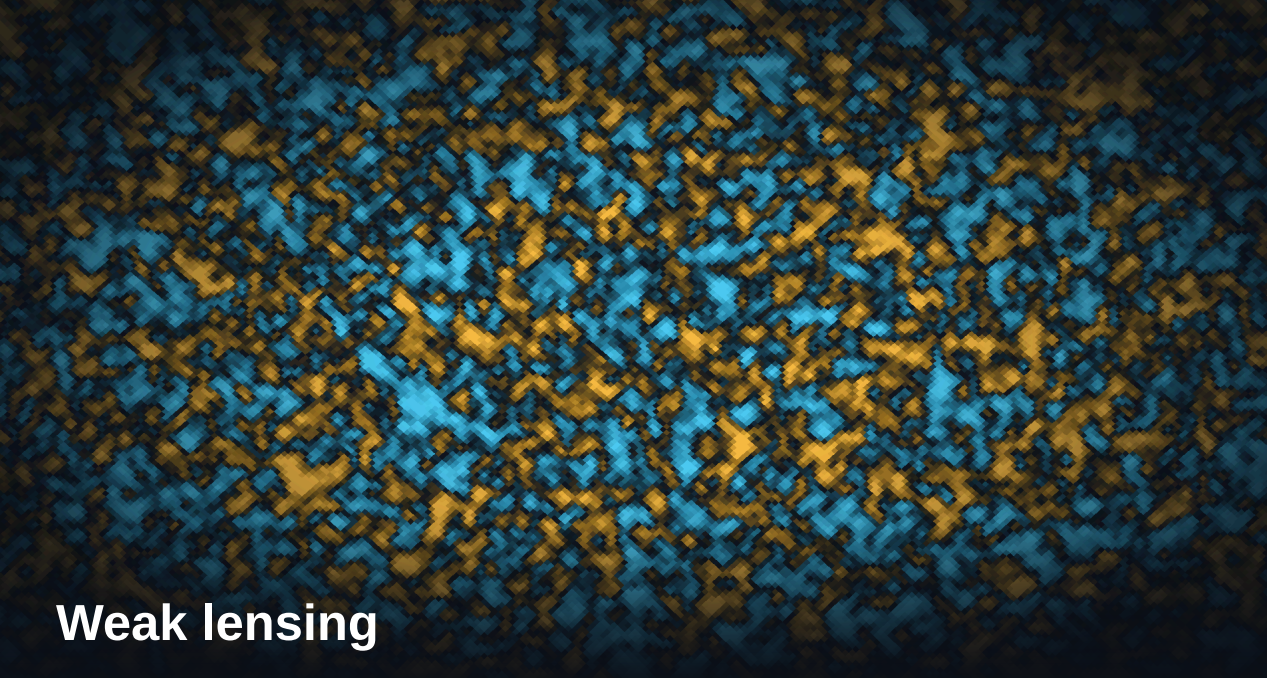




Weak lensing



Radio & 21-cm cosmology

From Stochastic Physics to Cosmological Inference

Yale University · 20 Oct

Zheng Zhang

Jodrell Bank Centre for Astrophysics
The University of Manchester

With support from



Funded by
the European Union

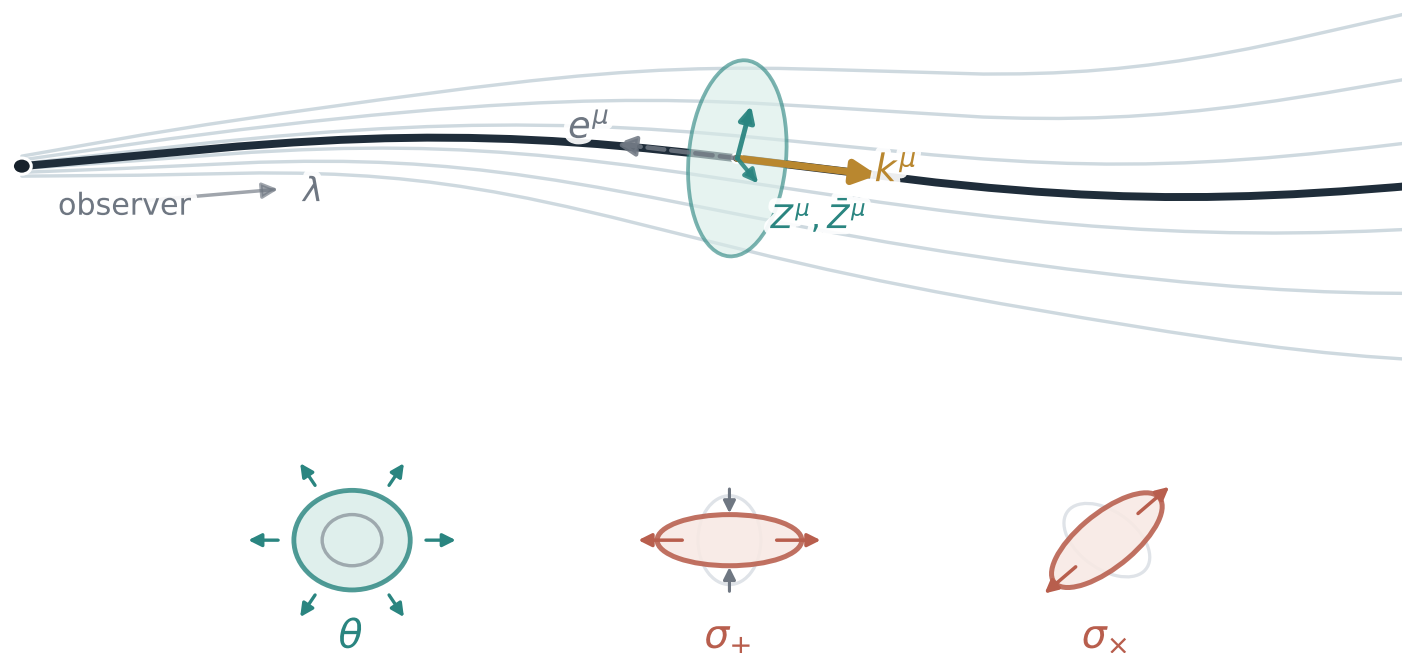


European Research Council
Established by the European Commission



ERC grant agreement No. 948764 · RadioForegroundsPlus GA 101135036

A decade of weak lensing is arriving

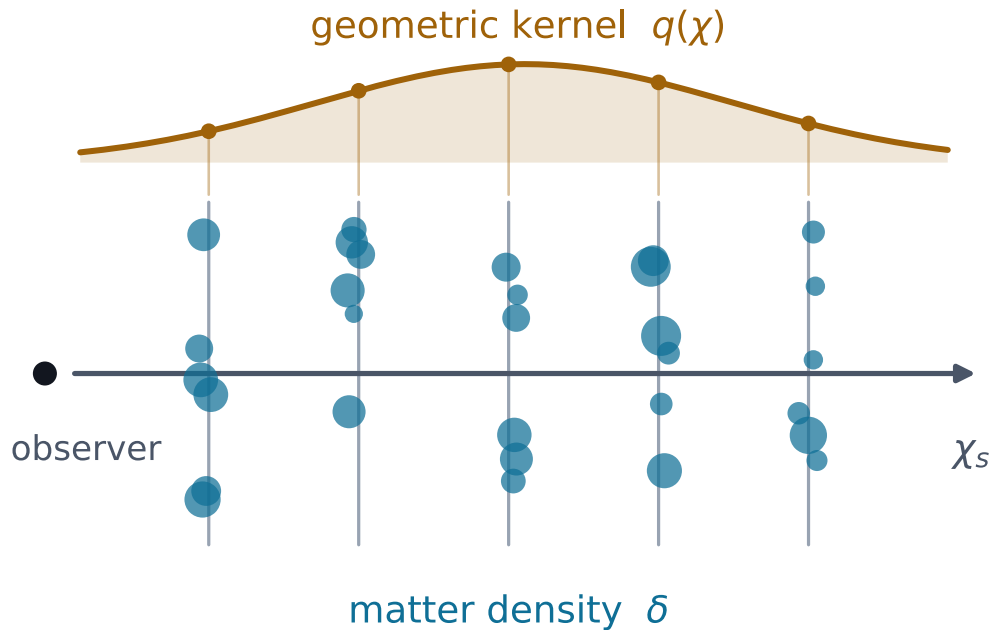


κ : convergence · magnification

γ : shear · shape distortion

Target: per-cent precision

Everything rests on a linear hierarchy



$$\kappa(\hat{n}) = \int_0^{\chi_s} d\chi q(\chi) \delta(\chi \hat{n}, \chi)$$

$$\xi_\kappa \leftarrow P_m(k)$$

$$\zeta_\kappa \leftarrow B_m(k_1, k_2, k_3)$$

Nothing mixes.

But the projection is an approximation

Known corrections, each in its own formalism:

post-Born & lens–lens coupling

beyond-Limber radial mixing

reduced shear & magnification bias

nonlinear growth of the matter field

What does the projection throw away?

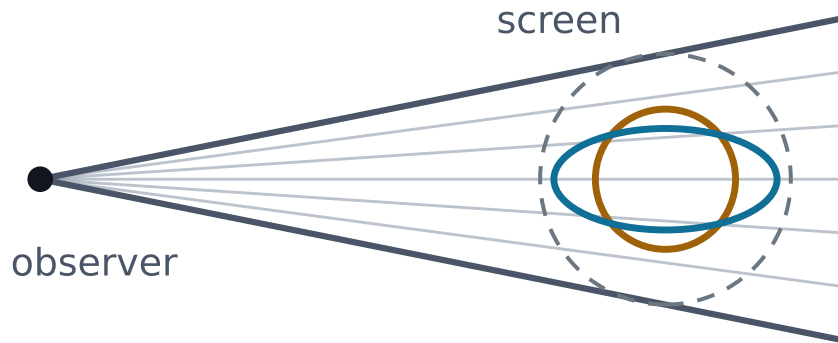
Does a two-point measurement still read only the matter two-point function?

A photon bundle **focuses** and **shears**

$$\dot{\theta} = -\theta^2 - \sigma_+^2 - \sigma_\times^2 + \Phi_{00}$$

$$\dot{\sigma}_{+,\times} = -2\theta\sigma_{+,\times} + \mathcal{W}_{1,2}$$

Sachs optical scalars



θ expansion

$\sigma_+, \sigma_\times$ shear

what drives each rate

Φ_{00}  $\rightarrow \dot{\theta} \rightarrow \kappa$
Ricci, parity even

$\text{Re } \Psi_0$  $\rightarrow \dot{\sigma}_+ \rightarrow \gamma_+$
Weyl, parity even

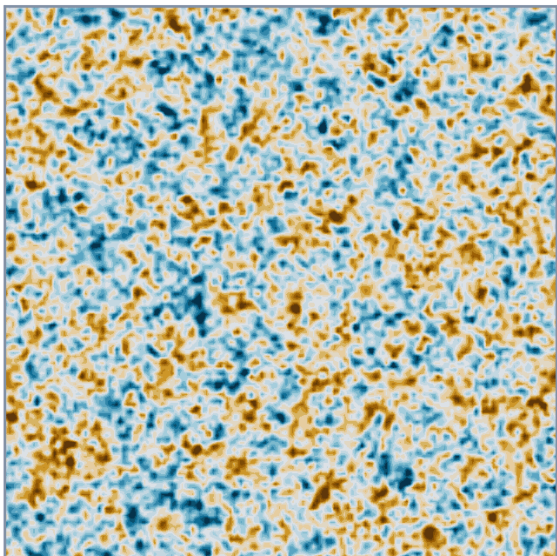
$\text{Im } \Psi_0$  $\rightarrow \dot{\sigma}_\times \rightarrow \gamma_\times$
Weyl, parity odd

