Next-generation astrophysical inference across the interdisciplinary frontier

Will Handley
<wh260@cam.ac.uk>

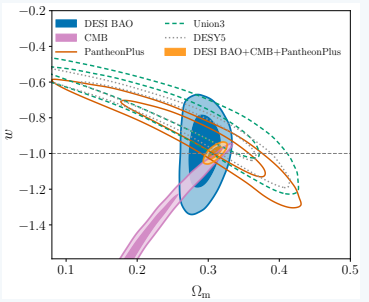
Royal Society University Research Fellow
Astrophysics Group, Cavendish Laboratory, University of Cambridge
Kavli Institute for Cosmology, Cambridge
Gonville & Caius College
willhandley.co.uk/talks

29th April 2024

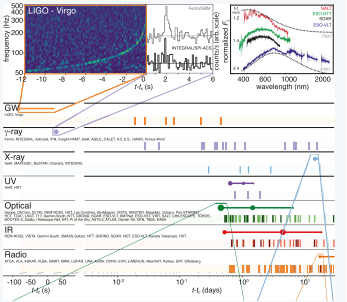

The future of astronomy is interdisciplinary

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

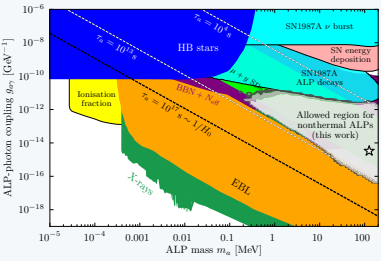
CMB+BAO



GW170817



HEP+Astro

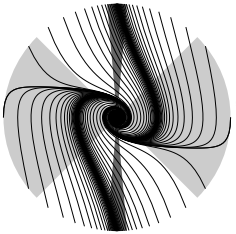


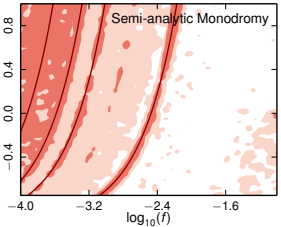
[2205.13549]

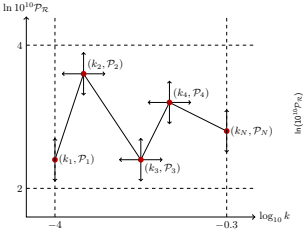
- We have spent the last 5 years hair-splitting "parkable" tensions.
- The next 25 years of data confront the real tensions in our understanding of the Universe.
- I aim to show how my research programme is preparing us for this interdisciplinary challenge.

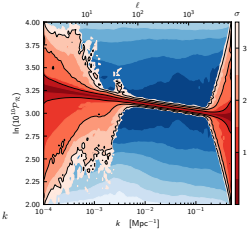
Planck: Inflation & primordial power spectrum

- Began theoretical PhD in 2012 investigating initial conditions for inflation.
- Joined Planck inflation team, working on Bayesian model fitting alongside theory.
- Found I enjoyed the observation & inference as much as the theory.
- `FlexKnots` were used to reconstruct the primordial power spectrum, inflationary potential & reionisation history (now used by Fialkov group) [1908.00906].
- `PolyChord` developed for model comparison, particularly axion monodromy.







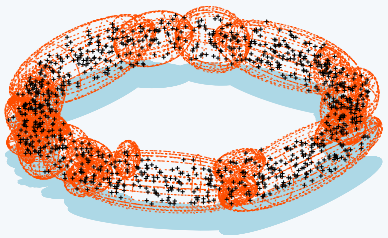


Analytic innovation: from MultiNest to PolyChord

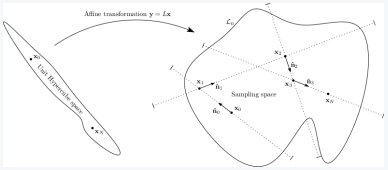
- MultiNest [0809.3437] was the leading Bayesian numerical model comparison tool in 2013.
  - A general purpose & performant implementation of John Skilling's nested sampling meta-algorithm.
  - Remains the leader of the pack in $n \sim \mathcal{O}(10)$ parameter fits.
  - Careful testing in *Planck* showed that it couldn't handle the many fast-slow nuisance parameters needed for systematics .
- I theoretically developed and numerically implemented PolyChord, which has polynomial scaling efficiency $f_{\mathtt{PC}} \sim \frac{1}{n}$ with model parameters (c.f. exponential $f_{\mathtt{MN}} \sim e^{-n/n_0}$).
- Exemplifies theoretical innovation & numerical implementation driven by astrophysical challenges.

