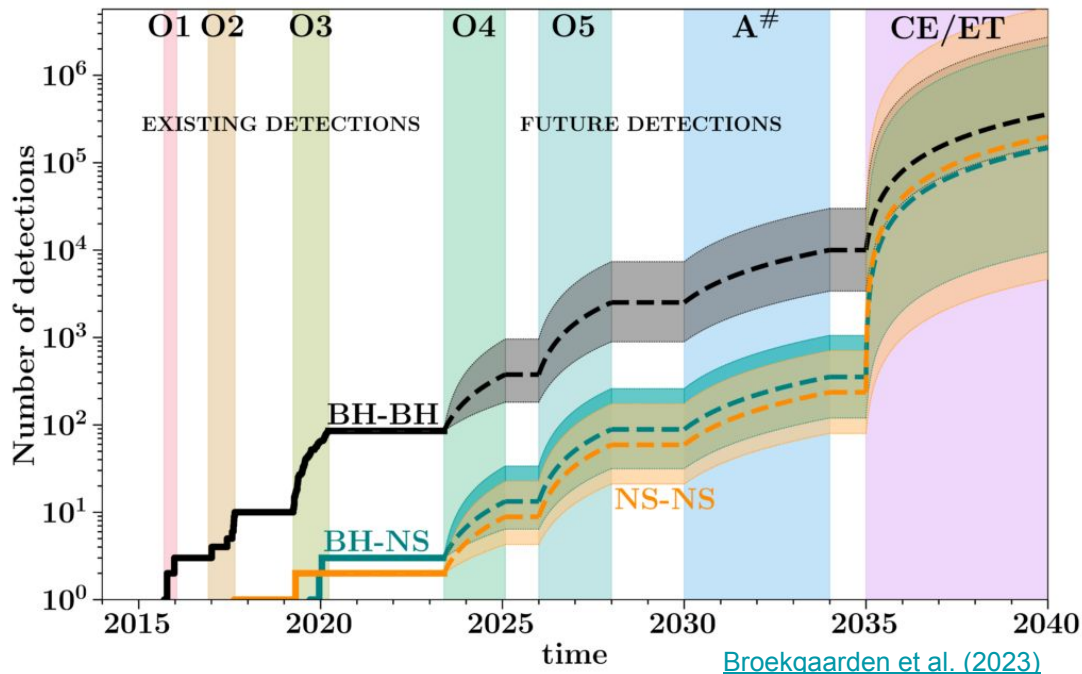
## **Aditya Vijaykumar**

Canadian Institute for Theoretical  
Astrophysics (CITA)

Interests: GW Parameter Estimation,  
Astrophysics of CBCs and their environments

# Challenge 1: Number and length of signals

More signals than we can handle? Is the **speed of PE** rising to the challenge?

