Next-generation astrophysical inference across the interdisciplinary frontier

Will Handley
<wh260@cam.ac.uk>

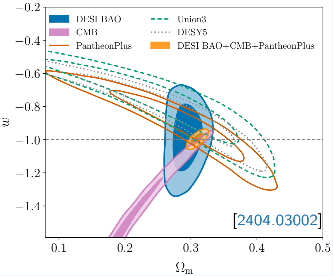
Royal Society University Research Fellow
Astrophysics Group, Cavendish Laboratory, University of Cambridge
Kavli Institute for Cosmology, Cambridge
Gonville & Caius College
willhandley.co.uk/talks

24th May 2024

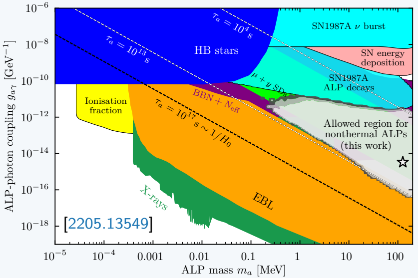
The future of astronomy is interdisciplinary

- Across astronomy, combining data and disciplines will be the key to the next breakthroughs.

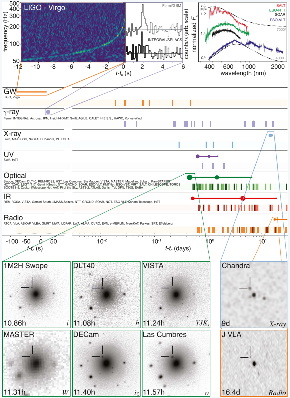
CMB+BAO



HEP+Astro



GW170817 [1710.05833]



- We have spent the last 5 years hair-splitting “parkable” tensions.
- The next 25 years of data confront the real tensions in our understanding of the Universe.
- I aim to show how my research programme is preparing us for this interdisciplinary challenge.

Interdisciplinary work to date

- CMB cosmology & inflation
- Cosmological tension quantification
- **21cm cosmology**
- Radio Instrumentation
- **Gravitational waves**
- Exoplanets
- **Particle physics**
- Theory of machine learning
- Gravitational Gauge theories
- Atomistic chemistry
- **Industrial applications**
- ...

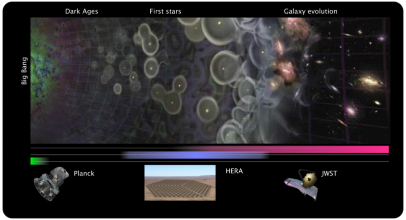
20 minutes is not enough time to cover a decade's publishing.

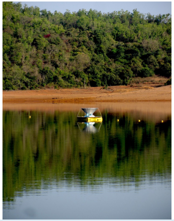
arxiv.org/a/handley_w_1.html

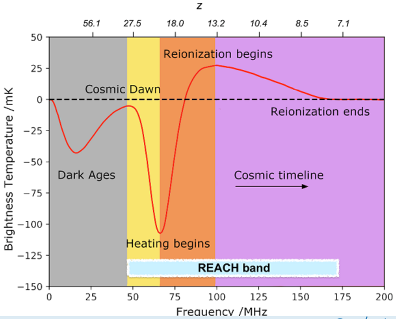
Will showcase a targeted subset.

REACH: Global 21cm cosmology [2210.07409]

- Imaging the universal dark ages using CMB backlight.
- 21cm hyperfine line emission from neutral hydrogen.
- Global experiments measure monopole across frequency.
- Challenge: science hidden in foregrounds $\sim 10^4\times$signal.
- Lead data analysis team (REACH first light in January)
- Nested sampling woven in from the ground up (calibrator, beam modelling, signal fitting, likelihood selection).
- All treated as parameterised model comparison problems.