PolyChord inspired a new generation of nested samplers (dynesty, UltraNest, nessai...), but remains the state of the art in high-dimensional model comparison.

MultiNest [0809.3437]



PolyChord [1506.00171]



Aside: theoretical work

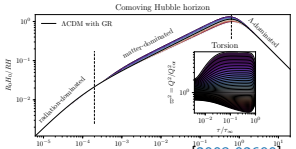
This talk will focus on my interdisciplinary work, but I have theoretical interests in:

- Quantum fields in curved spacetime
  (Mary Letey, Zak Shumaylov, Fruzsina Agocs)
- Poincaré Gauge Theory
  (Sinah Legner, Will Barker)
- Future conformal boundary/CPT universes
  (Metha Prathaban, Wei-Ning Deng)
- Curved finite inflation (Lukas Hergt)

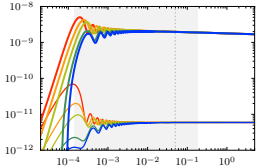
energy tensor (5). Inserting the mode function (26) into (32) and taking the coincidence limit, one finds:

$$\langle 0|T_{00}(x)|0\rangle_{\rm ren} = \frac{1}{2}\int \frac{d^3k}{(2\pi)^3 a^2}(\chi_{\mathbf{k}}' - \frac{a'}{a}\chi_{\mathbf{k}})(\chi_{\mathbf{k}}^{*\prime} - \frac{a'}{a}\chi_{\mathbf{k}}^*) + (k^2 + m^2a^2)\,\chi_{\mathbf{k}}\chi_{\mathbf{k}}^* + \tilde{T}, \qquad (34)$$

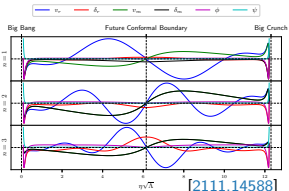
where $\tilde{T}$ signifies the plethora of additional terms arising from the renormalisation process that have no dependence on the variables $\mathcal{X}$. Minimising this with respect [1607.04148]



[2003.02690]



[2211.17248]

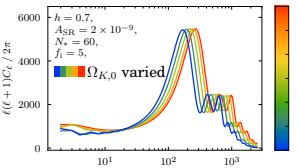


[2111.14588]

Substituting Eq. (24) into the above, and integrating by parts one more time returns the unusual action

$$\frac{1}{2}\int d^4x\sqrt{|c|}a^3\frac{\dot{\phi}^2}{H^2}\left\{\frac{1}{a^2}\mathcal{R}\mathcal{D}^2\mathcal{R} + \left(\mathcal{R} - \frac{K}{a^2}\frac{\mathcal{R}}{H}\right)\frac{\mathcal{D}^2}{\mathcal{D}^2 - K\mathcal{E}}\left(\mathcal{R} - \frac{K}{a^2}\frac{\mathcal{R}}{H}\right)\right\}. \qquad (26)$$

[1907.08524]



Interdisciplinary work to date

- CMB cosmology
- Cosmological tension quantification
- **21cm cosmology**
- Radio Instrumentation
- **Gravitational waves**
- **Exoplanets**
- **Particle physics**
- Theory of machine learning
- Nested sampling theory
- Atomistic chemistry
- **Industrial applications**
- ...

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

arxiv.org/a/handley_w_1.html

Will showcase a targeted subset.