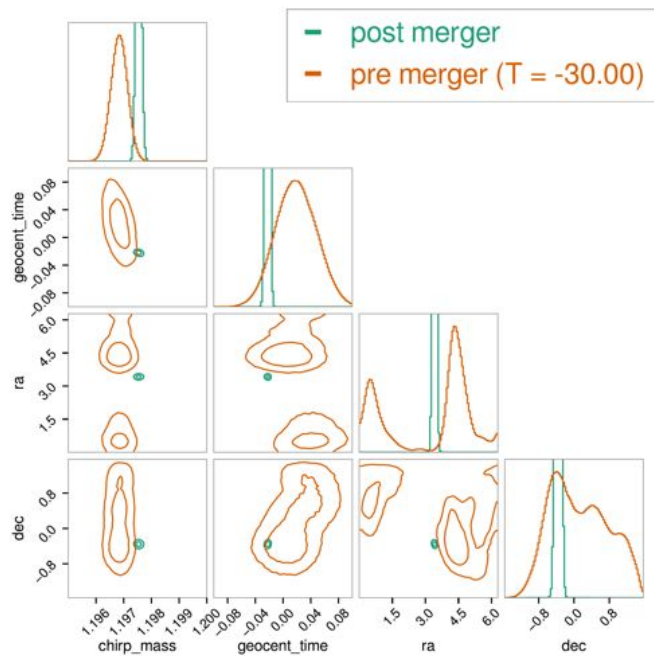


## Approaches

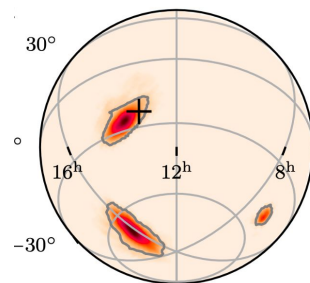
- **Faster likelihoods**
  - faster waveforms
  - multi-banding, heterodyning, ROQ, etc.
- **Efficient sampling**
  - Hamiltonian MC
  - AI-enhanced samplers
- **Simulation-based inference**
  - Neural networks / amortization

# Challenge 1b: Fast alerts for EM follow-up

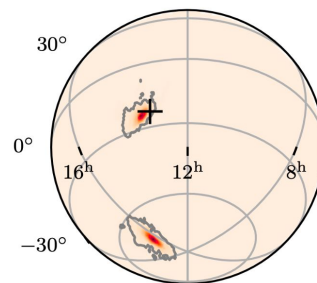
## Pre-merger analysis



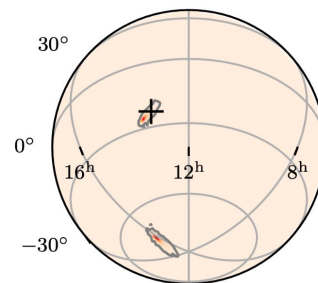
## BNS signals an hour before merger with XG



$T = -45$  min  
SNR = 82



$T = -30$  min  
SNR = 255



$T = -15$  min  
SNR = 664

[Dax et al \(2024\)](#)

# Challenge 2: Overlapping signals

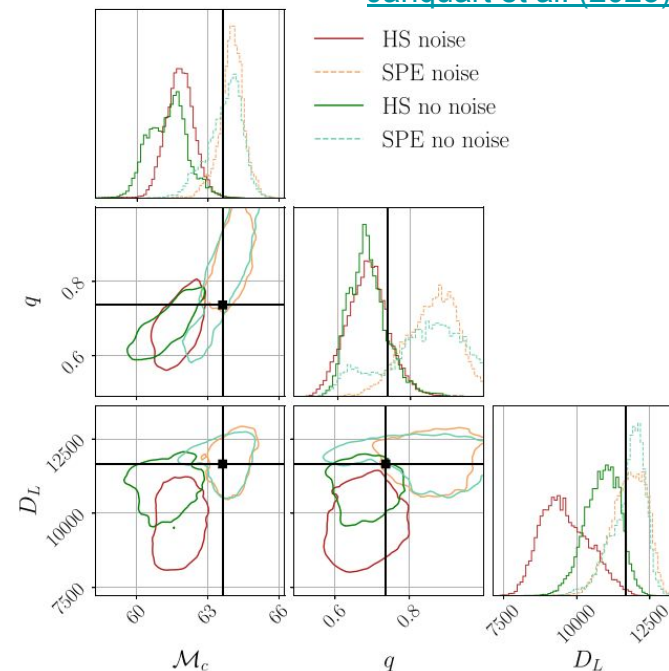
Standard PE frameworks: data = signal + noise

If data = *signal A* + *signal B* + noise, this leads to **bias**

Broadly two **approaches**:

- Hierarchical subtraction
  - Faster, but neglects correlations
- Joint analyses
  - Slower (2x dimensionality), but more precise