The spacetime is a **random field**

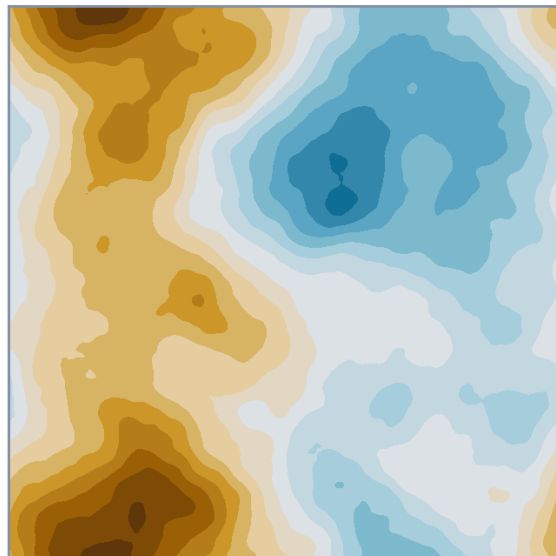
one realization of the driving fields

line-of-sight depth = 150 Mpc

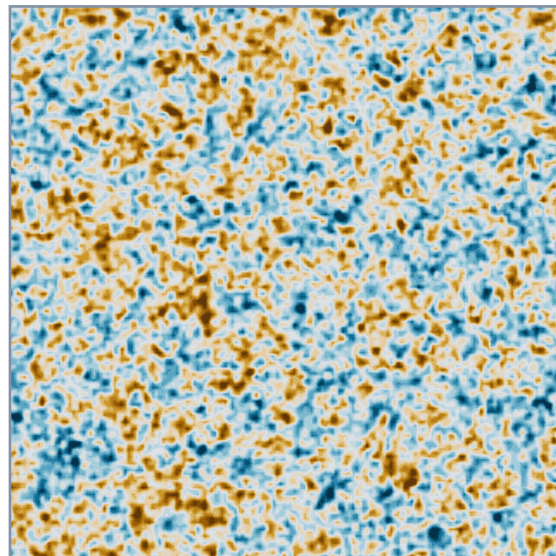
δ_m overdensity



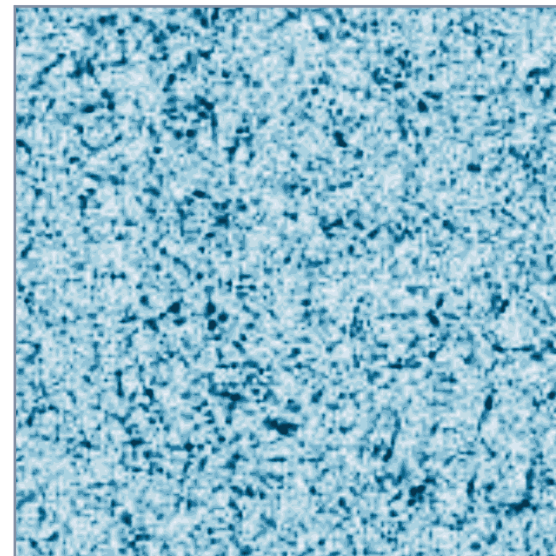
Φ potential



Φ_{00} Ricci $\rightarrow$ focusing



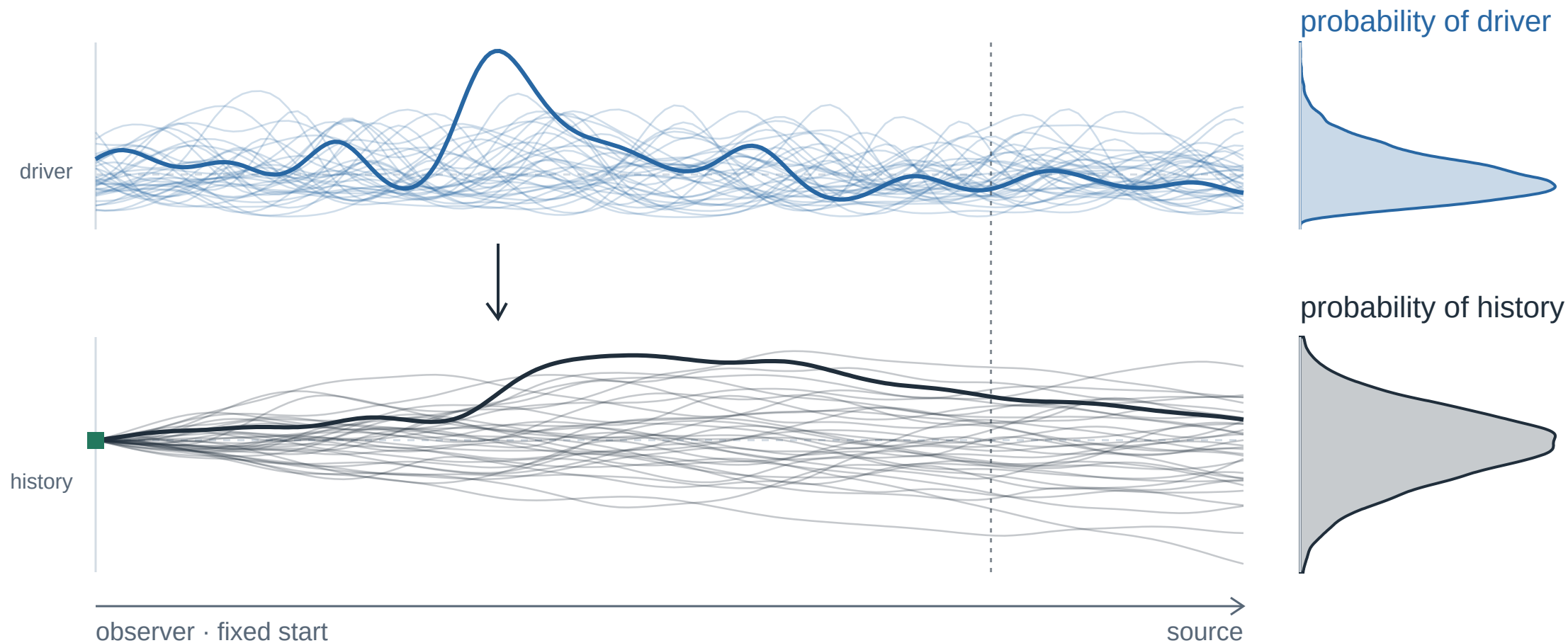
$|\Psi_0|$ Weyl $\rightarrow$ shear



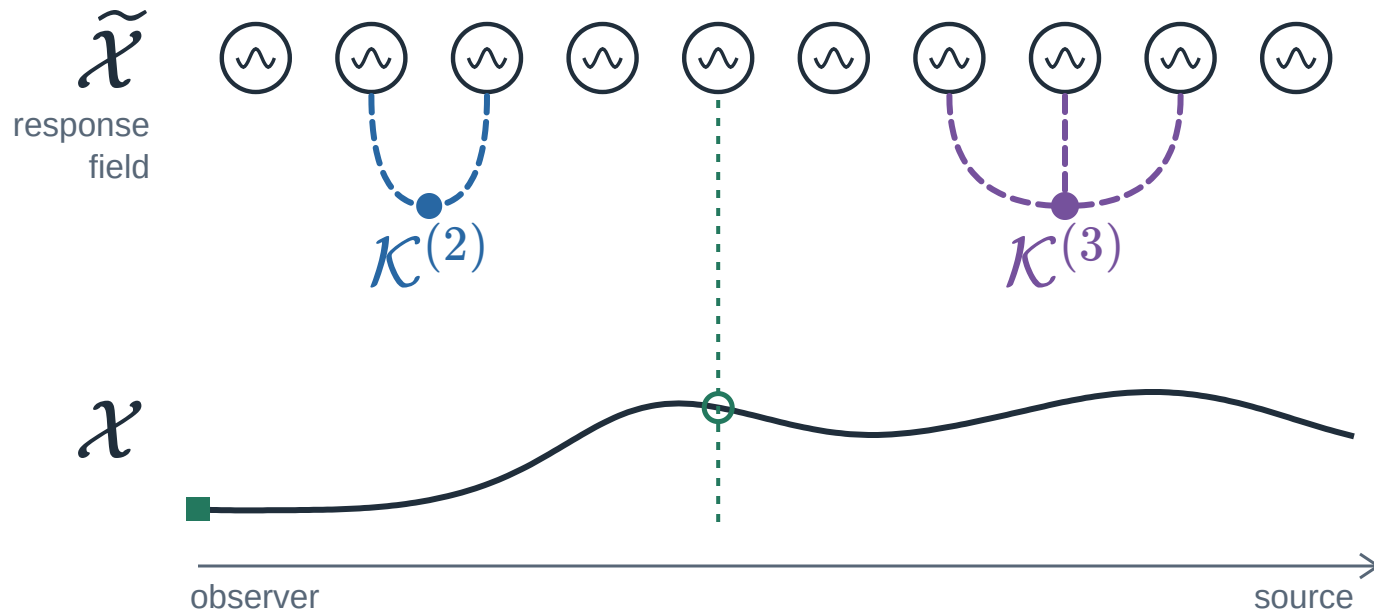
$$\dot{\chi}_a = \underbrace{A_{ab}\chi_b}_{\text{drift}} + \underbrace{F_{abc}\chi_b\chi_c}_{\text{Sachs vertex}} + \underbrace{\varphi_a}_{\text{driving field}}$$

$$\mathcal{X} = (\delta\theta, \sigma_+, \sigma_\times), \quad \varphi = (\Phi_{00}, \mathcal{W}_1, \mathcal{W}_2)$$

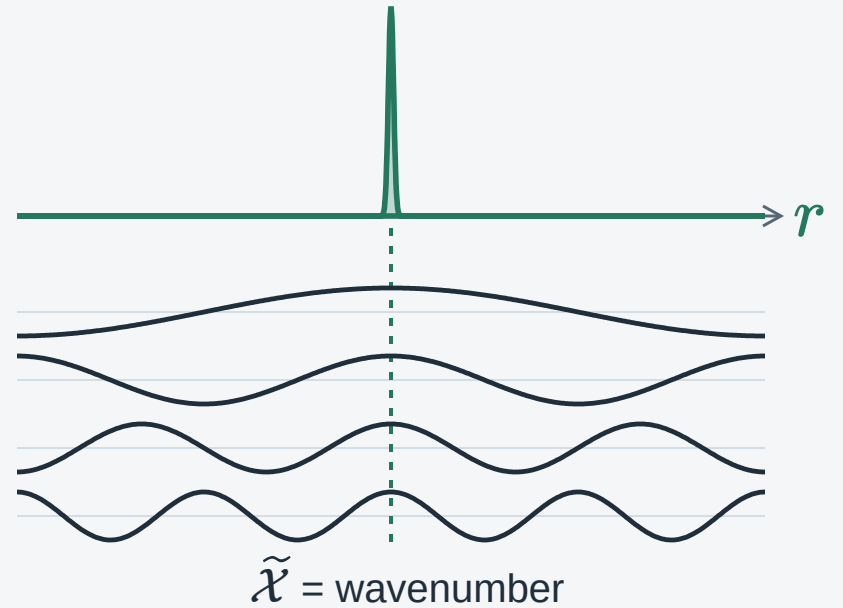
Random curvature in, random optical histories out



Averaging out the driver leaves only its cumulants



one point



Dynamics + driver statistics become an action.

$$\mathcal{P}[\chi] = \int \mathcal{D}\tilde{\chi} e^{-S[\chi, \tilde{\chi}]}$$

Expand around a Gaussian reference; pair the fields.

$$S[\chi, \tilde{\chi}] = \underbrace{i \int \tilde{\chi} (\dot{\chi} - A\chi - F\chi\chi)}_{\text{equation of motion}} - \underbrace{W_\varphi[i\tilde{\chi}]}_{\text{driver statistics}}$$

$$\langle \mathcal{O} \rangle = \int \mathcal{D}\chi \mathcal{D}\tilde{\chi} \mathcal{O}[\chi] e^{-S} \quad \langle 1 \rangle = 1$$

$$S = S_0 + S_{\text{int}}$$

linear dynamics + covariance

Sachs vertex + higher cumulants

$$e^{-S} = e^{-S_0} \left(1 - S_{\text{int}} + \frac{1}{2} S_{\text{int}}^2 - \dots \right)$$

Wick's theorem

Gaussian averages → allowed pairings → diagrams

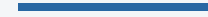


Building blocks: two propagators, two kinds of vertices



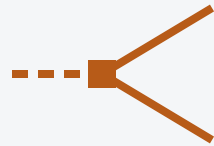
R

causal response: how a kick at λ' moves the bundle at $\lambda > \lambda'$



$$C = \iint R \mathcal{K}^{(2)} R$$

the driver covariance, propagated to both ends



F

the Sachs vertex: nonlinear propagation, present even in a Gaussian universe



$$K_n = \mathcal{K}^{(n)}, \quad n \geq 3$$

the driver's non-Gaussianity: a connected n -point cumulant, nonlocal

sft-wick: Wick contractions and diagrams by machine

sft-wick

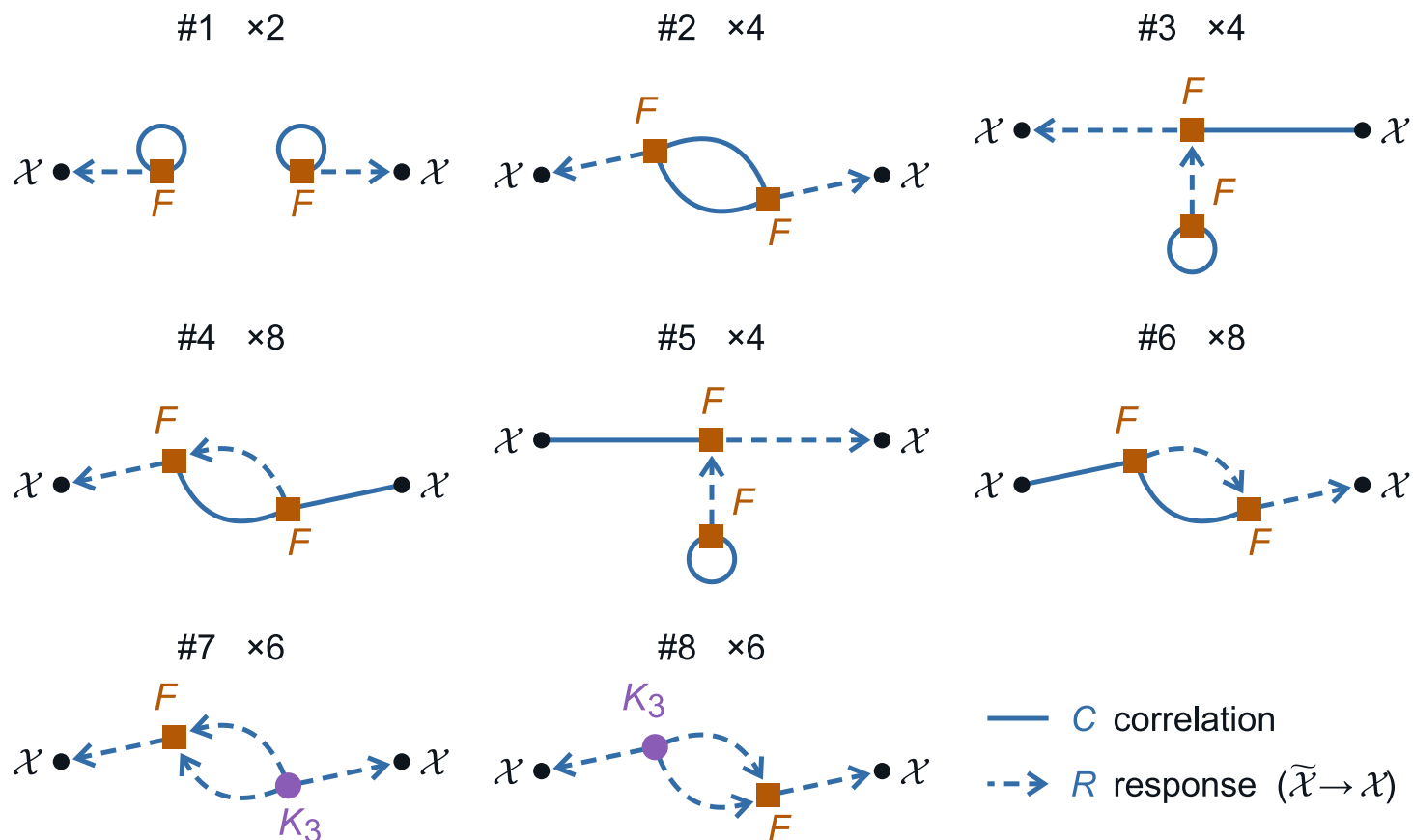
github.com/StatFieldTheory/sft-wick

`pip install sft-wick`

[arXiv:2606.19480](https://arxiv.org/abs/2606.19480)

EXAMPLE two-point function $\langle \mathcal{X}(\lambda_1) \mathcal{X}(\lambda_2) \rangle$ at second order

42 pairings $\rightarrow$ 8 diagrams



DEMONSTRATION

Does sft-wick work?

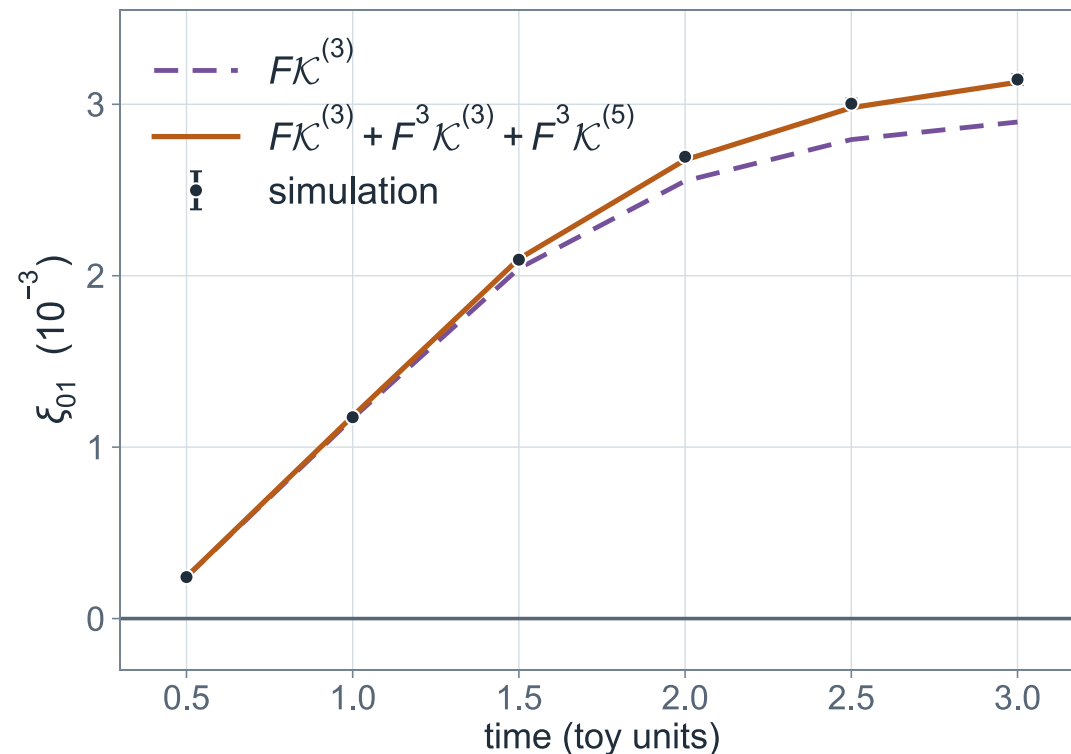
$$\dot{\chi}_0 = -\Gamma \chi_0 + s \chi_1^2 + \eta_0$$

$$\dot{\chi}_1 = -\Gamma \chi_1 + s \chi_0 \chi_1 + \eta_1$$

$$\Gamma = 1, \quad s = 0.2$$

$$\xi_{01} \equiv \langle \chi_0 \chi_1 \rangle = \underbrace{F\mathcal{K}^{(3)} + F^3\mathcal{K}^{(3)}}_{\propto \text{skewness } +0.62} + F^3\mathcal{K}^{(5)} + \dots$$

Gaussian drive: $\xi_{01} = 0$ by symmetry



≤ 0.90 σ theory vs simulation, all six times

A weak-lensing toy model

Matter fluctuations drive the Sachs equations along unperturbed light rays.

DRIVING STATISTICS

$$P_{\text{lin}} \longrightarrow \mathcal{K}^{(2)}$$

$$B_{\text{tree}} \longrightarrow \mathcal{K}^{(3)}$$

Power spectrum and bispectrum of matter

LENSING OBSERVABLES

$$\kappa = - \int_0^{\lambda_s} \delta\theta \, d\lambda$$

$$\gamma_{+, \times} = - \int_0^{\lambda_s} \sigma_{+, \times} \, d\lambda$$

Convergence and shear from the Sachs scalars

Born rays · linearised observable map · linear / tree-level driving statistics

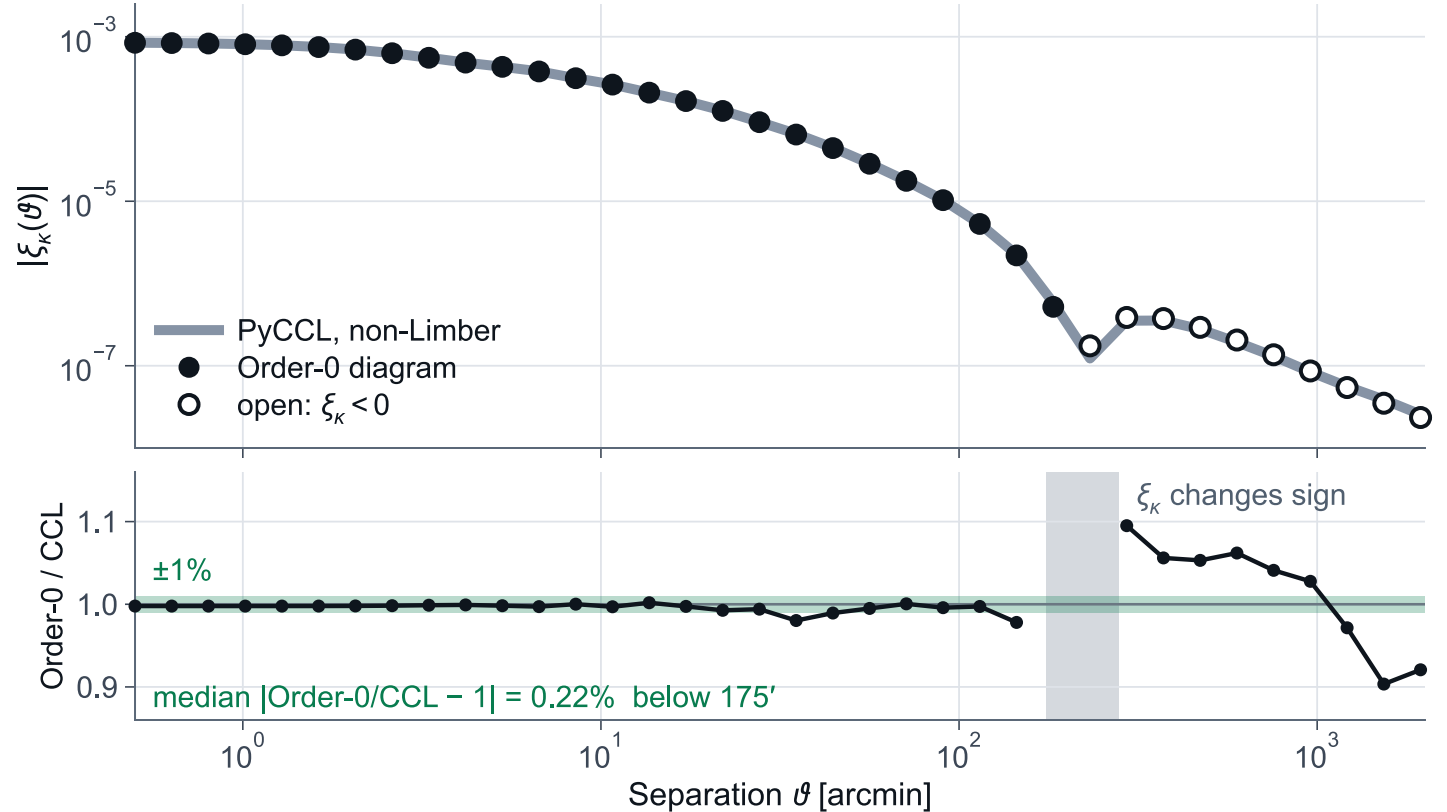
The textbook kernel is the leading diagram