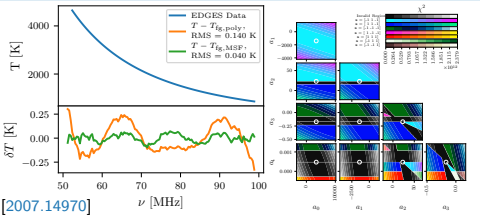
21cm cosmology

Harry Bevins
PhD→KICC fellow

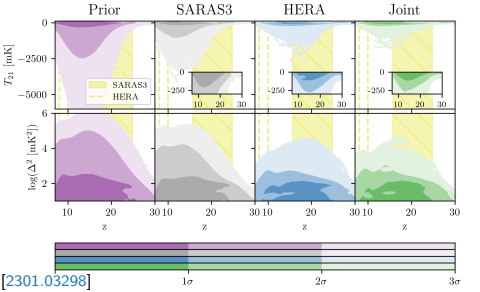
Collaboration with Anastasia Fialkov & Eloy de Lera Acedo transfers interdisciplinary ideas 21cm cosmology

- `maxsmooth` [2007.14970]
  - quadratic programming choice arose from quantitative finance consultancy work
- `FlexKnots`
  - importing ideas from inflationary reconstruction into reionisation [2310.05608](Heimersheim) & ionospheric reconstruction [2311.14537](Shen).
- `margarine` [2205.12841] [2207.11457]
  - combination of ideas from interdisciplinary fields (emulators, nested sampling, marginal density estimation)

These techniques are now widely used beyond the Cambridge 21cm community.



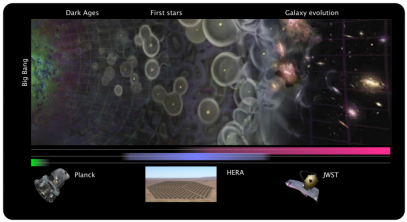
[2007.14970]

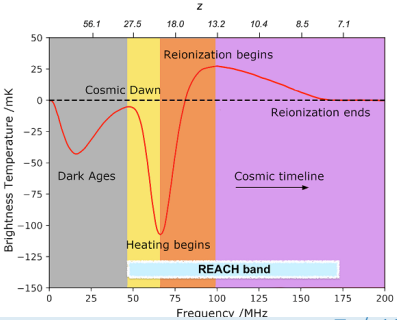


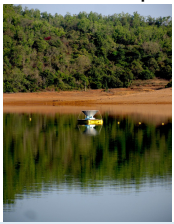
[2301.03298]

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.



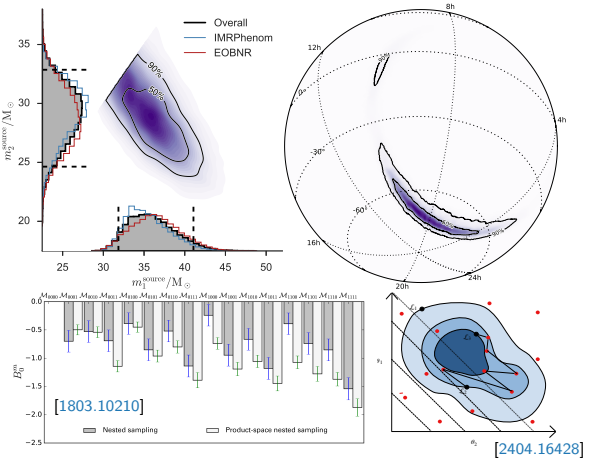




Gravitational waves

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



Exoplanets

Namu Kroupa
PhD

- Exoplanet science requires solution of subtle inference problems
  - Inference from RV data [1806.00518]
  - Survey challenges [2007.07278]
  - Stellar activity [2102.03387]
- Gaussian processes+Nested Sampling for transit astronomy [2311.04153]
- Potential for further collaboration with Madhu's group who are seeking beyond `MultiNest` solutions as their problems scale in dimensionality
- Ongoing cross-disciplinary theoretical chemistry work may be useful in Paul Rimmer's group.

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY
MNRAS **528,** 1232–1248 (2024) https://doi.org/10.1093/mnras/stae087
Advance Access publication 2024 January 12

Kernel-, mean-, and noise-marginalized Gaussian processes for exoplanet transits and H_0 inference

Namu Kroupa,[1,2,3★] David Yallup,[1,2] Will Handley [1,2] and Michael Hobson[1]

The HARPS search for southern extra-solar planets – XLV. Two Neptune mass planets orbiting HD 13808: a study of stellar activity modelling's impact on planet detection

E. Ahrer [1,2★] D. Queloz,[1,3] V. M. Rajpaul [1] D. Ségransan,[3] F. Bouchy,[3] R. Hall [1] W. Handley [1,4] C. Lovis,[3] M. Mayor,[3] A. Mortier [1,4] F. Pepe,[3] S. Thompson,[1] S. Udry[3] and N. Unger [3]