Overlap configuration	$N_{Overlaps}$ by region		
	Strong	Weak	Negligible
BBH+BBH	$5.6^{+9.1}_{-3.3}$	$17.0^{+27.0}_{-9.8}$	$13000.0^{+8300.0}_{-4900.0}$
BNS+BNS	$0.13^{+0.71}_{-0.12}$	$26.0^{+140.0}_{-25.0}$	$14000.0^{+22000.0}_{-11000.0}$
BNS+BBH	$0.11^{+0.18}_{-0.07}$	$22.0^{+36.0}_{-13.0}$	$13000.0^{+8300.0}_{-4800.0}$



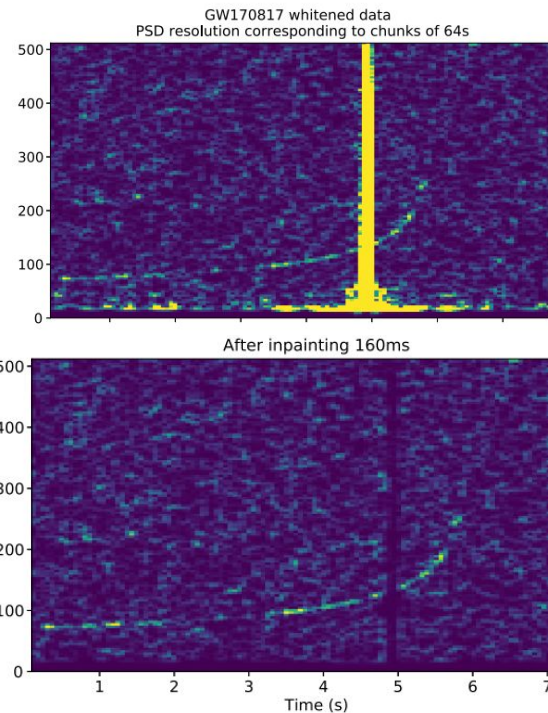
# Challenge 3: Non-stationary non-gaussian noise

For *long-lived* signals, the noise will be non-gaussian and non-stationary.

Failure to account for these could lead to imperfect estimates e.g. of luminosity distance. Impact on EM follow-up?

## Approaches:

- for non-gaussianities (glitches): inpainting (e.g. Zackay et al 2021) or similar methods
- for non-stationary (gaussian) noise: breaks the simple form of the frequency domain likelihood. Estimate PSD in segments (e.g. Kumar et al 2022) as a function of time? Fully time-domain analyses?
- Simulation-based inference