Order 0: one C line, no vertex

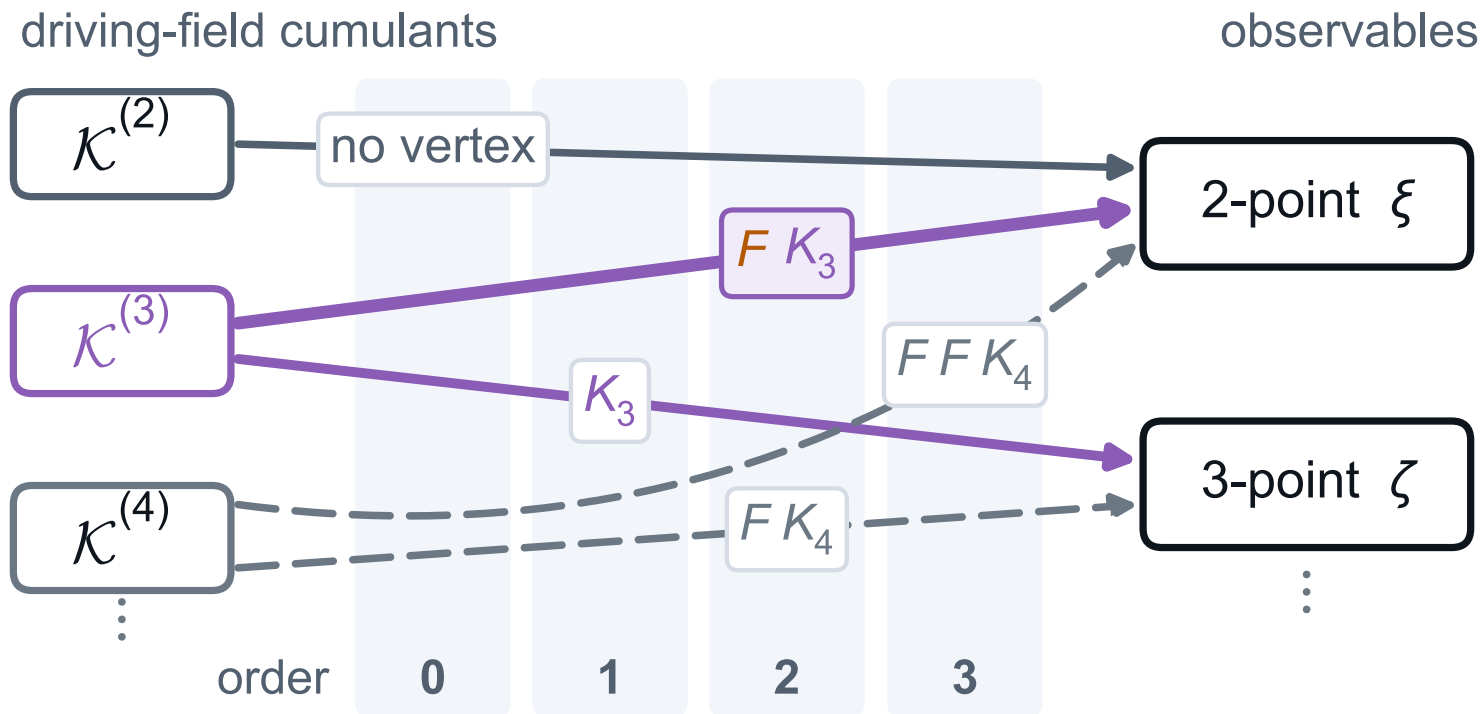
$$\langle \kappa \kappa \rangle_0 = \iint \mathcal{W} \mathcal{W} \kappa^{(2)}$$

$$\mathcal{W}(\lambda, \lambda_s) = \bar{D}^2 \int_{\lambda}^{\lambda_s} \frac{d\lambda_1}{\bar{D}^2(\lambda_1)}$$

$$= a^2 \underbrace{\frac{\chi(\chi_s - \chi)}{\chi_s}}_{\text{textbook lensing efficiency}}$$



Which cumulant reaches which observable, and at what order?



Hierarchy

$\mathcal{K}^{(n+j)} \rightarrow n\text{-point: } j \text{ vertices } F$

Selection rule

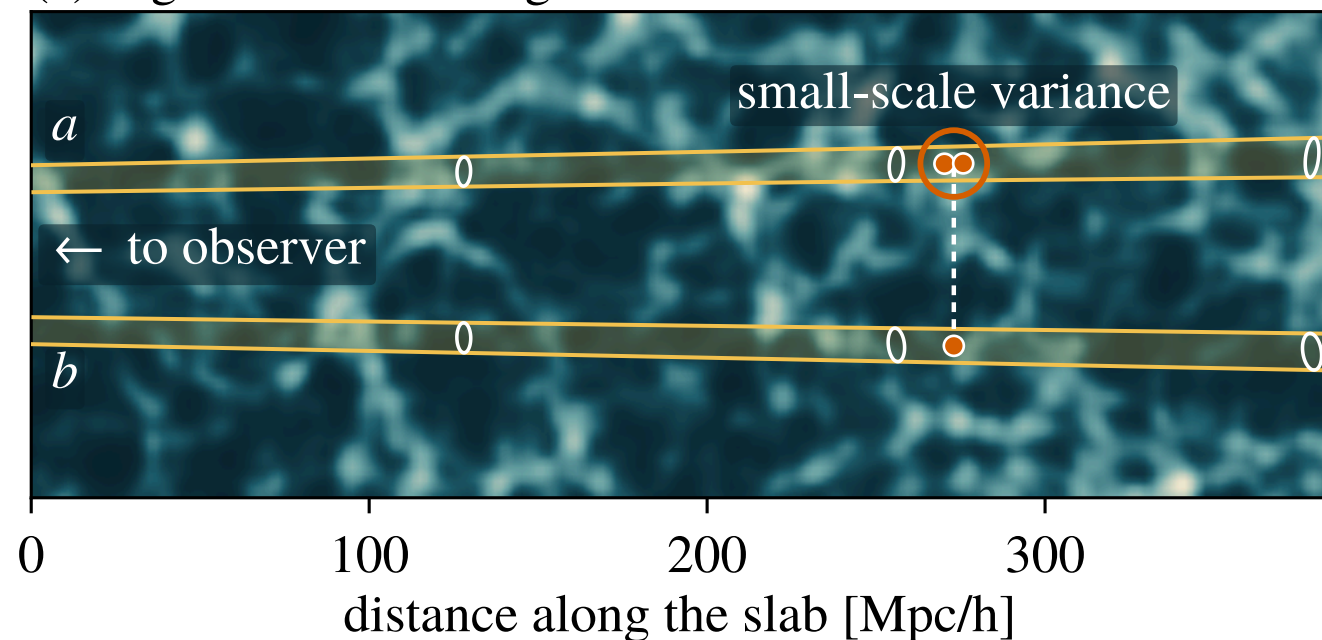
first mixing: $F \mathcal{K}^{(n+1)}$ only

For ξ : $\mathcal{K}^{(3)}$, the bispectrum

At the next order exactly two channels survive

FF · nonlinear propagation

(a) Light bundles through a simulated cosmic web



Hubble proposal, Fig. 1a · gevolution slab
Distortions $\times 200$ · bundle separation not to scale

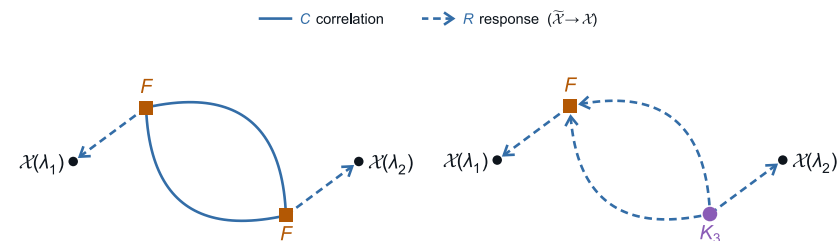
FK · hierarchy mixing

The bispectrum inside a two-point signal

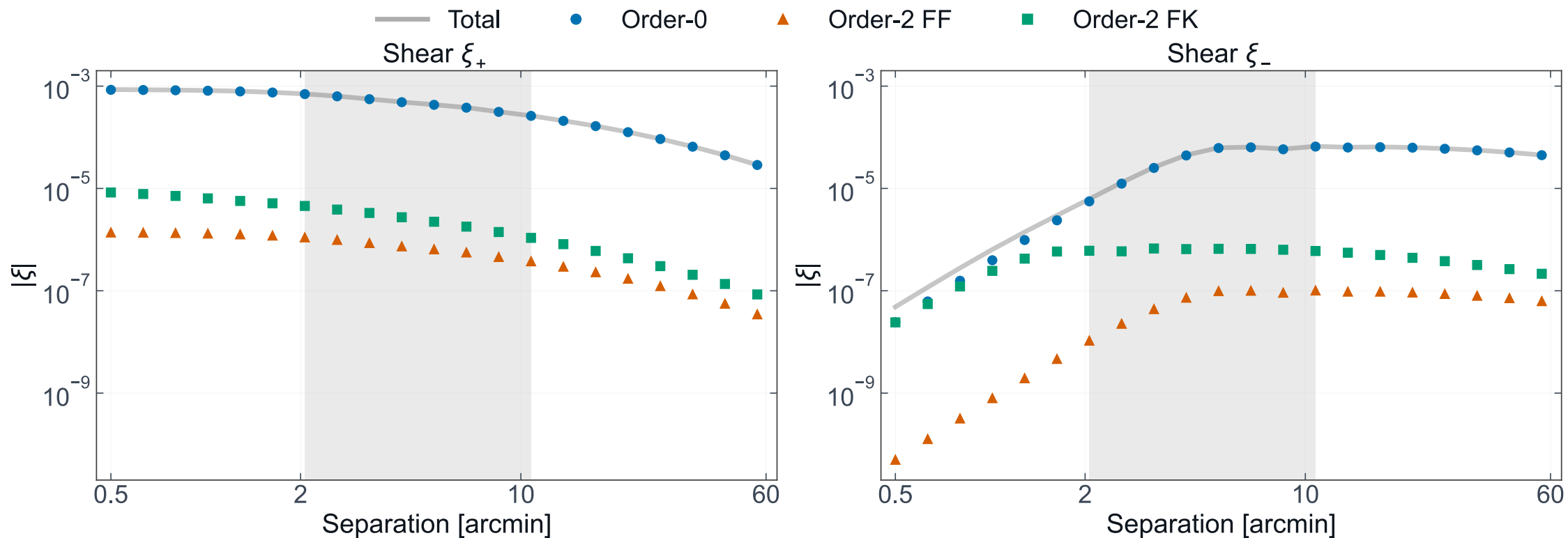
Two close points on a :
small-scale variance

Correlated with a point on b :
its environment

The FK channel carries this three-point information into the lensing two-point function.



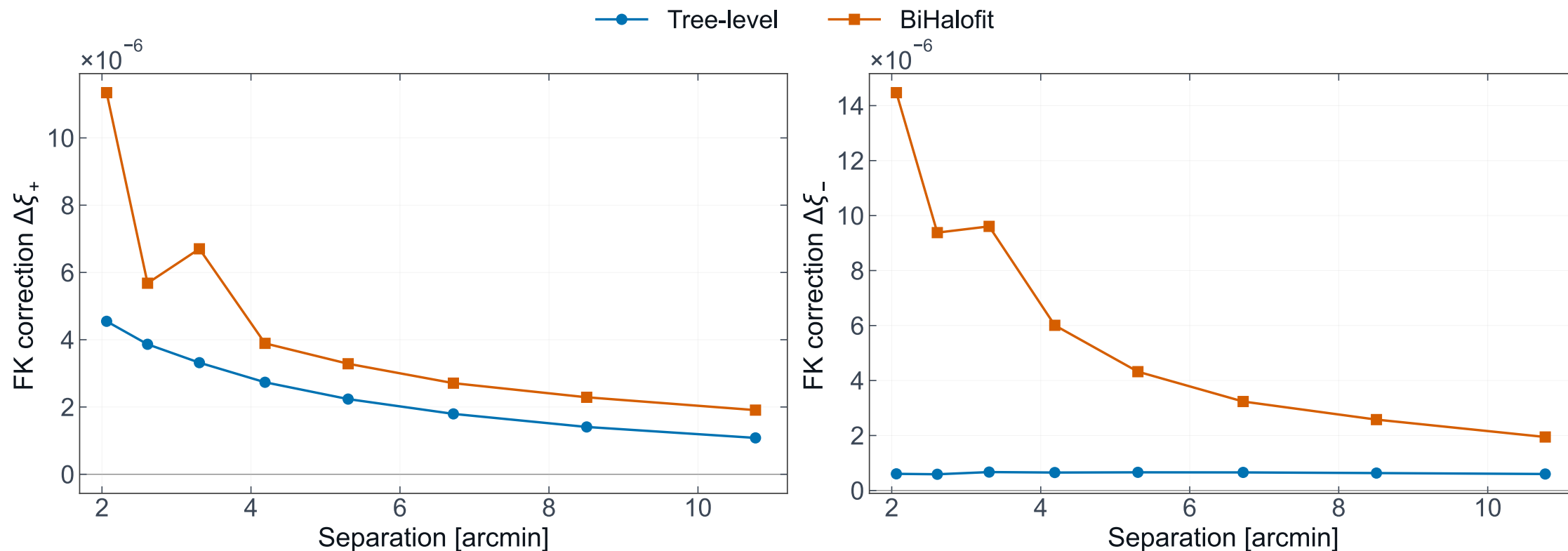
The FK contribution is present in both shear correlations



$$\xi_{\pm} = \langle \gamma_t \gamma_t \rangle \pm \langle \gamma_{\times} \gamma_{\times} \rangle \cdot \text{shaded: 2–11'} \cdot \text{filled / open: positive / negative}$$

$z_s = 5$, linear P , tree-level B , $\ell_{\max} = 15360$ · linearized shear readout

The shear correction depends on the input bispectrum



Fixed $P_{\text{lin}} \cdot z_s = 5$ · finite $\ell_{\text{max}} = 15360$

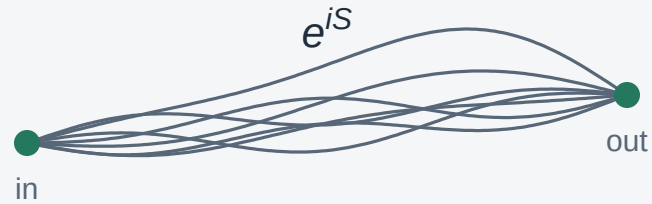
SFT-WL

Needed for precision weak lensing

Probes non-Gaussianity with the two-point function

Where does randomness enter the path integral?

Quantum field theory



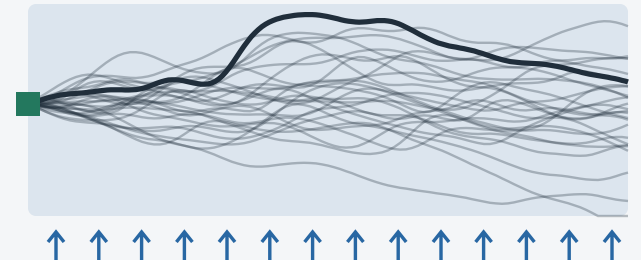
fixed in, fixed out
→ amplitude $\langle \text{out} | \text{in} \rangle$

Large-scale structure



random start, closed dynamics
→ average over initial fields

SFT-WL · this talk



fixed start, random kicks all along
→ probability of each history

SFT-WL averages optical histories driven by random curvature all along the ray.