Global analysis of the TRAPPIST Ultra-Cool Dwarf Transit Survey

F. Lienhard [1★] D. Queloz,[1] M. Gillon,[2] A. Burdanov [3,4] L. Delrez [1,2,5] E. Ducrot,[2] W. Handley,[1] E. Jehin,[5] C. A. Murray [1] A. H. M. J. Triaud [6] E. Gillen [1‡] A. Mortier [1] and B. V. Rackham [3,7†]

On the Feasibility of Intense Radial Velocity Surveys for Earth-Twin Discoveries

Richard D. Hall,[1★] Samantha J. Thompson,[1] Will Handley[1,2] and Didier Queloz[1,3]

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

GAMBIT: sub-GeV Dark matter constraints

Interdisciplinary case studies

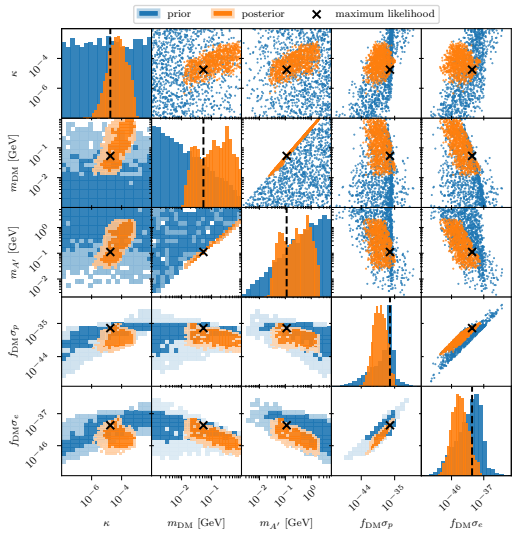
Felix Kahlhoefer et al
GAMBIT cosmo/DM working group

- 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).

$$\begin{aligned}\mathcal{L}_\psi =& \bar{\psi}(i\partial\!\!\!/ - m_{\text{DM}})\psi + g_{\text{DM}} A'^{\mu} \bar{\psi} \gamma_\mu \psi \,, \\ \mathcal{L}_\Phi =& |\partial_\mu \Phi|^2 - m_{\text{DM}}^2 |\Phi|^2 - g_{\text{DM}}^2 A'_\mu A'^{\mu} |\Phi|^2 \\ &+ i g_{\text{DM}} A'^{\mu} \left[\Phi^* (\partial_\mu \Phi) - (\partial_\mu \Phi^*) \Phi\right] \,,\end{aligned}$$

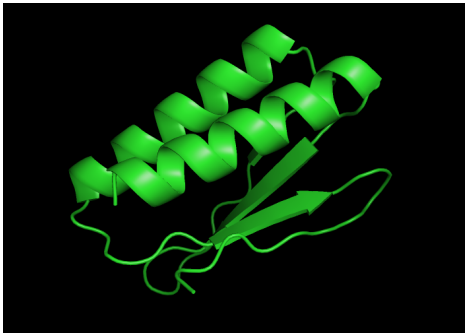


PolyChord Ltd: interdisciplinary R&D spinout

Catherine Watkinson
Senior Data Scientist

Interdisciplinary case studies

- 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.

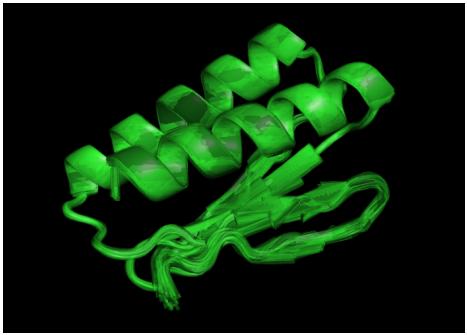


PolyChord Ltd: interdisciplinary R&D spinout

Catherine Watkinson
Senior Data Scientist

Interdisciplinary case studies

- 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.