Gravitational waves

Metha Prathaban
PhD

- Nested sampling has been used in GW since the beginning [GW150914]
- Work with Alvin Chua & Chris Moore on transdimensional sampling for EMRI [1803.10210]
- Recent work with Metha Prathaban showing new chain-based approaches for improving precision [2404.16428]

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

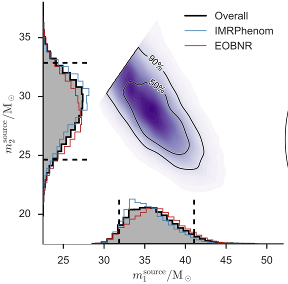
MNRAS **478,** 28–40 (2018) doi:10.1093/mnras/sty1079
Advance Access publication 2018 April 28

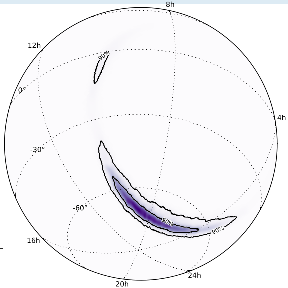
Towards a framework for testing general relativity with extreme-mass-ratio-inspiral observations

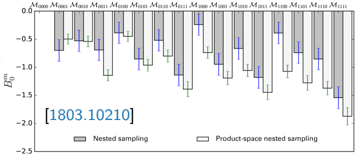
A. J. K. Chua,[1,2★] S. Hee,[3,4] W. J. Handley,[3,4,5] E. Higson,[3,4] C. J. Moore,[6,7] J. R. Gair,[8] M. P. Hobson[3] and A. N. Lasenby[3,4]

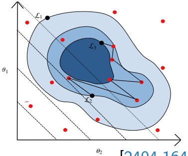
Costless correction of chain based nested sampling parameter estimation in gravitational wave data and beyond

Metha Prathaban[1,2,3★] and Will Handley[1,2,4†]









[2404.16428]

- Discussed use of `margarine` [2207.11457] as alternative to GW hierarchical modelling at inaugural data science discussion group

GAMBIT

Interdisciplinary case studies

- GAMBIT is an interdisciplinary community and software framework.
- Like `CosmoMC`/`Cobaya`/`Bilby`, an organiser of data, likelihoods & theory, including:
  - Collider data (e.g. LHC)
  - Direct detections (e.g. XENON1T)
  - Cosmology (`MontePython`)
  - Astrophysics (e.g. Bullet Cluster, Supernovae)
  - Pulsar timing
  - . . . & much more
- `GravBit` and `LowEnergyBit` arising from GAMBIT@KICC workshop

Felix Kahlhoefer et al

GAMBIT cosmo/DM working group

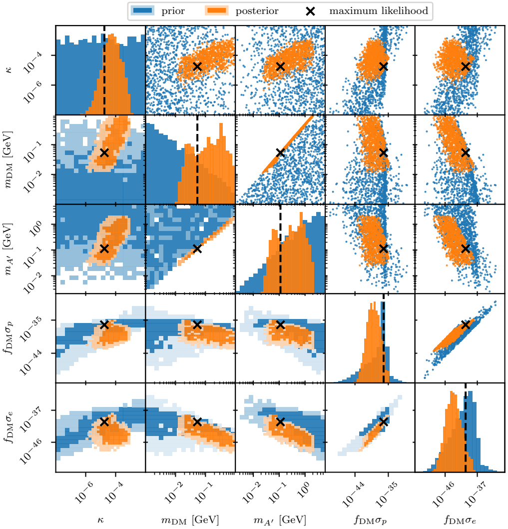
- Physical model of sub-GeV thermal dark matter with a dark photon mediator A:

$$\mathcal{L}_{\text{int}} = -\frac{1}{2}m_{A'}^2 A'^{\mu}A'_{\mu} - \frac{1}{4}A'^{\mu\nu}A'_{\mu\nu} - \kappa e A'^{\mu}\sum_f q_f \bar{f}\gamma_\mu f\,,$$

- Constrain using cosmological, astrophysical, accelerator & direct detection data.
- Bayesian Model comparison of Fermion ψ vs scalar Φ models (scalar preferred).

$$\mathcal{L}_\psi = \bar{\psi}(i\partial\!\!\!/ - m_{\text{DM}})\psi + g_{\text{DM}}A'^{\mu}\bar{\psi}\gamma_\mu\psi\,,$$

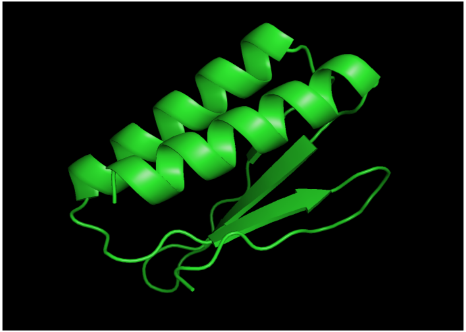
$$\mathcal{L}_\Phi = |\partial_\mu\Phi|^2 - m_{\text{DM}}^2|\Phi|^2 - g_{\text{DM}}^2 A'_\mu A'^{\mu}|\Phi|^2 + i g_{\text{DM}} A'^{\mu}\left[\Phi^*(\partial_\mu\Phi) - (\partial_\mu\Phi^*)\Phi\right]\,,$$