RHINO

The Remote HI eNvironment Observer

An independent approach to the
sky-averaged 21-cm signal

Horn antenna

Directive, shielded, chromatic

Hybrid continuous-wave calibration

Less switching, different systematics







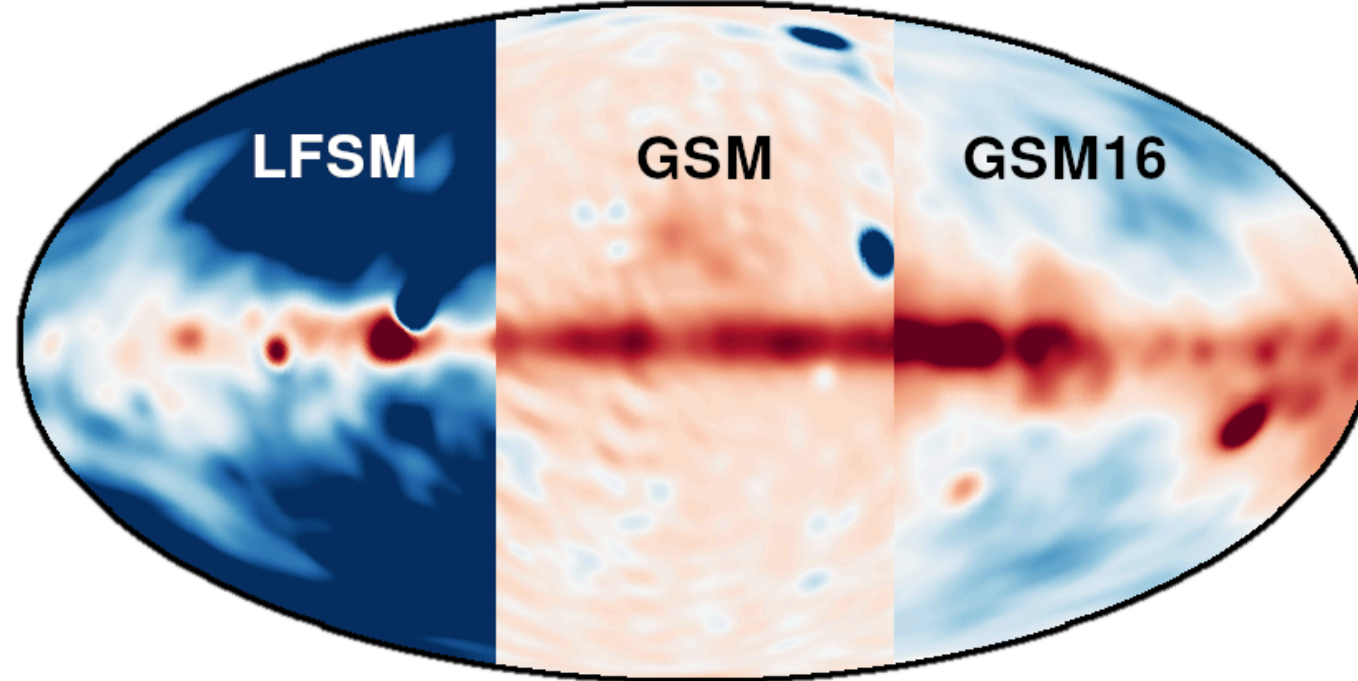
Faint cosmological signals share a foreground problem

Global 21-cm
mean intensity

Radio excess
absolute intensity

CMB B-modes
polarisation

Spectral distortions
CMB intensity spectrum



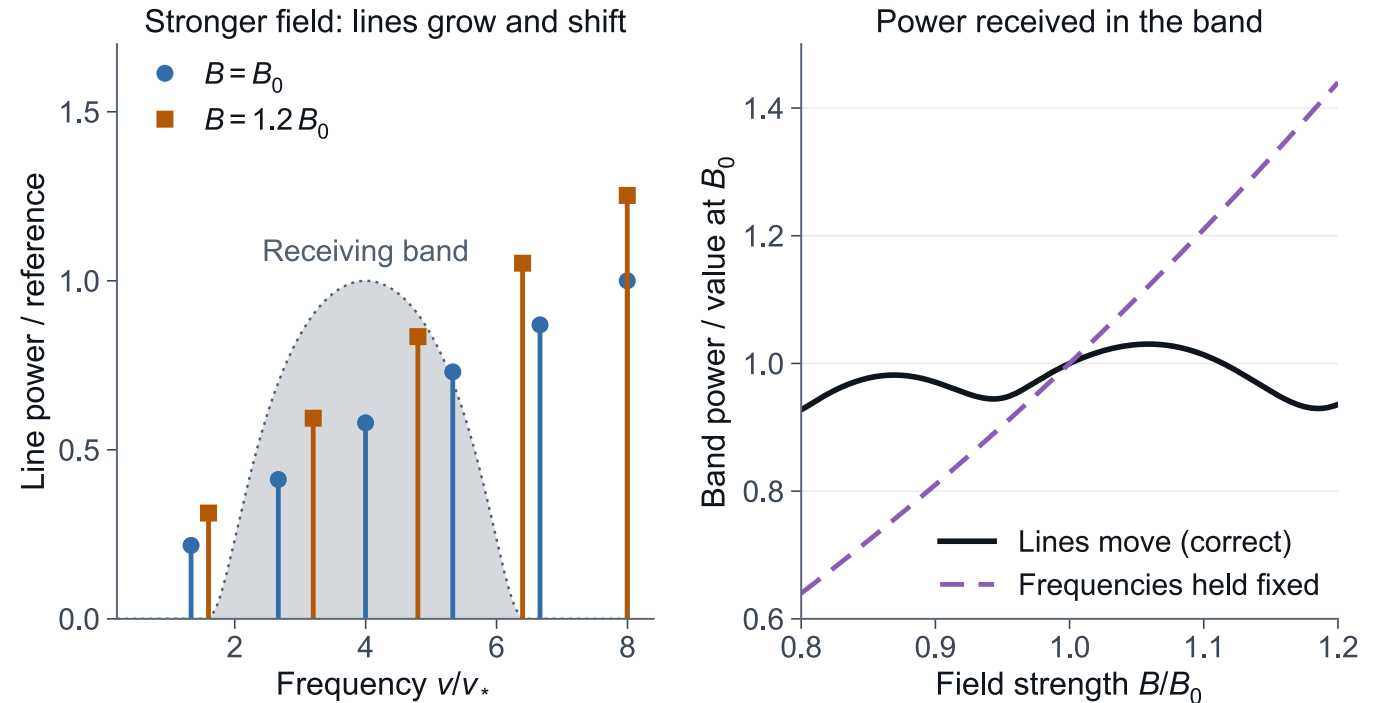
Existing low-frequency sky models disagree at the 10% level
Connect the spectrum to the emitting population.

One electron radiates lines; a channel sees them move

$$\mathcal{K}_{S,\omega}(\mathbf{p}) = \sum_{m \geq 1} S_m(\mathbf{p}) \delta(\omega - \omega_m)$$

$$\omega_m = \frac{m \omega_B}{1 - \beta \cos \alpha \cos \theta}$$

$\mathbf{p} = (\gamma, B, \cos \alpha, \cos \theta)$: energy, field, pitch angle, viewing angle. Each harmonic is fully polarised, $I_m^2 = Q_m^2 + V_m^2$; the projected-field azimuth ϕ rotates Q into $Q + iU = e^{2i\phi} Q^{\text{nat}}$.

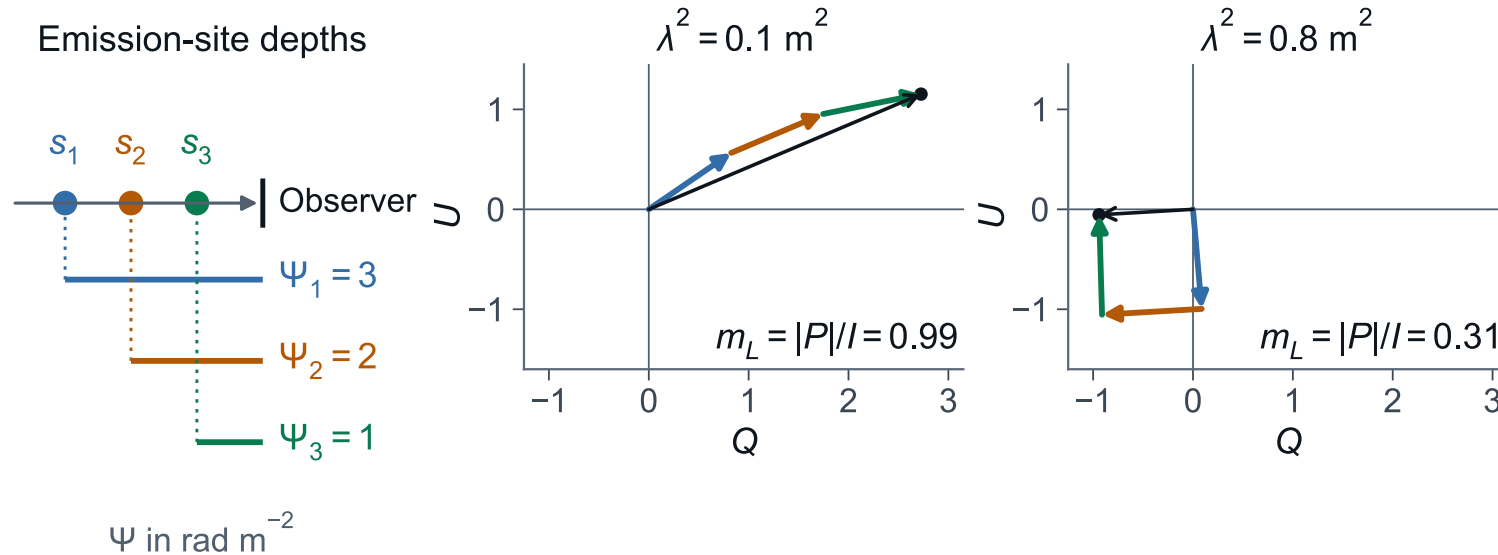


$\gamma = 20$: a 20% stronger field moves lines through the band; holding their frequencies fixed is wrong by up to 28%

Many electrons, one joint population

$$P_j^{\text{obs}} = \mathcal{N}_{\text{src}} \left\langle e^{2i\phi} \mathcal{K}_{P,j}^{\text{ch}}(\mathbf{p}, \Psi) \right\rangle_{\mathcal{P}_F}$$

Average the channel response over energy, field, angles and Faraday depth.



Emission and Faraday rotation can overlap along the same ray.

Physics fixes the basis; the population fixes the moments

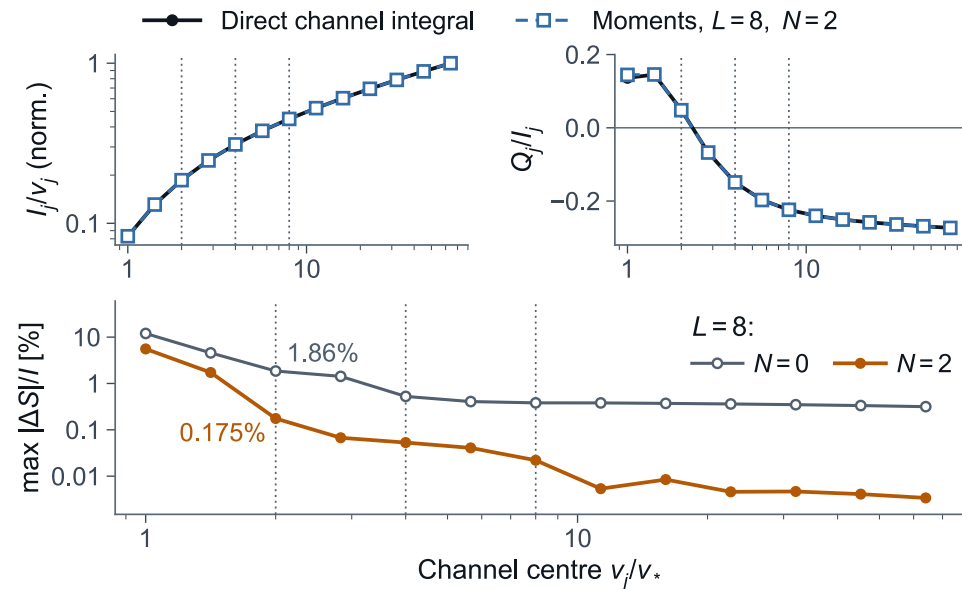
$$P_j^{\text{obs}} = \mathcal{N}_{\text{src}} \sum_A \tilde{P}_{jA} M_A + \epsilon_{P,j}$$

Response coefficients $\tilde{P}_{jA}$

Fixed by radiation physics and channels.

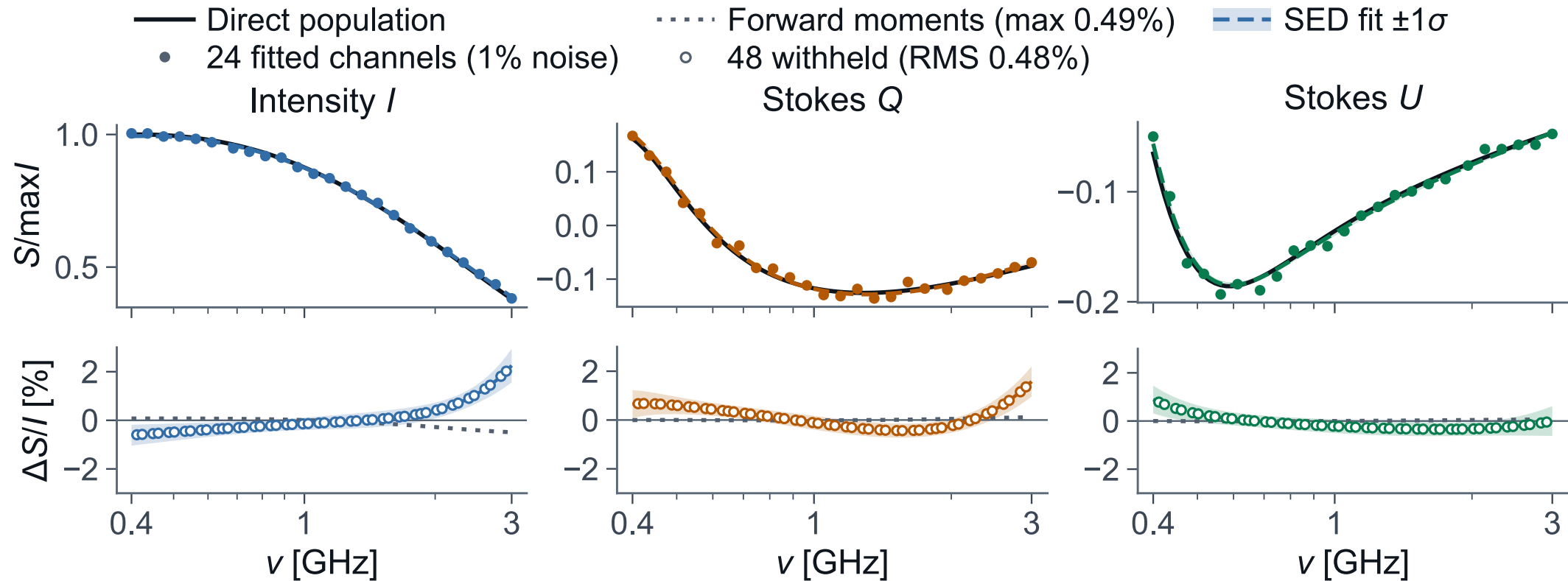
Joint moments M_A

Encode the population and its correlations.