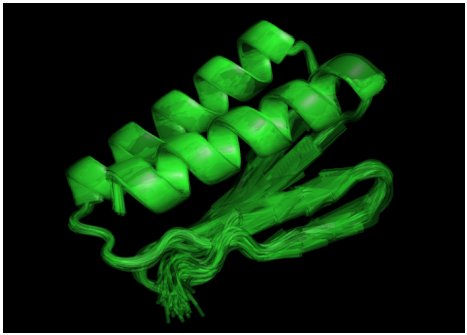


PolyChord Ltd: interdisciplinary R&D spinout

Catherine Watkinson
Senior Data Scientist

Interdisciplinary case studies

- 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.

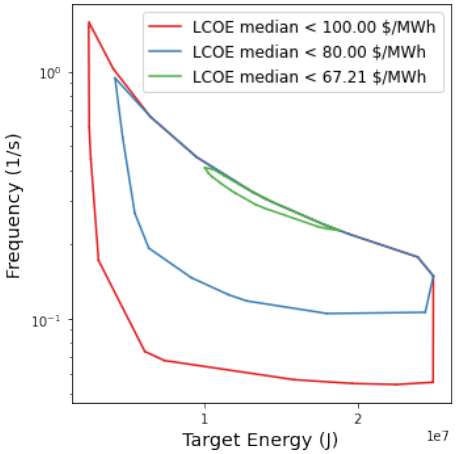


PolyChord Ltd: interdisciplinary R&D spinout

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.

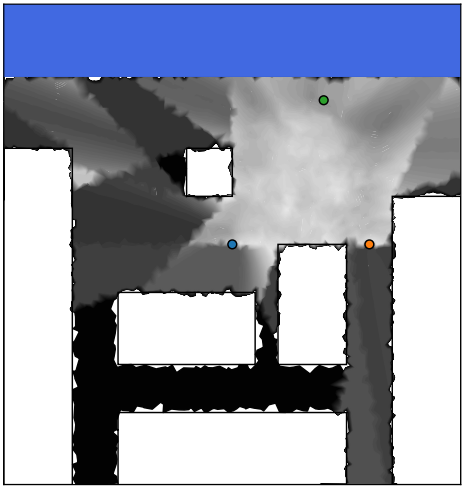


PolyChord Ltd: interdisciplinary R&D spinout

Interdisciplinary case studies

Thomas Mcaloone
PhD → Data Scientist

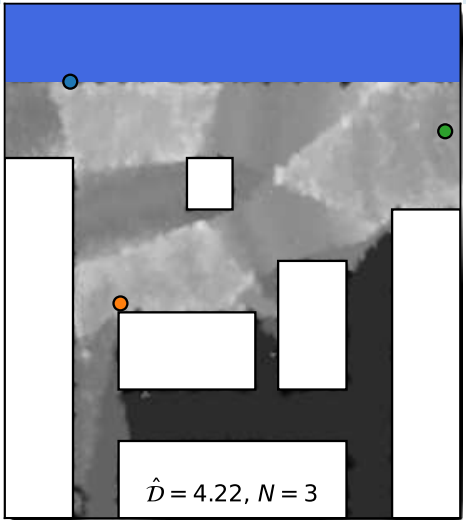
- Techniques have been spun-out (PolyChord Ltd) to:
- Protein folding
  - Navigating free energy surface.
  - Computing misfolds.
  - Thermal motion.
- Nuclear fusion reactor optimisation
  - multi-objective.
  - uncertainty propagation.
- Telecoms & DSTL research (MIDAS)
  - Optimising placement of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



PolyChord Ltd: interdisciplinary R&D spinout

Interdisciplinary case studies

- 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.