PolyChord Ltd: interdisciplinary R&D spinout
Interdisciplinary case studies

Catherine Watkinson
Senior Data Scientist

- Techniques have been spun-out (PolyChord Ltd) to:
- Protein folding
  - Navigating free energy surface.
  - Computing misfolds.
  - Thermal motion.
- Nuclear fusion reactor optimisation
  - multi-objective.
  - uncertainty propagation.
- Telecoms & DSTL research (MIDAS)
  - Optimising placement of transmitters/sensors.
  - Maximum information data acquisition strategies.

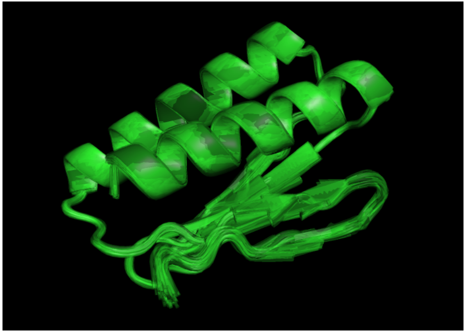


PolyChord Ltd: interdisciplinary R&D spinout

Interdisciplinary case studies

Catherine Watkinson
Senior Data Scientist

- Techniques have been spun-out (PolyChord Ltd) to:
- Protein folding
  - Navigating free energy surface.
  - Computing misfolds.
  - Thermal motion.
- Nuclear fusion reactor optimisation
  - multi-objective.
  - uncertainty propagation.
- Telecoms & DSTL research (MIDAS)
  - Optimising placement of transmitters/sensors.
  - Maximum information data acquisition strategies.

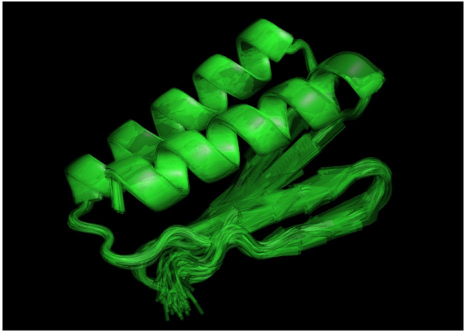


PolyChord Ltd: interdisciplinary R&D spinout

Interdisciplinary case studies

Catherine Watkinson
Senior Data Scientist

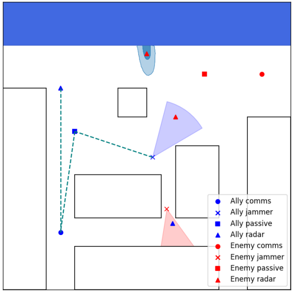
- Techniques have been spun-out (PolyChord Ltd) to:
- Protein folding
  - Navigating free energy surface.
  - Computing misfolds.
  - Thermal motion.
- Nuclear fusion reactor optimisation
  - multi-objective.
  - uncertainty propagation.
- Telecoms & DSTL research (MIDAS)
  - Optimising placement of transmitters/sensors.
  - Maximum information data acquisition strategies.



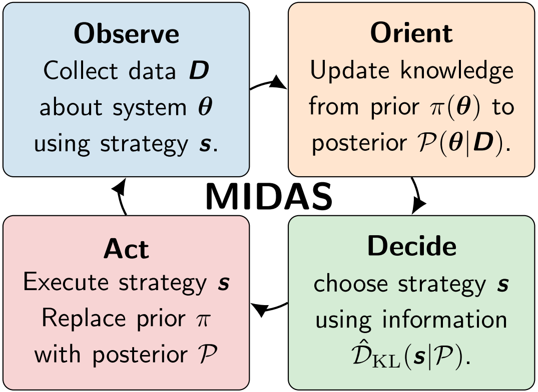
DSTL: Bayesian OODA loops

Interdisciplinary case studies

- Work through Isaac Newton Institute with Defence Science & Technology Laboratory.
- Quantification of "OODA" loop concept from litigation, business, law enforcement, management and military strategy