Correlated harmonic-model population · $L = 8, N = 2$

The linear basis can fit an SED



Synthetic I , Q , U spectra · fitted channels and withheld-channel checks

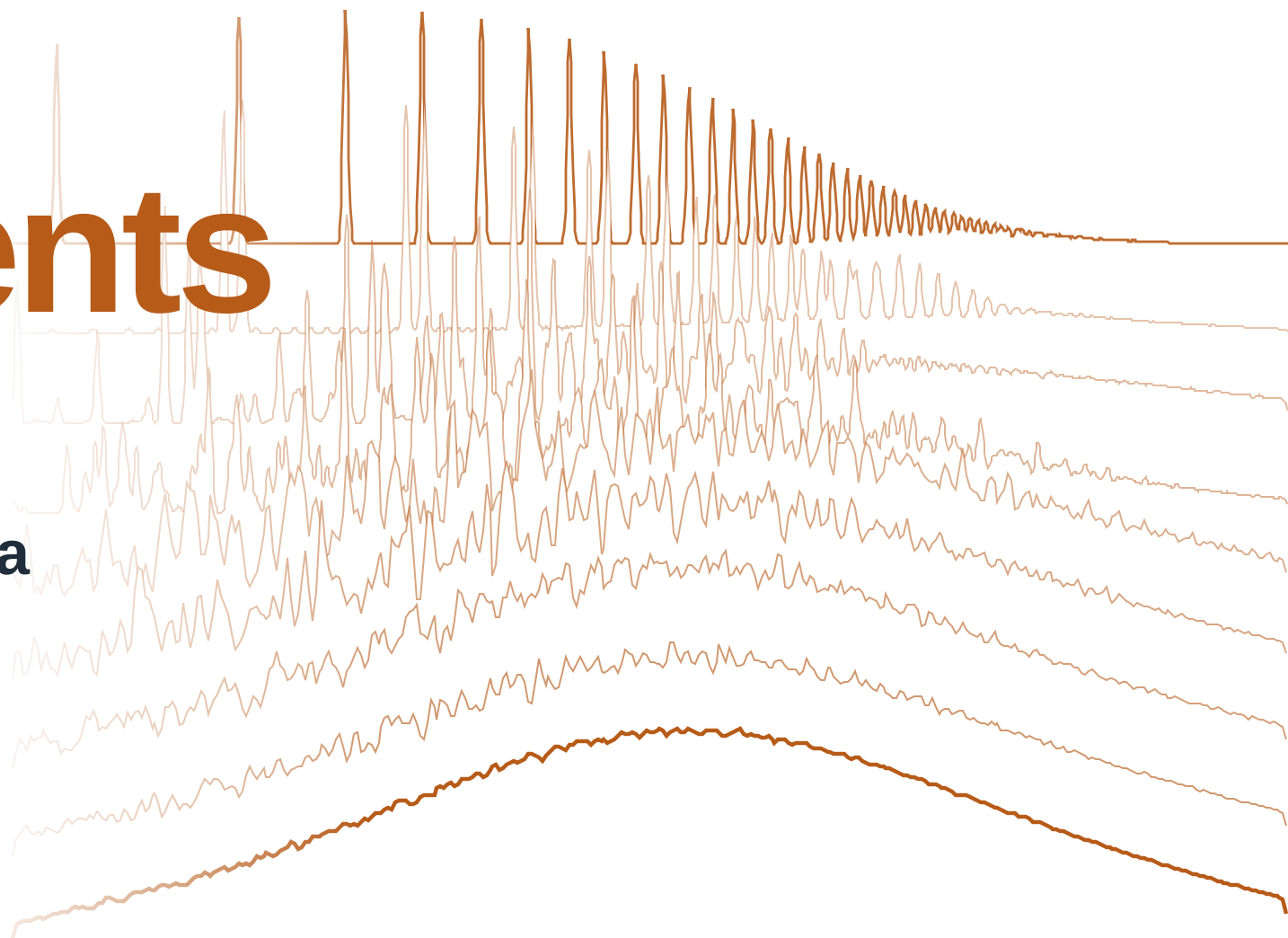
PART 2 · MOMENT EXPANSION

SyncMoments

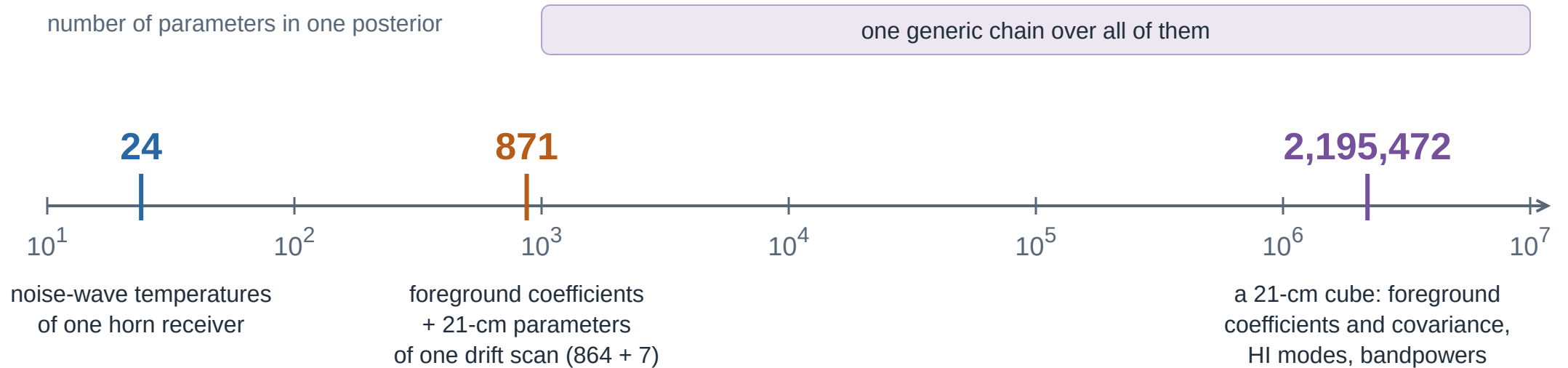
a moment expansion of channel-integrated synchrotron emission

Kernel + derivative spectra = a linear basis for the SED

Moments are its coordinates; data fix a few combinations



From 24 to two million parameters in one posterior

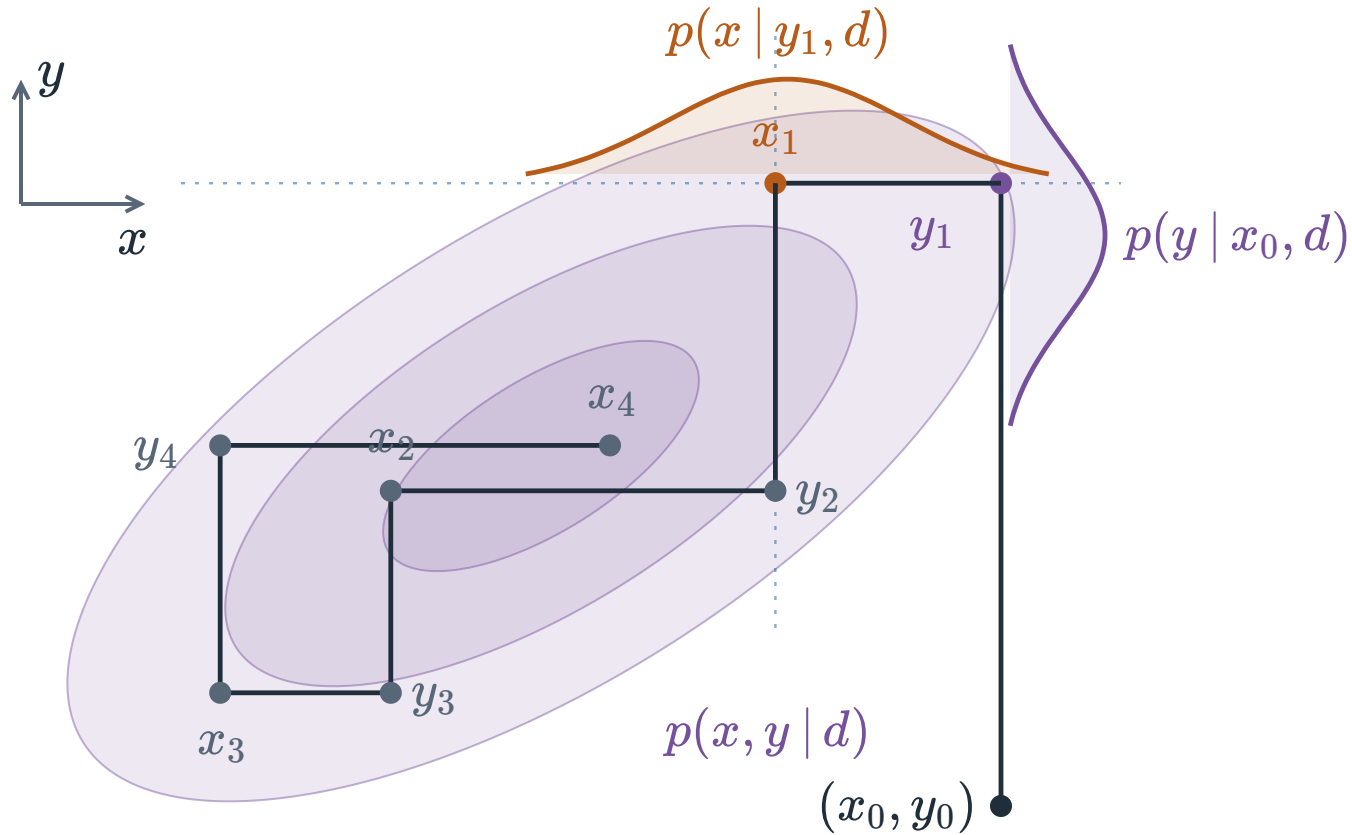


a dense covariance: 4.8×10^{12} numbers, 39 TB in float64

a random-walk chain mixes slower with every dimension

and yet: under 30 s per sweep on one 32-core node

One block at a time; the chain still targets the joint



$$y_1 \sim p(y | x_0, d)$$

$$x_1 \sim p(x | y_1, d)$$

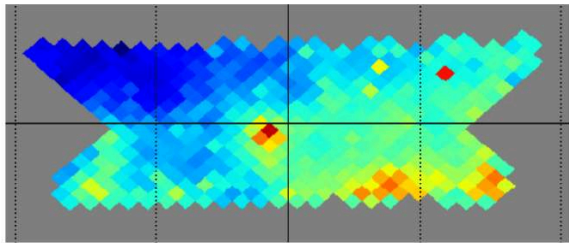
$$(x_i, y_i) \rightarrow p(x, y | d)$$

coupled blocks: correlation along the chain

affine block + Gaussian noise $\Rightarrow$ Gaussian conditional, solved or drawn exactly at any dimension

Gains, temperatures, noise, sky map: one sweep in 12 s

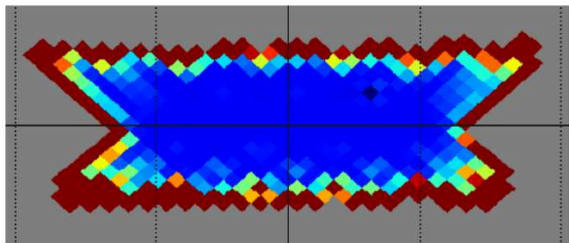
posterior mean



5 K

14 K

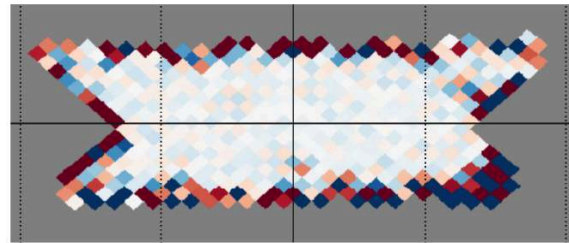
uncertainty



0 K

0.2 K

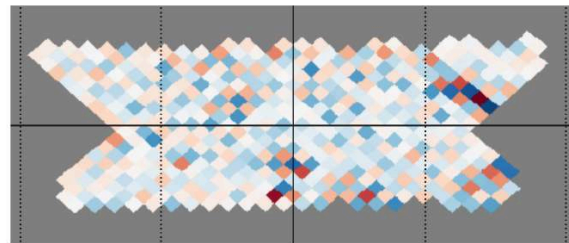
residual



-0.2 K

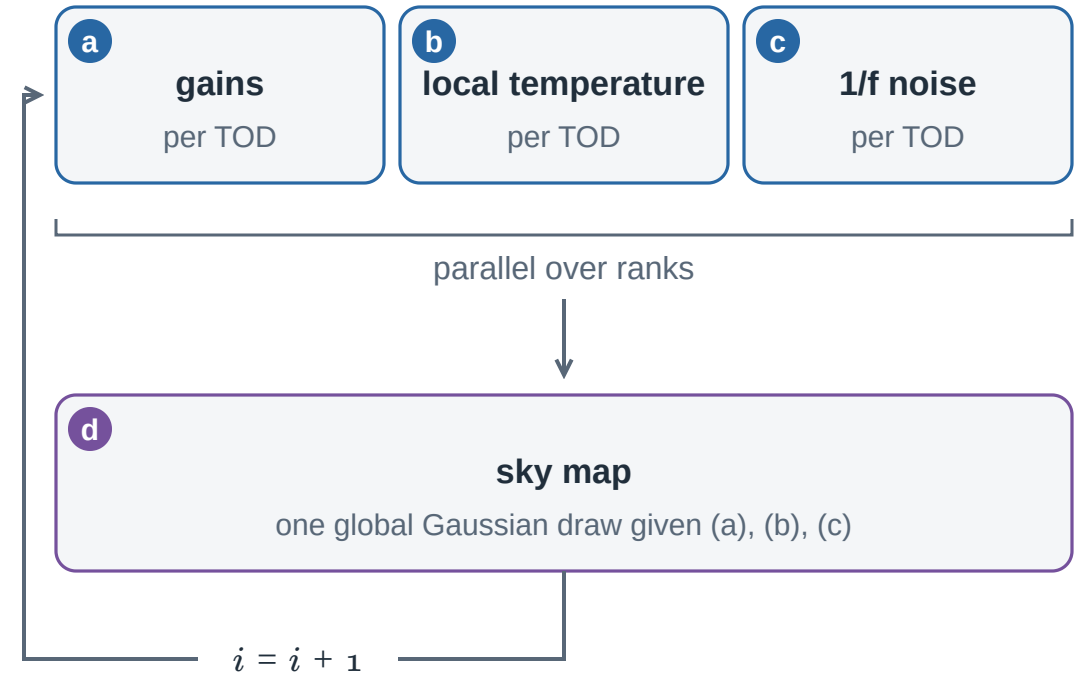
0.2 K

Z-score



-3

3



2,000 sweeps at 12 s · diode 15.0 K recovered as 14.994 K · a second scan geometry was decisive

Hydra: a family of Gibbs samplers



Hydra · several heads

2023 → 2024

hydra-pspec · 2023

Visibilities + power spectra
Kennedy et al.

hydra-beam I · 2024

Parametric HERA beam
model
Wilensky et al.

2025

hydra-beam II

Beam inference from
visibilities
Wilensky et al.

Foreground-map recalibration

Nasirudin & Bull

2026

Component separation

Murphy et al.
2.2m parameters · <30 s /
sweep

hydra-tod

Calibration + sky map · Zhang
et al.

Systematics

Dutta et al.

Commander BeyondPlanck

CMB map, power spectrum
and foregrounds drawn in turn

Planck LFI: calibration and
component separation in an
end-to-end Gibbs analysis

Eriksen et al. (2004, 2008)
BeyondPlanck (2023)

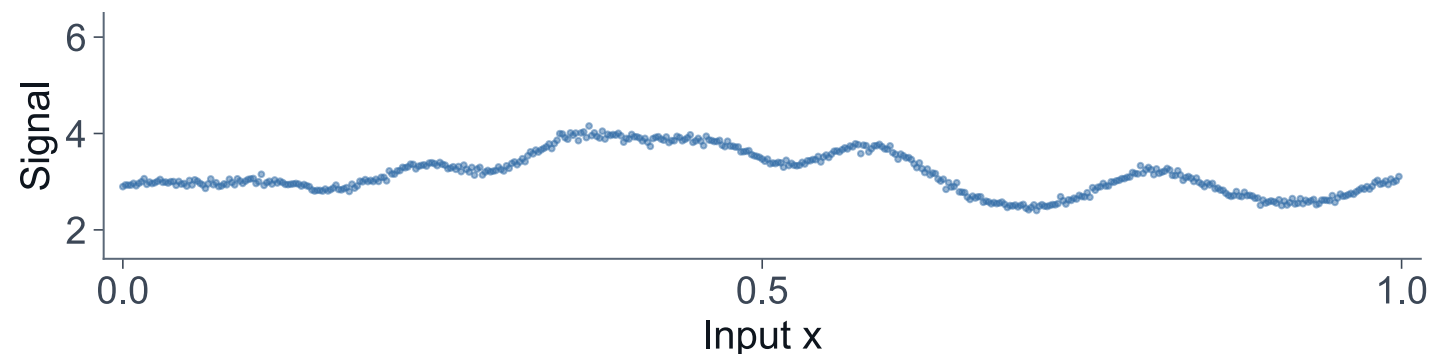
Exploit the model's Gaussian conditionals inside a joint analysis

Designing a sampler for the model

BlackJAX · NumPyro

Powerful sampling
and posterior analysis.

One model: field + nonlinear shape + gain + noise



BLOCKS

Which parameters
should move together?

METHODS

What can be sampled
or integrated exactly?

PROPOSALS

How do we make the
hard updates tractable?

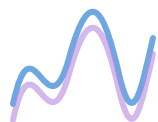
Bayesmith

Analyse the model. Compile its sampling plan.

Reusable design, before sampling.

Numerical backends execute the updates.

Model structure guides the sampling strategy



Whole-field draws

Gaussian conditionals



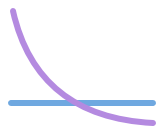
Log-space structure

Products → linear form



Analytic collapse

Fewer sampled variables



Prior analysis

Explicit choices



Automatic blocking

A model → a sweep



Structured proposals

Easier moves + correction

One model. A reusable sampling plan.

Field draws

Log space

Collapse

Priors

Blocks

Proposals

Sky and instrument in one model

Rheplicant

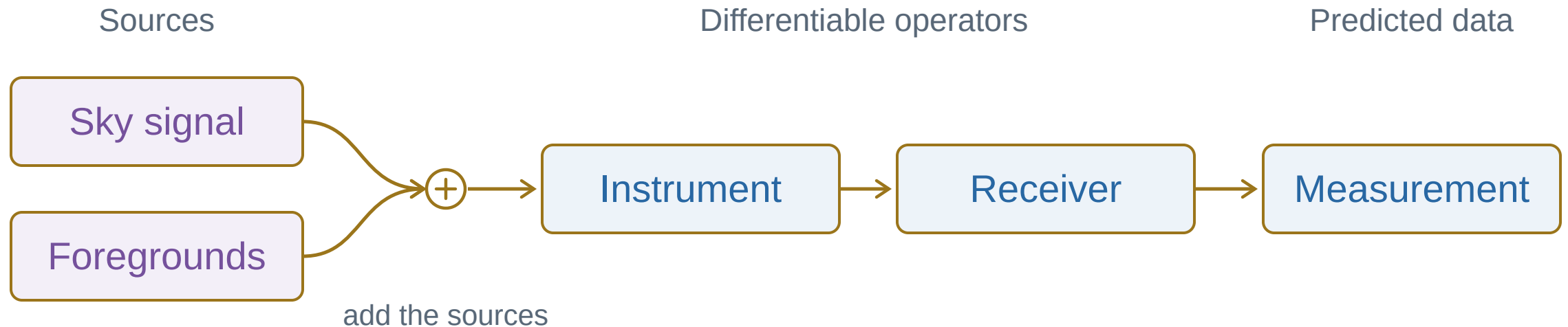
A DAG is an observation.

The same graph language, carried into the observing system.

RHINO observation and inference

Rheplicant

Compose a data model
from differentiable operators.