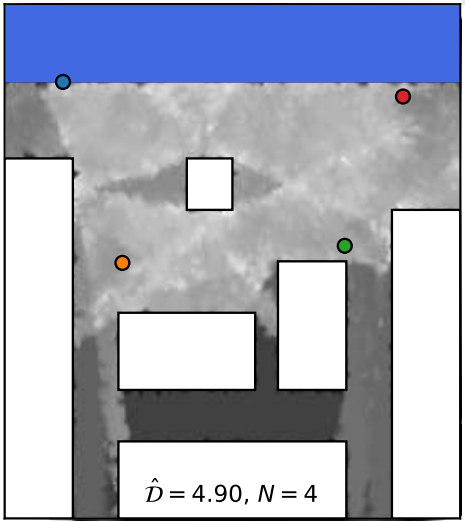


PolyChord Ltd: interdisciplinary R&D spinout

Thomas Mcaloone
PhD → Data Scientist

Interdisciplinary case studies

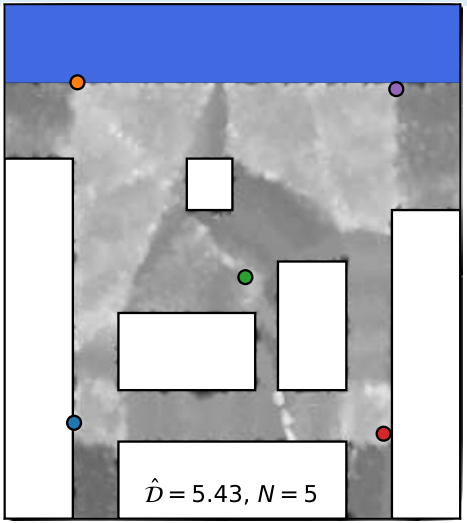
- Techniques have been spun-out (PolyChord Ltd) to:
- Protein folding
  - Navigating free energy surface.
  - Computing misfolds.
  - Thermal motion.
- Nuclear fusion reactor optimisation
  - multi-objective.
  - uncertainty propagation.
- Telecoms & DSTL research (MIDAS)
  - Optimising placement of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



PolyChord Ltd: interdisciplinary R&D spinout

Interdisciplinary case studies

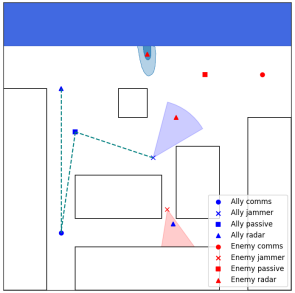
- Techniques have been spun-out (PolyChord Ltd) to:
- Protein folding
  - Navigating free energy surface.
  - Computing misfolds.
  - Thermal motion.
- Nuclear fusion reactor optimisation
  - multi-objective.
  - uncertainty propagation.
- Telecoms & DSTL research (MIDAS)
  - Optimising placement of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



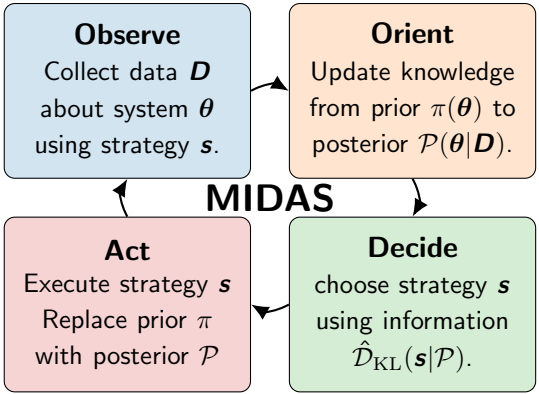
DSTL: Bayesian OODA loops

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



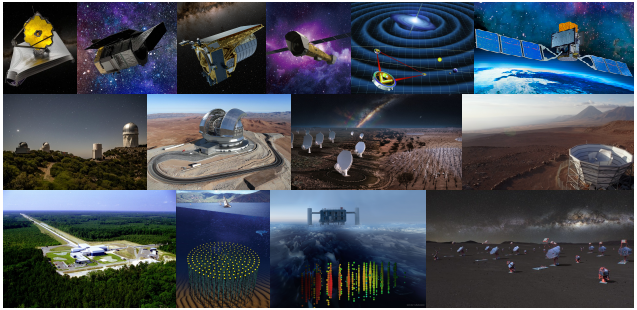
- Two-way research interaction between government and academia.
- techniques now being used in REACH antenna design [2309.06942]



Beginning the golden age of astronomy data

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

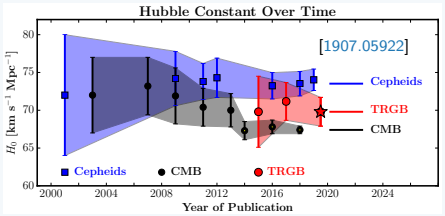
Radio	SKA *et al*
Micro	SO/CMB-S4
IR	JWST, Roman (WFIRST)
Optical	Euclid, DESI, Rubin (LSST), EELT
X-ray	Athena
Gamma-ray	e-ASTROGAM
Gravitational	LIGO/Virgo/Kagra + LISA
Particle	CTA, IceCube, KM3NeT