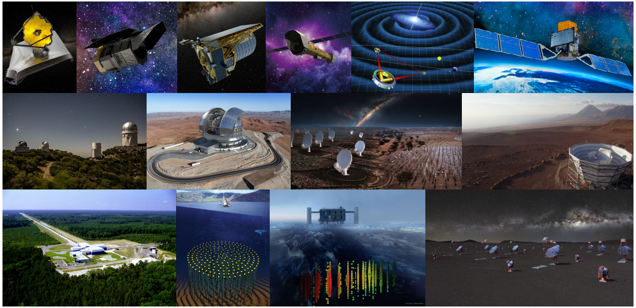
- Two-way research interaction between government and academia.
- techniques now being used in REACH antenna design [2309.06942]



Beginning the golden age of astronomy data

- Over our research lifetimes we will see next-generation data rates across the electromagnetic spectrum & beyond:

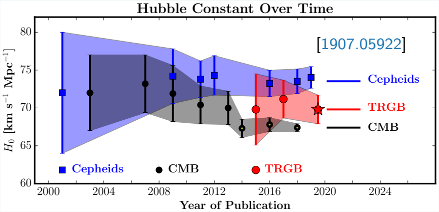
Radio	SKA *et al*
Micro	SO/CMB-S4/LiteBIRD
IR	JWST, Roman
Optical	Euclid, DESI, Rubin, EELT
X-ray	Athena
Gamma-ray	e-ASTROGAM
Gravitational	LIGO/LVK$^+$/LISA
Particle	CTA, IceCube, KM3NeT



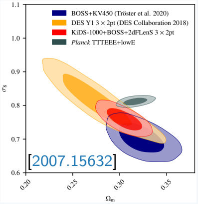
- This ever-increasing statistical weight will mean true accuracy demands rigorous attention to systematics.
- This applies to all of cosmology, astrophysics, particle physics and beyond.

Tensions in cosmology

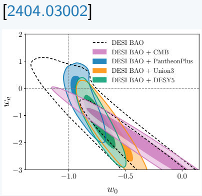
Hubble

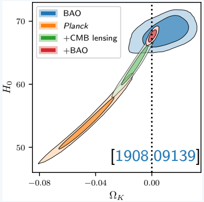


Weak lensing



other $w_0/\Omega_K/\nu$?





- Tensions have been appearing in cosmology over the last five years.
- Though their significance may be debatable, they have revealed a gap in the armour of modern data analysis.
- Likelihood-based methods have to assume a fiducial cosmology.
- This will cast greater doubt on the robustness of our conclusions from stage 4 surveys and beyond.

DESI BAO paper: [2404.03000]

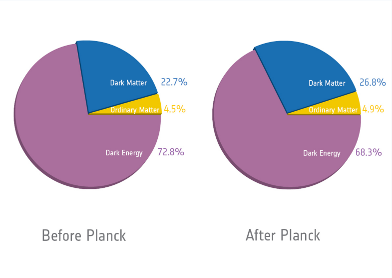
5.4 Systematics due to the assumption of the fiducial cosmology

In this subsection, we summarize the impact of using a wrong fiducial cosmology in the BAO analysis, which were extensively studied in [51].

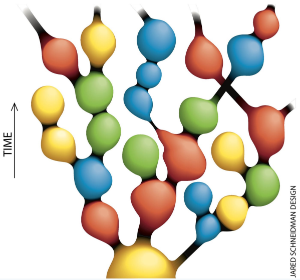
The choice of reference cosmology comes into play at three different stages. First, we assume a set of cosmological parameters when converting redshift measurements into distances; we term this the **grid cosmology**. The difference between the grid and true cosmologies causes a distortion of the BAO scale along and across the line of sight [10], which is quantified by the parameters $q_{\parallel} = H_{\rm grid}(z)/H_{\rm true}(z)$ and $q_{\perp} = D_{\rm true}(z)/D_{\rm grid}(z)$ respectively. Without the three-dimensional standard ruler like BAO, this effect is somewhat degenerate with the redshift-space distortions, but with a sufficiently large data set, such as we have with DESI, and with the BAO feature it is possible to distinguish between the two [145, 146]. Second, a **template cosmology** is chosen in order to compute the linear power spectrum, which is then used to create the model power spectrum for the fitting ($P_{\rm nw}$ and P_w in Eq. (4.1)). The effect of fixing the template is interpreted as an additional isotropic rescaling of the distances by a factor of $r_d^{\rm tem}/r_d^{\rm true}$. Lastly, the values for the linear bias b_1 and the growth rate $f(z)$ input into the reconstruction algorithm are cosmology dependent, affecting the estimation of the displacement field. The separate effect of both the template and the values assumed for reconstruction have been comprehensibly studied in the past [147–150], as well as the joint effect of consistently changing the reference cosmology in the whole pipeline [114], while the separate effect of the grid is less explored [151]. Potential systematic shifts of the order of a few tenths of a percent in the alpha values have been reported in the most extreme scenarios.