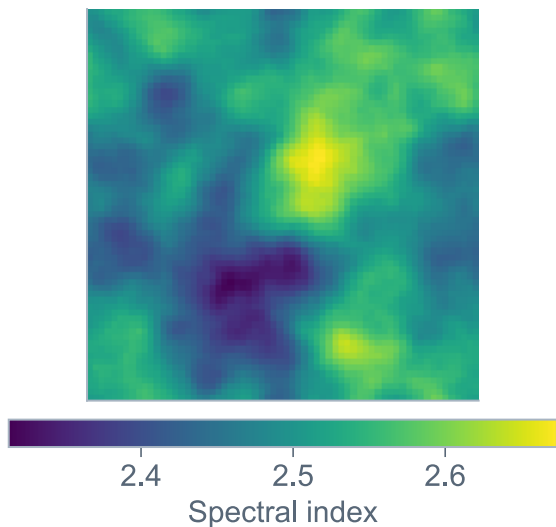
Bayesmith

Turn model structure
into a sampling strategy.

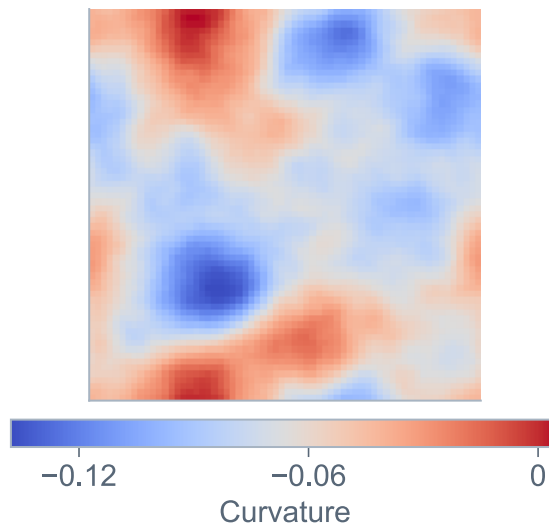
Analyse dependencies → group parameters → choose updates

A synthetic sky and its observation

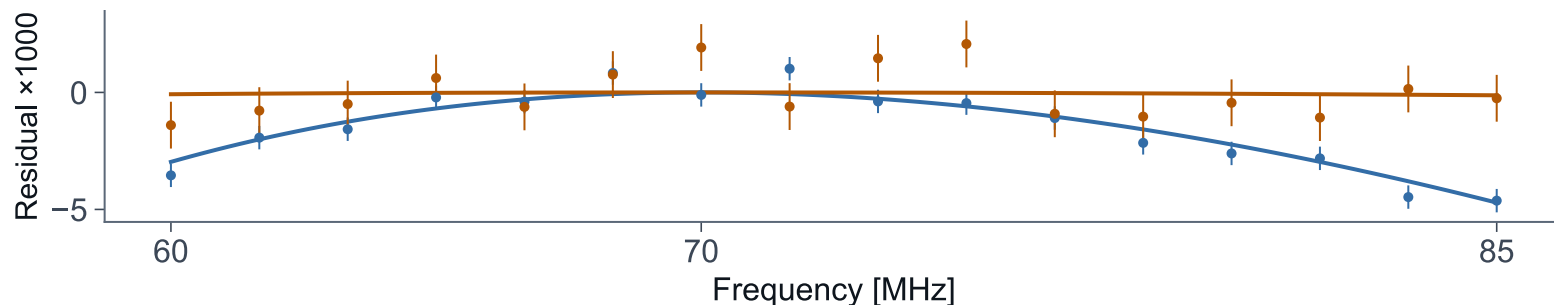
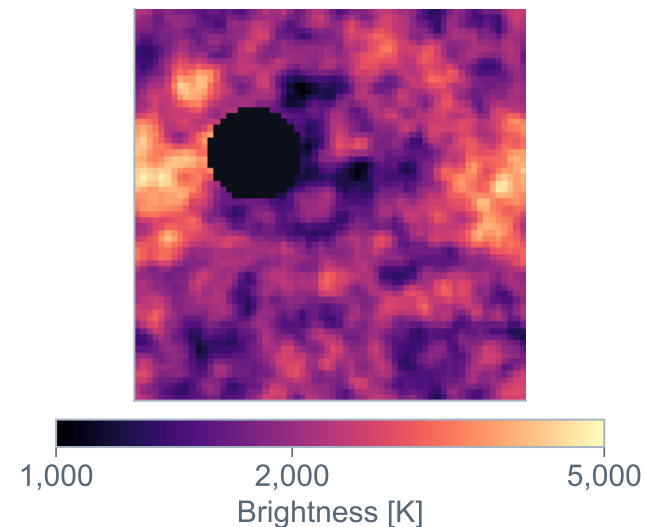
Generating spectral index



Generating curvature



Observed survey · 70 MHz

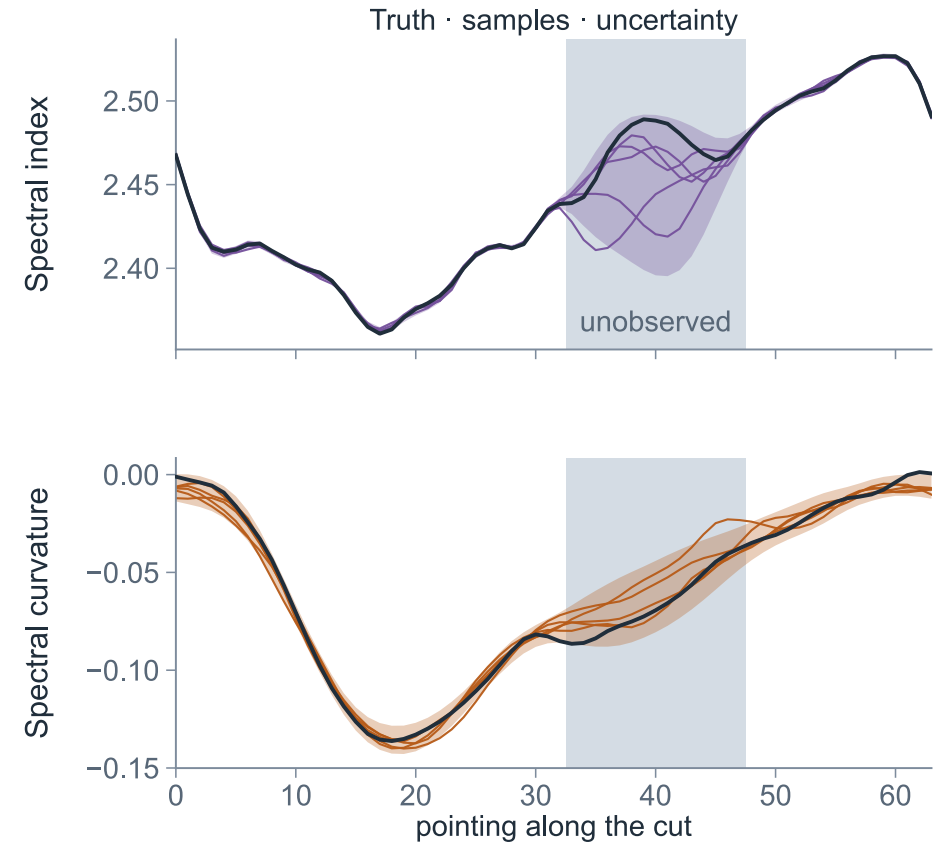
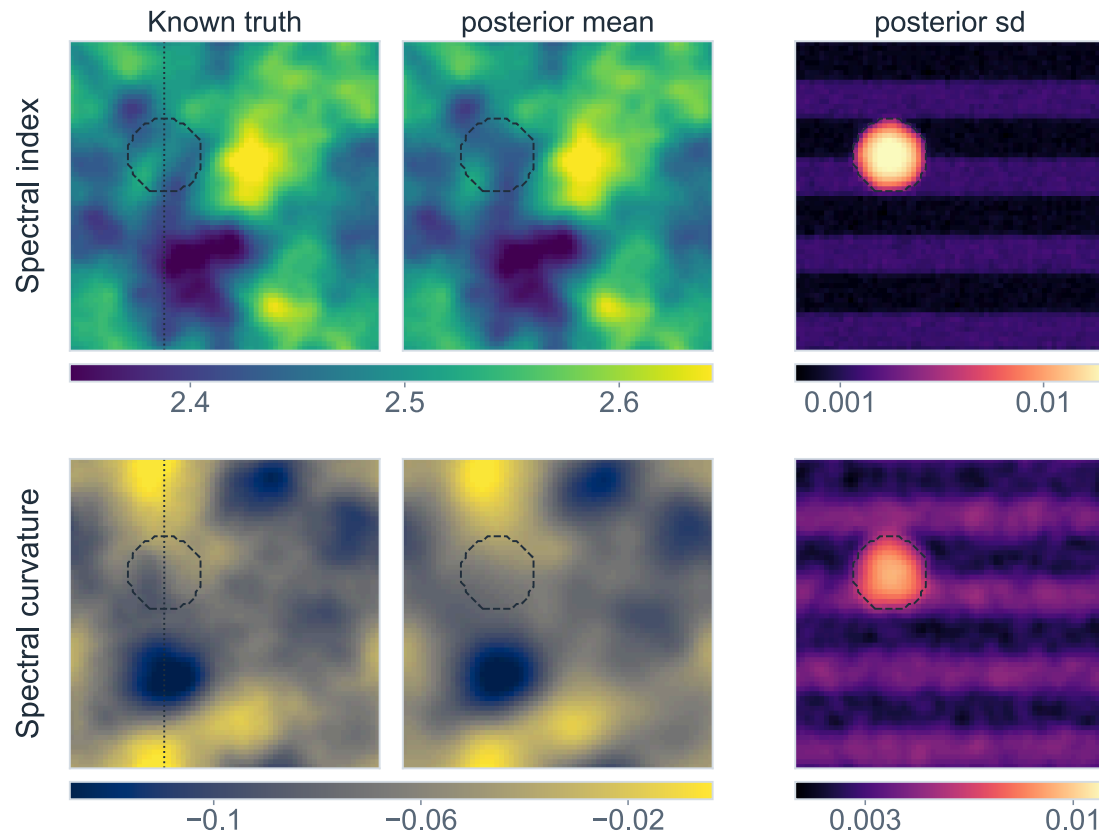


**Joint inference
across the survey**

Brightness · spectral shape

All positions together

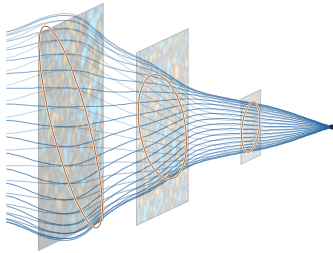
One posterior draw is a whole sky



Joint sky maps · independent draws · uncertainty follows coverage

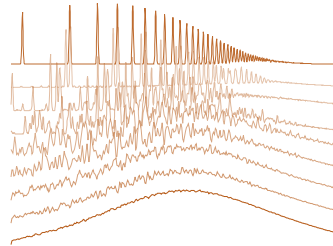
From stochastic models to inference

Stochastic optics



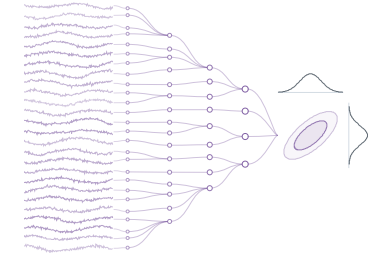
Connect matter statistics
to light propagation.

Astrophysical emissions



Connect population statistics
to observable spectra.

Bayesian inference



Connect executable models
to reusable samplers.

Make the physics explicit.
Let the model guide the inference.

Thank you. Questions welcome.

Rheplicant · Bayesmith · [Technical backup](#) ↗

Choose a layer to open

Part 1 · Lensing

Toy: Gaussian baseline	UV sensitivity, MC check
Toy: exact tier	E/B
Toy: one dial	Map of corrections
Toy: channel budget	Completeness rule
Jacobi and Sachs	What moves the amplitude
Rates, maps, images	One-point PDF
Response construction	PDF toy model
Order counts	PDF propagation
Driver statistics from $P(k)$	PDF tails
Two-point scope	The toolchain

Part 2 · Synchrotron

Faraday screen cumulants	Assumptions and limits
Correlations change spectra	Error in the science parameter
What varies the spectrum	Sky models disagree

Part 3 · Inference

Observation + inference: examples	Spectral maps: what fits measure
Historical Demo C: plan	Spectral maps: exactness and cost
bayesmith: five routes	Spectral maps: the prior
Example 03: chains, record	Spectral maps: 16 realisations
RHINO pipeline	RHINO-sized beam, global β
Agent: the signal path	Gaussian field reweighting
Agent: before and after	Mixture population: numerical test
Agent: the session view	Poisson population: numerical test
RHINO identifiability	Spectral maps: setup
Exact block vs NUTS	Spectral maps: cost and diagnostics
Historical Demo C: combinations	Spectral maps: physical/model checks
Historical Demo C: stress band	Bayesmith: model and generated sweep
Poisson sky: the model	Bayesmith: saved posterior evidence
Poisson sky: rate inference	References
Campbell cumulants	
Reweighting validity	
Draw cost	

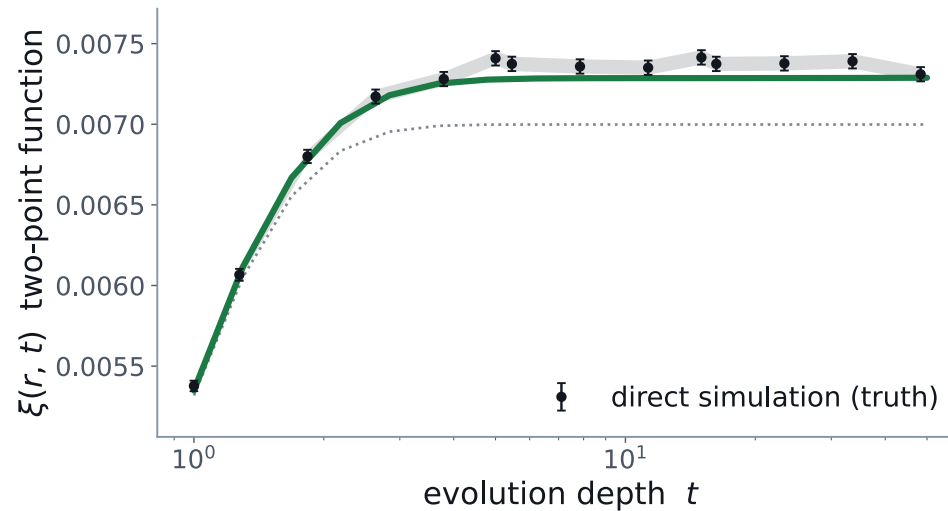
Gaussian baseline: the diagram sum approaches the simulation

$$\dot{\phi}_a = -\gamma \phi_a + s F_{abc} \phi_b \phi_c + \eta_a$$

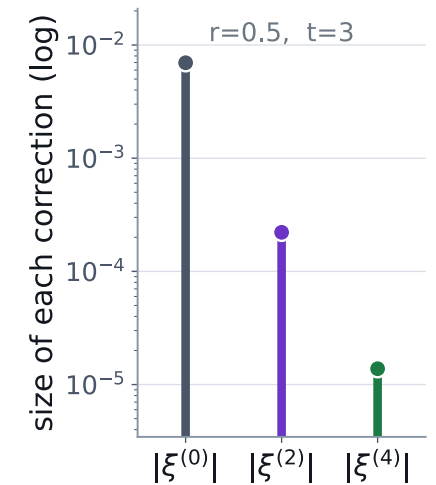
Two coupled fields on a line: the same shape as the Sachs system, but small enough to integrate directly. For now the drive η is **Gaussian**:

$\mathcal{K}^{(n)} = 0$ for $n \geq 3$, so there are no cumulant vertices.

summing diagrams to order 4

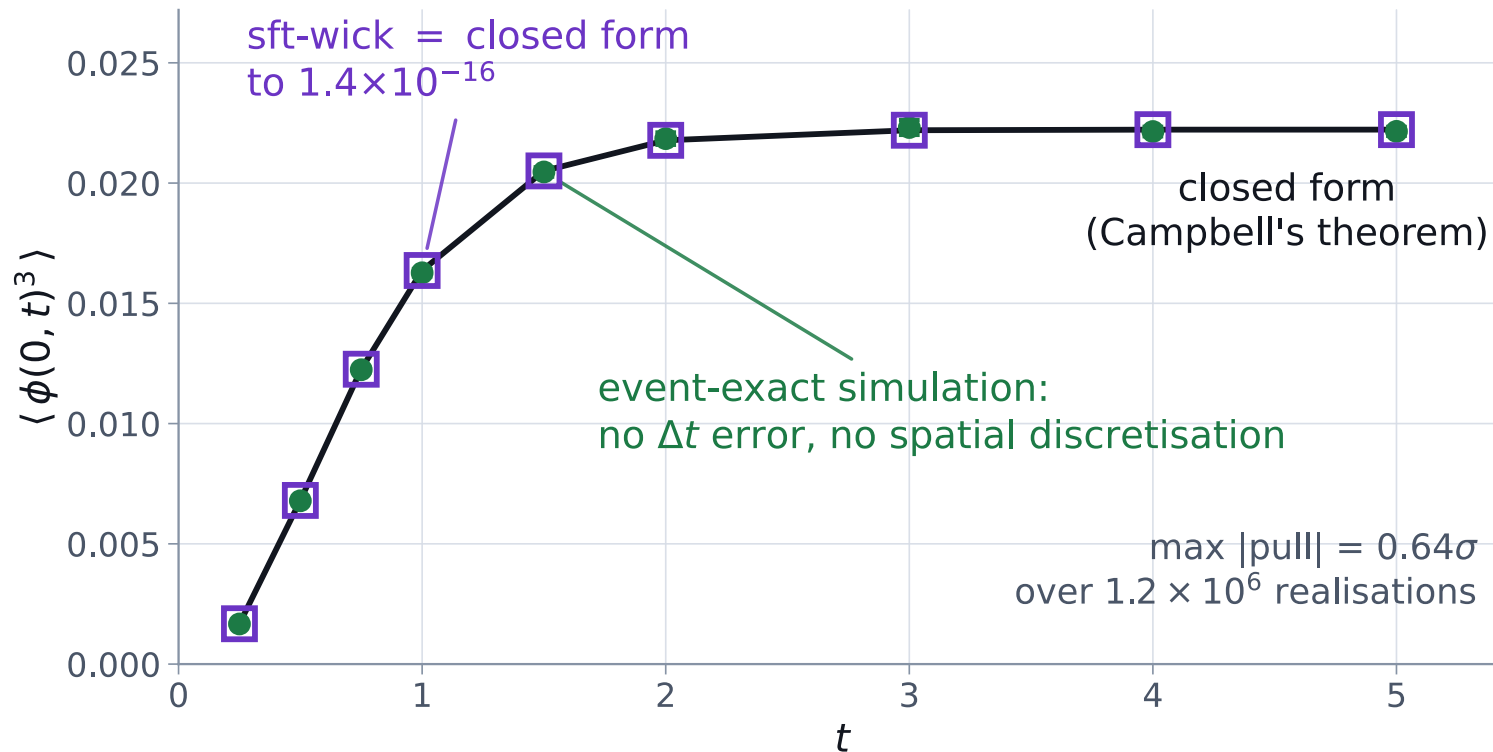


71 Feynman diagrams



Each order brings the sum closer to the simulation. At order 4 the simulation still lies above it at 12 of 14 times, by up to 3.0 Monte-Carlo error bars. A Gaussian drive cannot show the effect we are after.