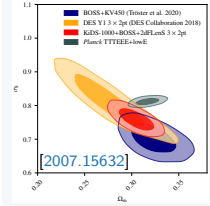
- We are moving from an age of **precision** cosmology to **accurate** cosmology.
- **Systematics** $\gtrsim$ **statistics**.
- Tools risk lagging behind hardware

Tensions in cosmology

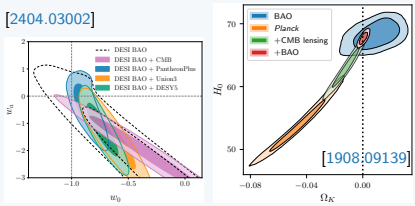
Hubble



Weak lensing



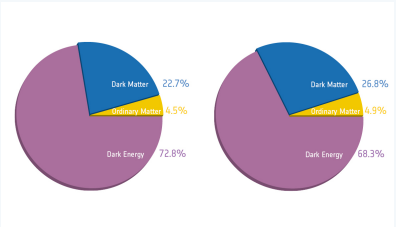
other $w_0/\Omega_K/\nu$?



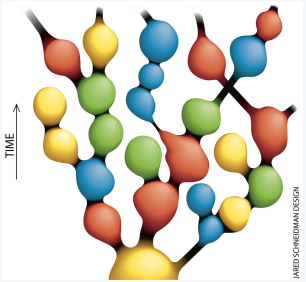
- Tensions of debatable significance have been plaguing cosmology for the past five years
- Independent of veracity, they have revealed a gap in the armour of modern data analysis.
- Likelihood-based methods have assume a fiducial cosmology.
- This will cast greater doubt on the robustness of our conclusions from stage 4 surveys.

The real tensions in the room

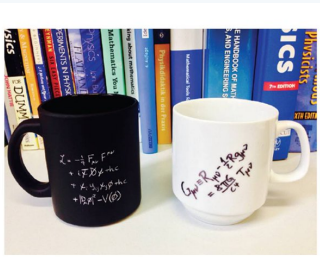
Dark tension



Initial conditions



Quantum gravity



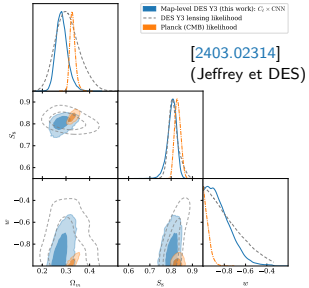
- These are existential questions that don't go away if ΛCDM is phenomenologically correct...
- ...and if it is, we will need an interdisciplinary approach that goes far beyond cosmology (using astronomy, collider & experimental physics).
- The GW redshift frontier, time domain astronomy & 21cm SKA will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

The future: simulation-based inference

Kilian Scheutwinkel
PhD

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

- Can extract nonlinear information from data
- Users do not need to know advanced statistics
  - For this reason alone, it will come to dominate the next generation, so it is imperative to build a principled understanding!
- The current state of the art is achieved with *Neural* estimation
- 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` [2403.02314] (presenting at EuCAIFCon)



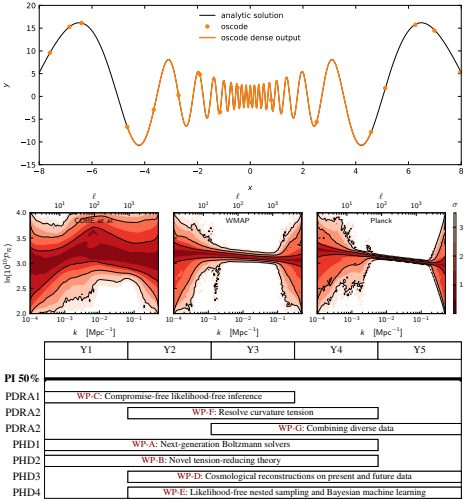
[2403.02314]
(Jeffrey et DES)

ERC grant: COSMOTENSION

willhandley.co.uk/ERC.pdf

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

- ERC starting grant ⇒ 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.

DiRAC erc UKRI THE ROYAL SOCIETY PolyChord