The real tensions in the room

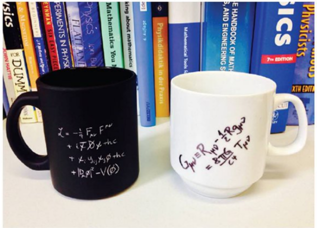
Dark tension



Initial conditions



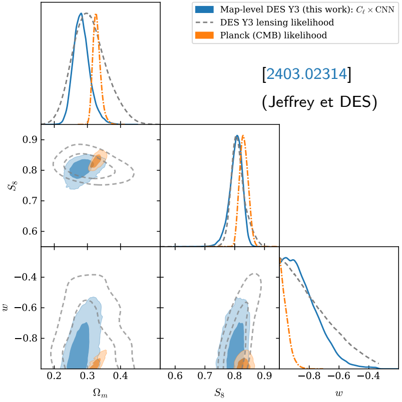
Quantum gravity



- These are existential questions that don't go away if ΛCDM is phenomenologically correct. . .
- . . . and if it is, we will need an interdisciplinary approach that goes far beyond cosmology (using astronomy, collider & experimental physics).
- The GW redshift frontier, time domain astronomy & 21cm SKA will be the key tools for challenging our astronomical understanding.

- Traditional Likelihood-based inference (LBI) requires knowledge of the likelihood $P(D|\theta)$
  - For the CMB, it is possible to compute Probability(Sky$|\Lambda$CDM) [with caveats]
  - For almost everybody else (different models, more evolved systems), the likelihood is approximate
- Simulation-based inference learns the likelihood from physical simulations $\theta \to D$.

- Can extract nonlinear information from data.
- Do not need to assume a fiducial model, since data covariance is implicit in simulations.
- Users do not need to know advanced statistics
  - For this reason alone, it will come to dominate the next generation $\Rightarrow$ imperative to build a principled understanding!
- My interest has been in understanding how essential machine learning is by building analytics `github.com/handley-lab/lsbi`
- Have also made progress on the first simulation-based nested sampler `PolySwyft` (presenting at EuCAIFCon)



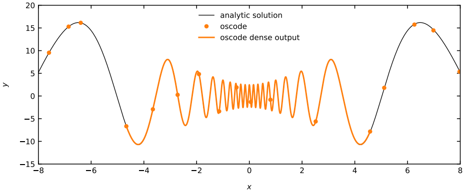
[2403.02314]
(Jeffrey et DES)

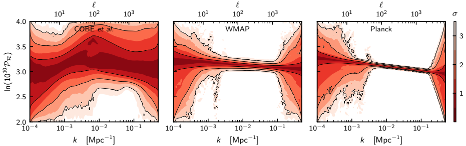
ERC grant: COSMOTENSION

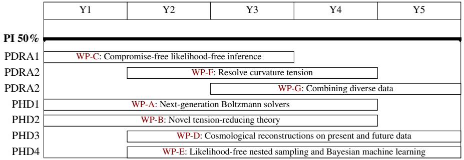
willhandley.co.uk/ERC.pdf

Resolving cosmological tensions with diverse data, novel theories and Bayesian machine learning

- ERC starting grant $\Rightarrow$ UKRI Frontier, commencing October 2024.
- Funds 3 PDRAs and 4 PhDs over 5 years.
- Research programme centered around combining novel theories of gravity, Boltzmann solvers [1906.01421], reconstruction [1908.00906], nested sampling & simulation-based inference (née LFI).
- Aims to disentangle cosmological tensions H_0, σ_8, Ω_K with next-generation data analysis techniques.







Conclusions

github.com/handley-lab

- The astronomy challenges of our generation are interdisciplinary.
- Interdisciplinary approaches are two-way streets for innovation.
- Simulation-based inference represents the frontier of how we will all be doing data analysis.
- Our research programme at the interface between theory, observation and inference is well-placed to explore the interdisciplinary frontier.