Where truth is known, we hit it exactly



$$1.4 \times 10^{-16}$$

sft-wick vs closed form

0.64 σ

vs an event-exact simulation, 1.2×10^6
realisations

Jacobi and Sachs describe the same propagation

Jacobi map D

$$D'' = \mathcal{T} D$$

linear in D , random multiplicative tidal matrix; angular separation at the observer $\rightarrow$ transverse separation at the source

Optical rates S

$$S' + S^2 = \mathcal{T}$$

Riccati form: additive forcing, quadratic drift; the Sachs scalars are the components of S

$$S = D' D^{-1}, \quad \mathcal{A} = D / \bar{D}, \quad \mu = (\det \mathcal{A})^{-1}$$

A rate integral recovers the matrix map only through ordered evolution; magnification and reduced shear are readouts of $\mathcal{A}$, not of $\int S$.

Rates, maps and image observables differ beyond leading order

$$\mathcal{A} = (1 - \kappa_A)I - \gamma_+ \sigma_3 - \gamma_\times \sigma_1 + \omega \epsilon$$

$$g = \frac{\gamma}{1 - \kappa_A}, \quad \mu = (\det \mathcal{A})^{-1}$$

$$\kappa = - \int_0^{\lambda_s} \delta\theta \, d\lambda, \quad \kappa_{\ln} = - \ln(D_A/D_0)$$

The two-point slides compute $\langle \kappa \kappa \rangle$ with this κ , the paper's linear-order convergence. Beyond linear order the same integral equals $\frac{1}{2} \ln \mu$ at fixed affine endpoint on the positive-parity branch and differs from κ_A . All of these coincide at leading order and differ beyond it.

The auxiliary field enforces the optical equation

$$1 = \int \mathcal{D}\mathcal{X} \, \delta[f[\mathcal{X}] - \varphi] \, \mathcal{J}, \quad f = \dot{\mathcal{X}} - A\mathcal{X} - F\mathcal{X}\mathcal{X}, \quad \mathcal{J} = \text{Det} \frac{\delta f}{\delta \mathcal{X}}$$

$$\delta(r) = \int \frac{d\tilde{\mathcal{X}}}{2\pi} e^{-i\tilde{\mathcal{X}}r}, \quad W[J] = \sum_{n \geq 2} \frac{1}{n!} \int \mathcal{K}^{(n)}(1, \dots, n) J_1 \cdots J_n$$

$$S = i \int \tilde{\mathcal{X}}(\dot{\mathcal{X}} - A\mathcal{X} - F\mathcal{X}\mathcal{X}) - W[i\tilde{\mathcal{X}}]$$

$$\langle \mathcal{X}\tilde{\mathcal{X}} \rangle_0 = -iR, \quad \langle \mathcal{X}\mathcal{X} \rangle_0 = R\mathcal{K}^{(2)}R^\top, \quad \langle \tilde{\mathcal{X}}\tilde{\mathcal{X}} \rangle_0 = 0$$

Fixed initial state at the observer; mean driver matched to the background. Pre-point (Itô) slicing: the Jacobian $\mathcal{J}$ is field-independent, and in the time-sliced form it is 1, so $\langle 1 \rangle = 1$.

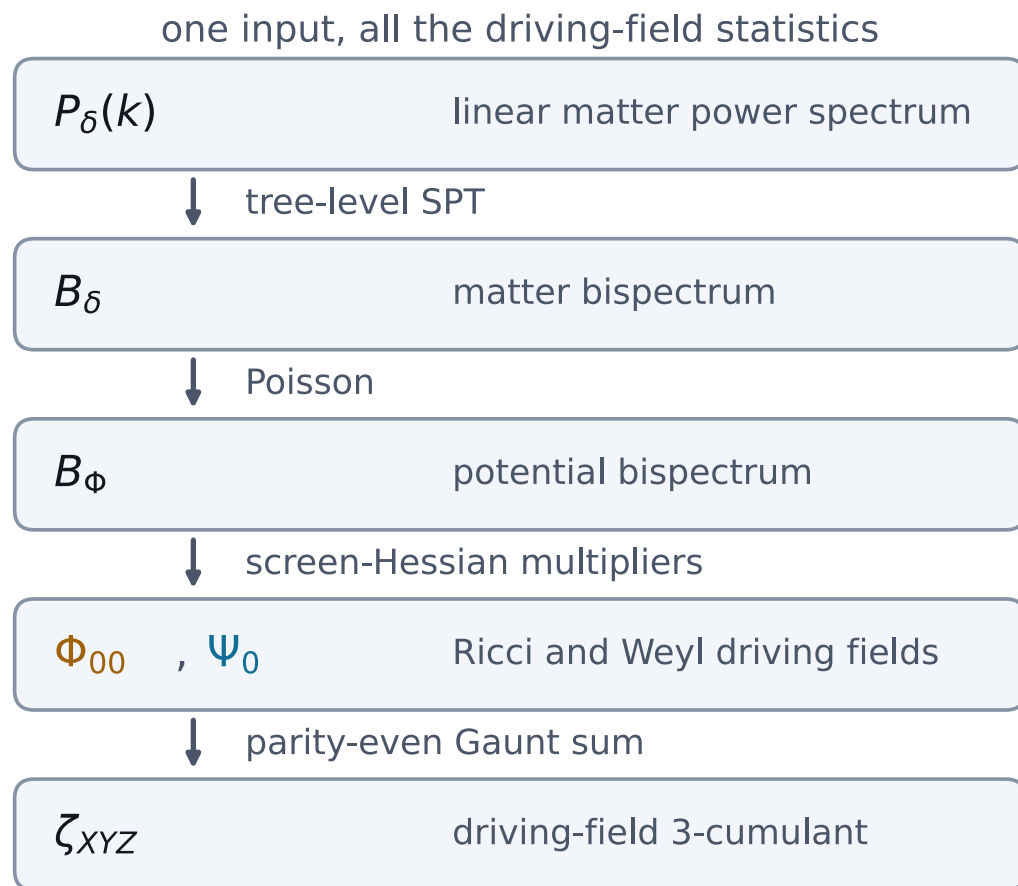
Several independent order counts, and a one-line mixing identity

Count	Example
Matter perturbations	linear P , tree-level B from gravity
Driver cumulants	$\mathcal{K}^{(2)}, \mathcal{K}^{(3)}, \mathcal{K}^{(4)}, \dots$
Optical insertions	FF, FK_3, F^3K_3, F^3K_5
Source-field order	$\gamma^{(1)}, \gamma^{(2)}, \dots$ in the readout

Why a two-point output reads higher input statistics, in one line: take a centred unit-variance X with cumulants c_3, c_4 and a quadratic response $Y = X + a(X^2 - 1)$.
Exact for the polynomial map; a toy for the structure, not an optical coefficient.

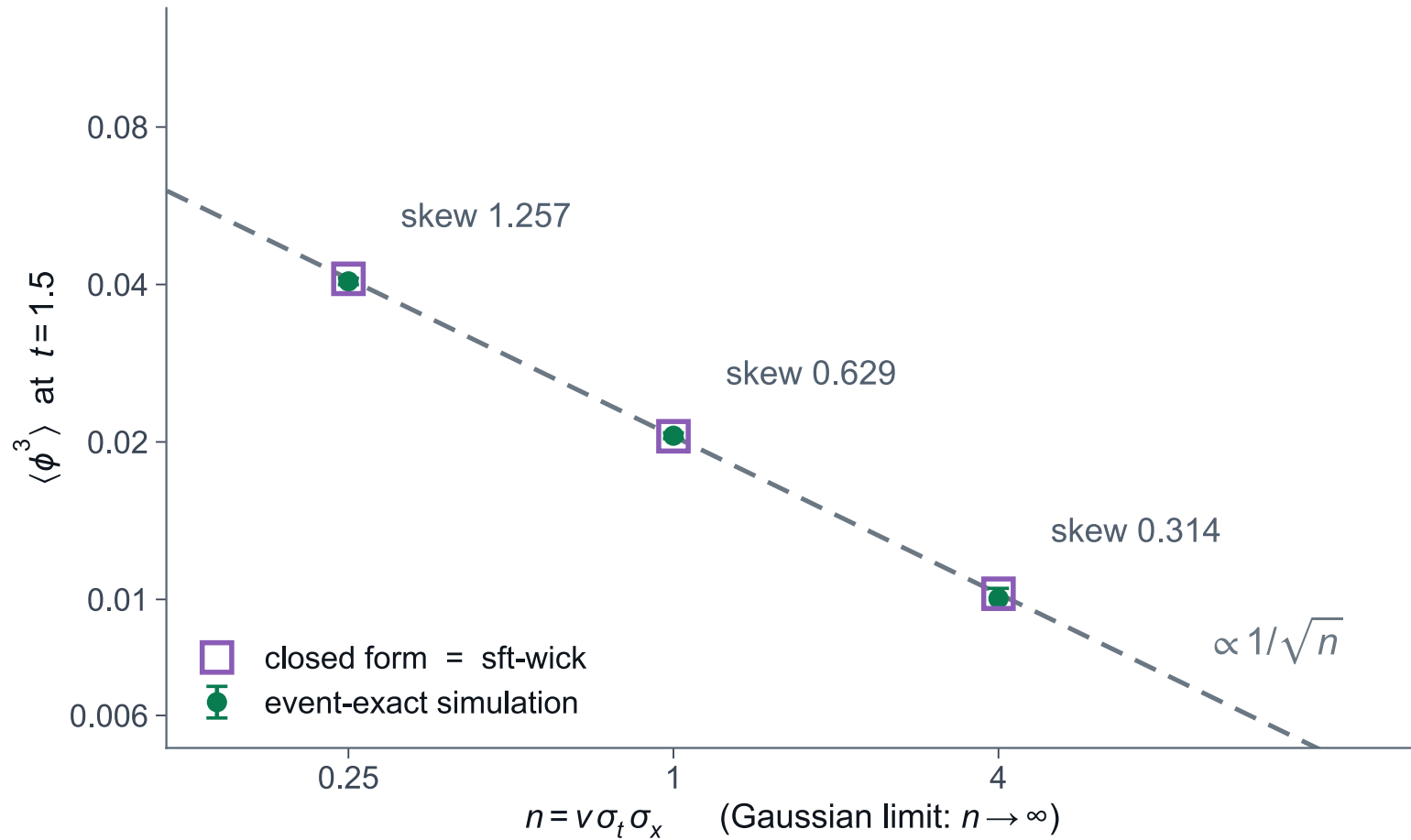
$$\text{Var } Y = 1 + \underbrace{2a c_3 + a^2 c_4}_{\text{non-Gaussian driver}} + \underbrace{2a^2}_{\text{Gaussian, nonlinear}}$$

Driving-field statistics from one input



$P_\delta \rightarrow B_\delta$ (tree-level SPT) $\rightarrow B_\Phi$ (Poisson)
 $\rightarrow (\Phi_{00}, \Psi_0)$ through screen-Hessian
 multipliers L^2/χ^2 and $\sqrt{L^2(L^2 - 2)}/\chi^2 \rightarrow$
 driver cumulants by a parity-even Gaunt sum.
 Exact Wigner-3j below $\ell = 60$, Limber above;
 FFTlog double-Bessel radial integrals.

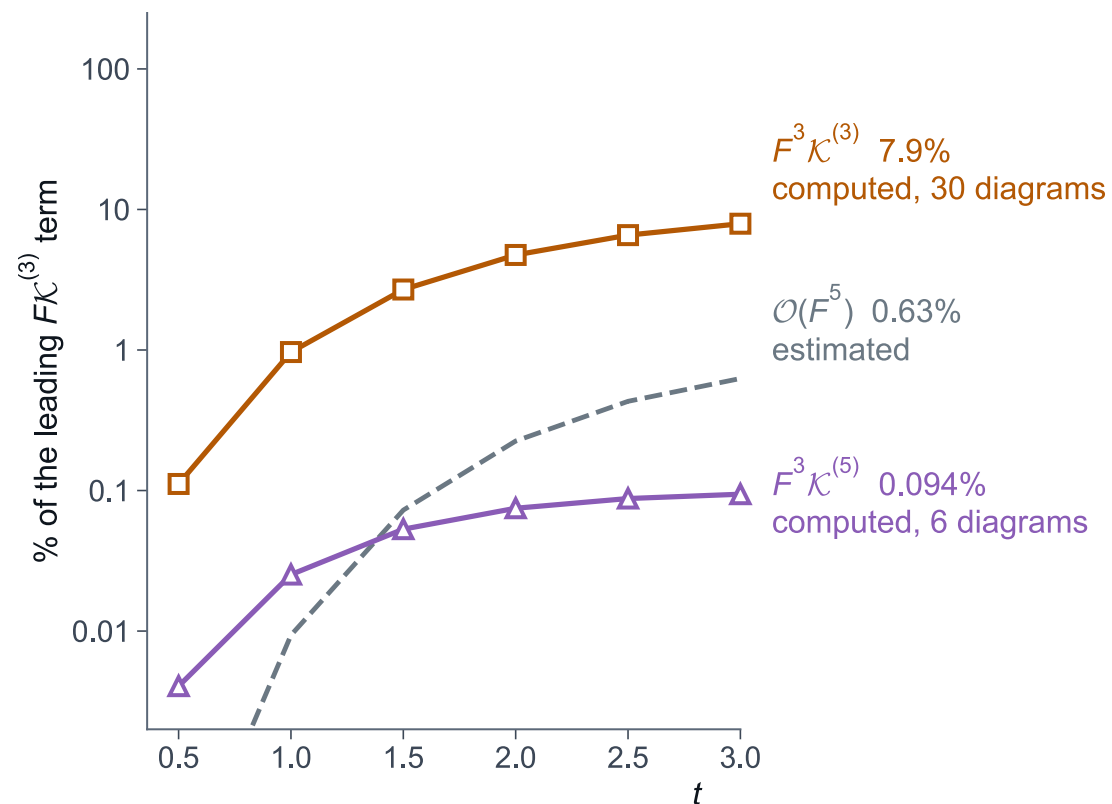
One dial moves the non-Gaussianity and nothing else



$\mathcal{K}^{(2)}$, and every Gaussian scale, is held fixed by a compensating h : only the non-Gaussian channel moves

$\times 4$ in $n \Rightarrow \div 2.0000$ in $\langle \phi^3 \rangle$
(predicted $\div 2$ exactly)

What is computed, and what is left



The $K^{(5)}$ ladder is computed: 0.09% of the leading channel at $t = 3$.