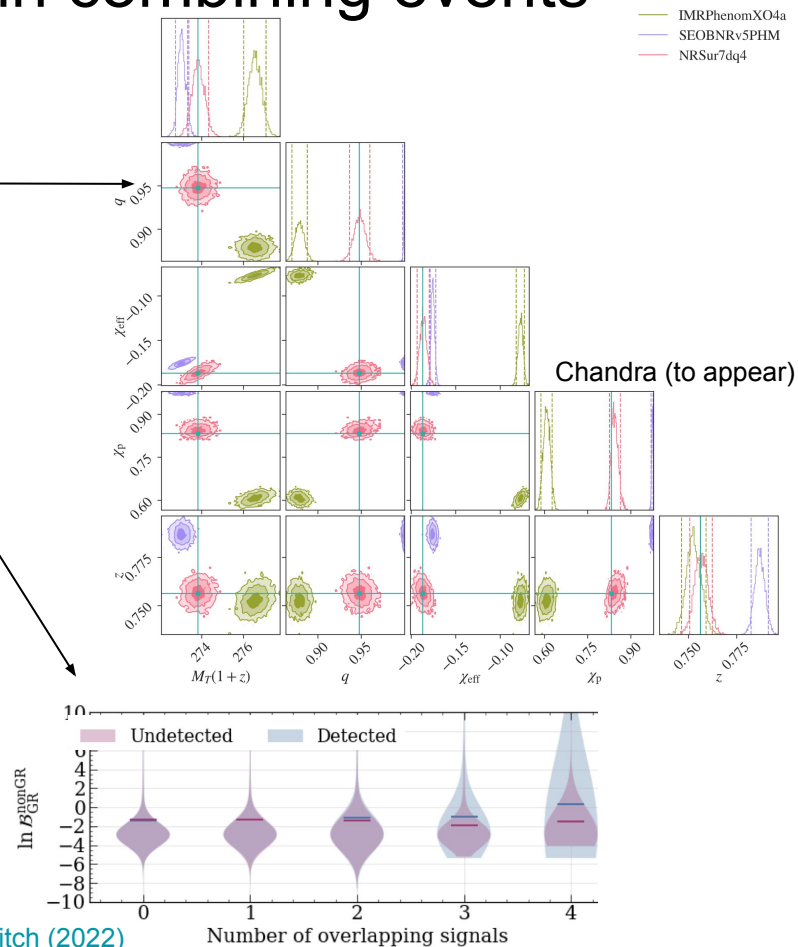


Zackay et al 2021



# Challenge 4: Role of systematics in combining events

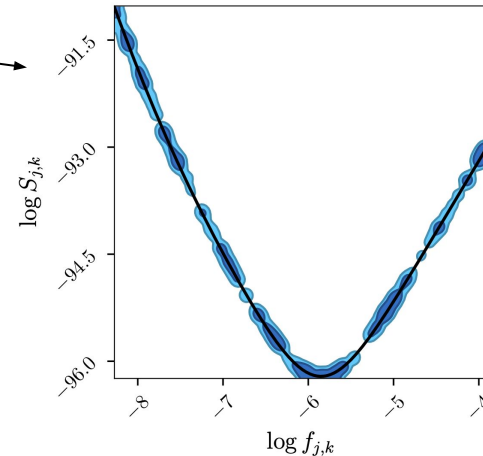
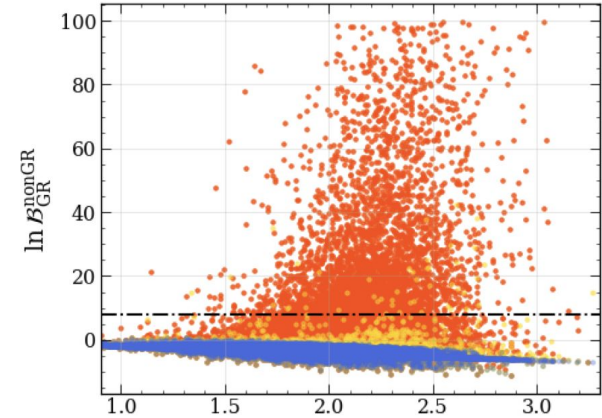
- Which **systematics**:
  - *waveform systematics*
  - *overlapping signals* and non-stationary noise: do these average...or not?
  - Data calibration, noise estimation
- Events **combined** for:
  - TGR
  - astrophysical population
  - cosmology
  - ...
- **Mitigation strategies**:
  - calibration of waveforms, especially for high SNR events
  - bias-variance trade-off?
  - marginalization over waveforms systematics?



# Challenge 5: Search for new (astro)physics

[Hu and Veitch \(2022\)](#)

- **Example: tests of GR:**
  - “**golden events**” (high SNR and no overlap): more vulnerable to waveform systematics
  - **overlapping events:** marginalise?
- How do we search for **stochastic backgrounds**?  
Model the *noise and the background together*?  
LISA: [Baghi et al. \(2023\)](#)
- How do we search for **non-CBC events** (e.g. hyperbolic encounters)?
- Numerous **quiet** events, potentially astrophysically interesting (e.g. at high-z).
  - But also more affected by e.g. noise model assumptions and glitches.
  - Can detect, but hard to characterize/estimate parameters precisely



$\log_{10} \text{SNR}$

**Blue:** “perfect waveform”  
(but overlapping signals)  
**red:** “current waveform”  
**yellow:** “future waveform”

To the polls...

[To Slido](#)



## Challenge 6: Searches for signals