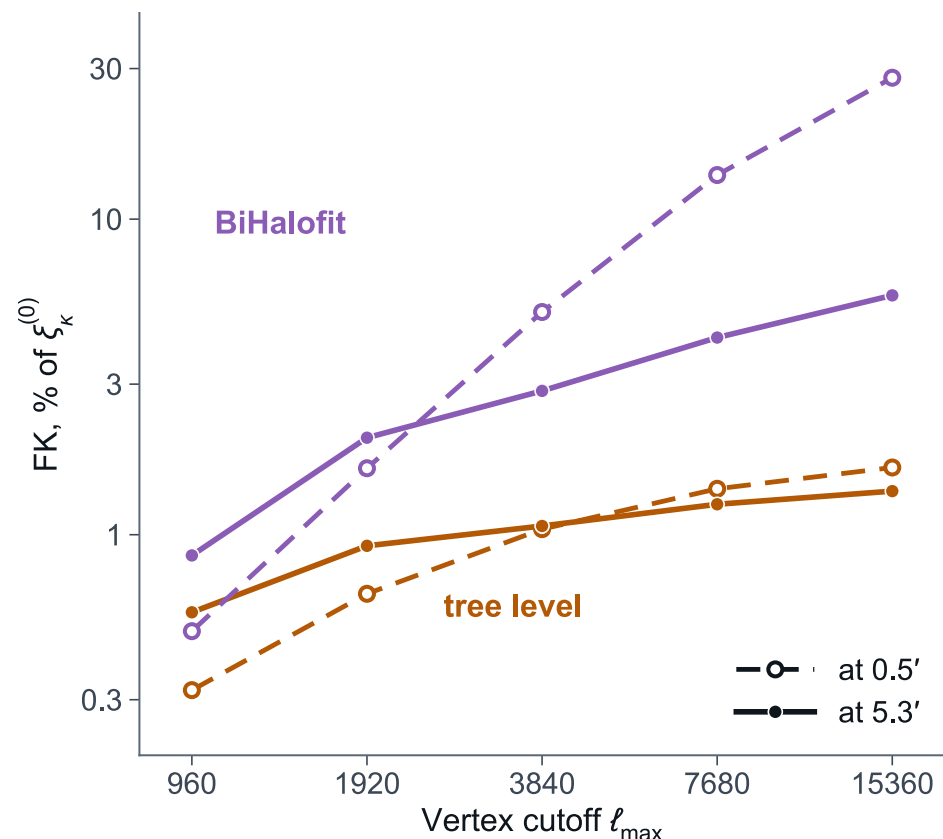
Not computed: $F^4 K^{(3)} K^{(3)}$, the next order the selection rule and the symmetry allow, and $\mathcal{O}(F^5)$, estimated geometrically at 0.6% at $t = 3$.

The three-point check at $F = 0$ and the two-point leak test the expansion, the enumeration and the evaluation. They do not test the lensing vertex table, the Limber reduction, the cutoff or the bispectrum model.

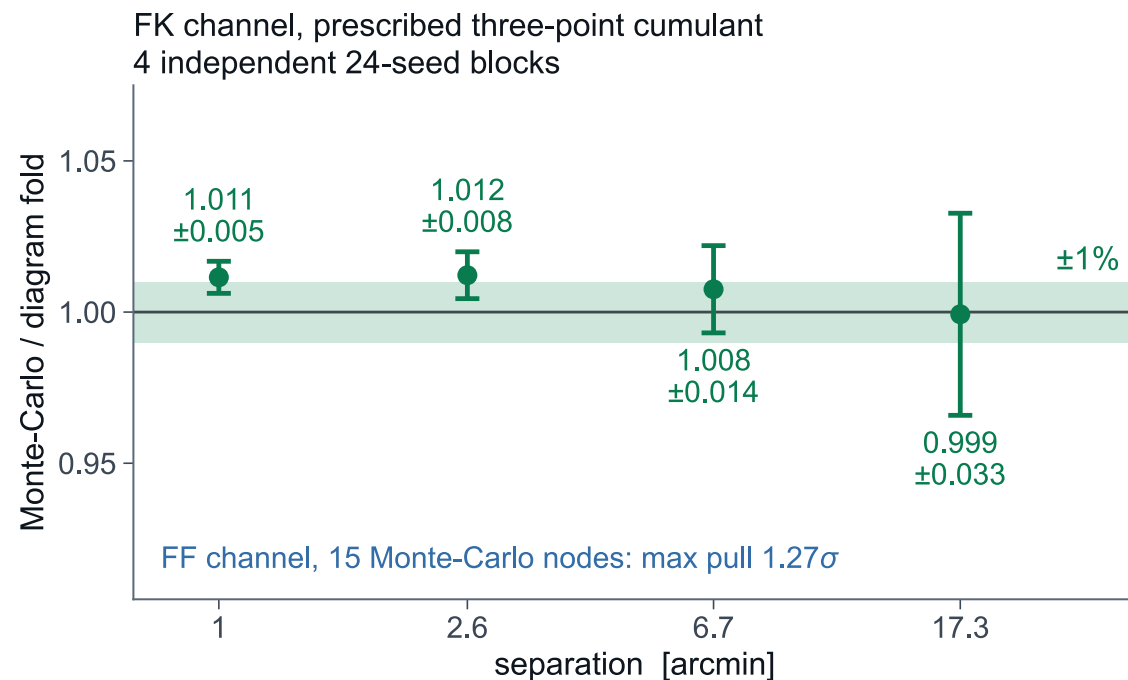
The saved FF/FK calculation has a precise boundary

Statistic	raw $\langle \kappa_1 \kappa_2 \rangle$ and ξ_+ , $\kappa = - \int \delta \theta \, d\lambda$ at linear order; FF includes an induced-mean product
Path & reference	Born-path Sachs sector, fixed affine endpoint 2313.03 Mpc; nominal $z_s = 5$
Input	linear P , tree-level B ; Limber equal-shell $\mathcal{K}^{(3)}$; vertex cutoff 15360
FK / Order-0, ξ_κ	1.79% (0.5') · 1.46% (2.06') · 1.47% (5.30') · 1.43% (10.77') · 1.42% (27.7') · 1.28% (56.3')
FK / Order-0, ξ_+	1.08% · 0.71% · 0.56% · 0.44% · 0.34% · 0.31%
Code	sft-wick pinned at a4befdc with the nonlocal-coupling leg-order fix; anchors reproduced at rtol 10^{-12}

UV sensitivity of the collapsed configuration, and a Monte-Carlo check of the fold



post-fix cutoff ladder (5 Oct), a separate product
(true $z = 5$ plane): tree level 1.34–1.40% in 2–11'



direct Monte Carlo of the Sachs equation vs the diagram fold

Field order and vertex order are different counts

$$\gamma = \gamma^{(1)} + \gamma^{(2)} + \dots, \quad B^{(1)} = 0$$

$$\Delta C_\ell^{BB} = \epsilon^4 \langle B^{(2)} B^{(2)*} \rangle + 2\epsilon^5 \operatorname{Re} \langle B^{(2)} B^{(3)*} \rangle + \dots$$

Channel	Vertex order	Source-field order	BB status
FK_3	2	3	zero: one external linear B leg
Gaussian FF	2	4	first allowed $\langle B^{(2)} B^{(2)*} \rangle$; legacy amplitude not promoted
FFK_4	3	4	allowed; not computed
$FFFK_3$	4	5	first bispectrum BB channel; not computed

Tail predictions require a law or a controlled closure

Finite-memory toy

positive compound-Poisson law; zero start; independent linear and full-Jacobi banks; six models ($\ell \in \{0.03, 0.1, 0.3\}$, two mark scales) all pass the campaign checks

Finite samples

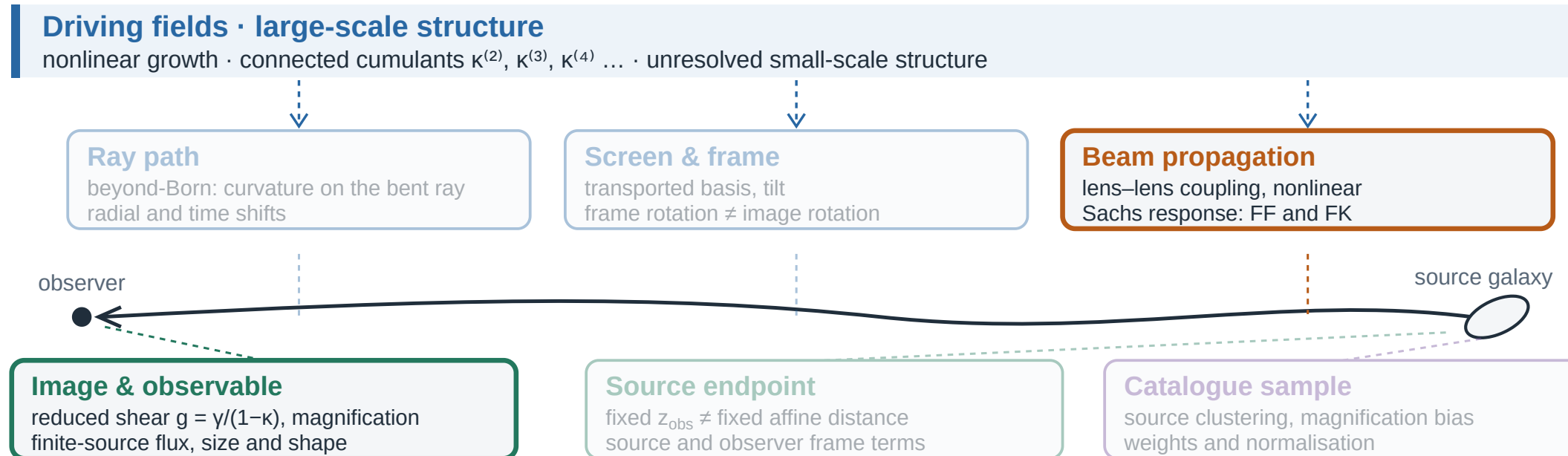
histograms estimate populated bins; two-sample KS distance within $3\sqrt{1/N_1 + 1/N_2}$; rare-event tails need more control

Finite cumulants

moment agreement does not establish positivity or a unique tail

The no-event atom has probability e^{-10} for the first law; an exact smooth density is not assumed. A 40-Mpc-beam cosmological white-driving completion exists in the preparation assets and is kept out of the main sequence.

Completeness is a rule at fixed observable and order



This talk: driver statistics $\times$ nonlinear optical response, read out as $\kappa = -\int \delta\theta \, d\lambda$ (two-point) and $\kappa_{\text{ln}} = -\ln(D_A/D_0)$ (PDF)

$$g = \frac{\gamma_A}{1 - \kappa_A}, \quad \xi_w = \frac{\langle w_1 w_2 O_1 O_2 \rangle}{\langle w_1 w_2 \rangle}$$

Choose the observable and the amplitude order; then every retained contribution enters **once**, including means, denominators and cross terms. The two applications open the highlighted corner with everything else held fixed.

A model where every link of the chain is specified exactly

Driver

compound-Poisson events
along the ray, rate $\nu = 10$:
Ricci mark $J = 0.02$, Weyl
modulus $|K| = 0.04$, random
angle; exponential memory
 $\ell = 0.1$

→

Dynamics

full Jacobi matrix $D'' = \mathcal{T}D$,
integrated path by path (50k
paths); or the **linear** projection
through the exact response
window (200k paths)

→

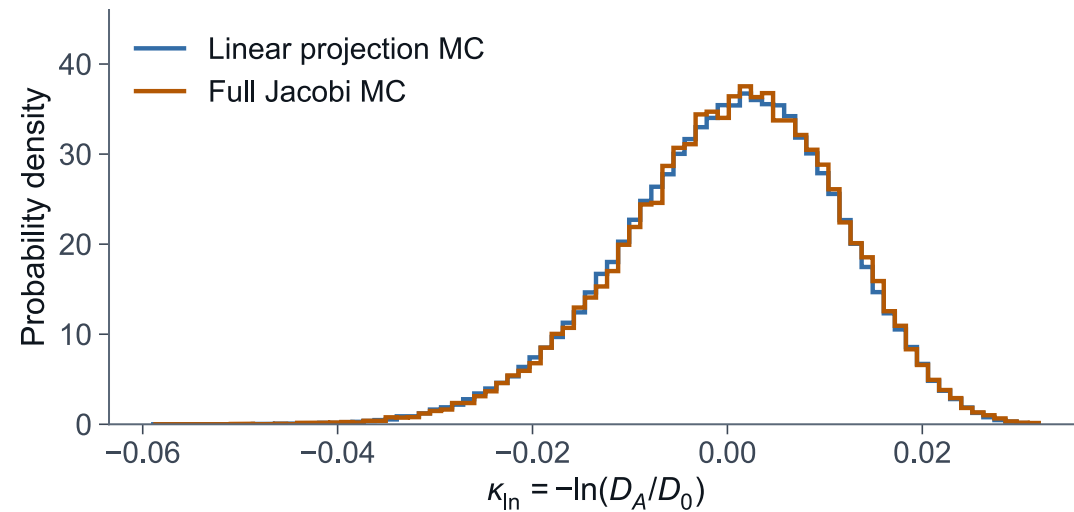
Readout

$\kappa_{\ln} = -\ln(D_A/D_0)$ at
 $v_s = 1$, reference $D_0(v) = v$

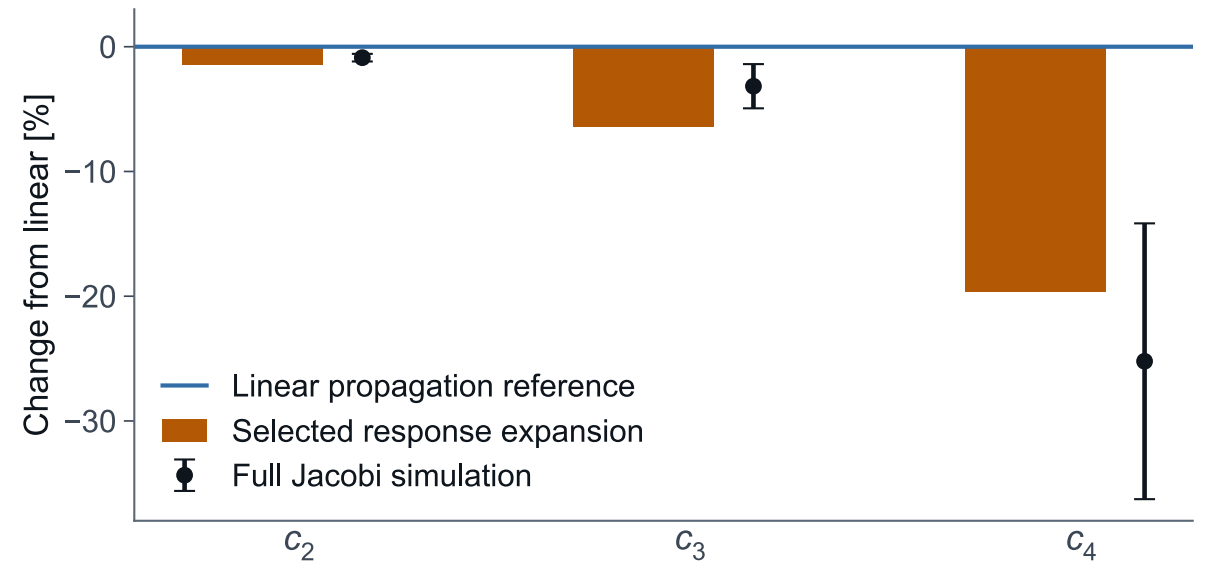
The law is positive and known in closed form, so every driver cumulant is exact and the **linear** output distribution is an exact prediction, not a fit. Nonlinearity enters only through the propagation.

Two experiments: hold the law, change the propagation; hold the covariance, change the law. Then ask where measurement noise would enter.

Histogram barely moves; fourth cumulant drops by 25%

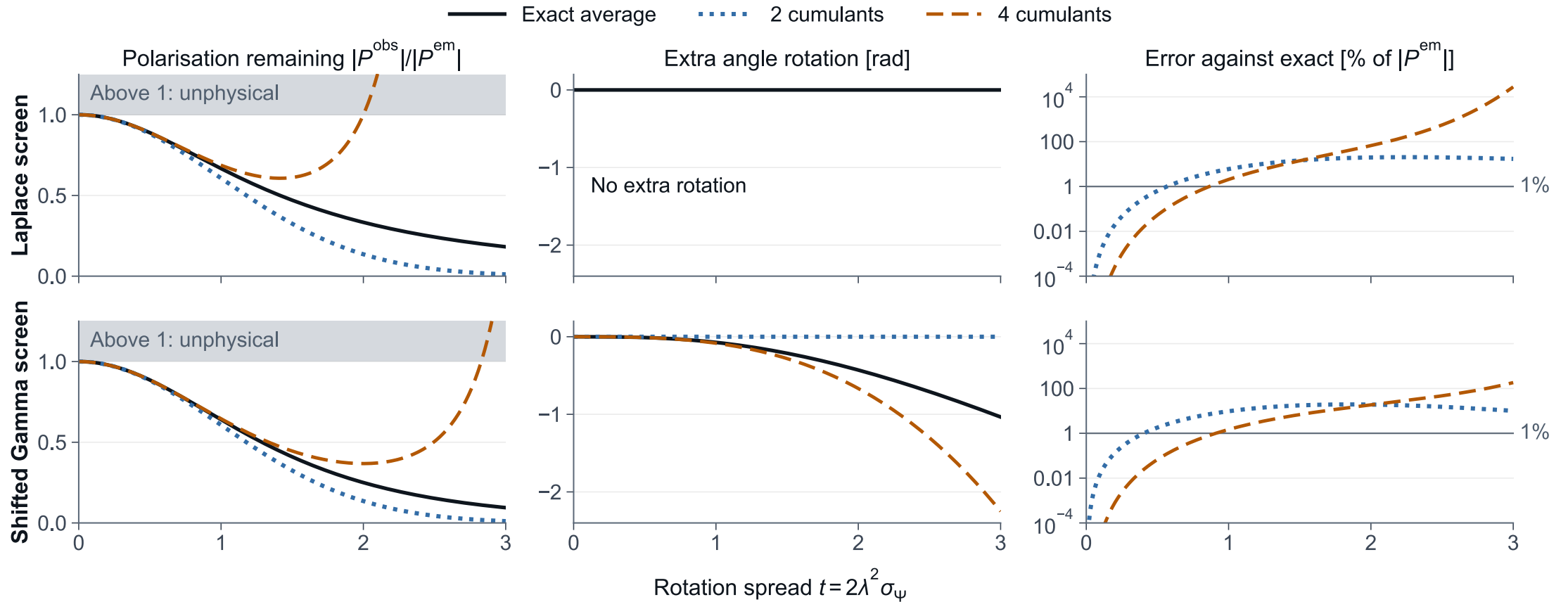


linear projection vs full Jacobi, same driver



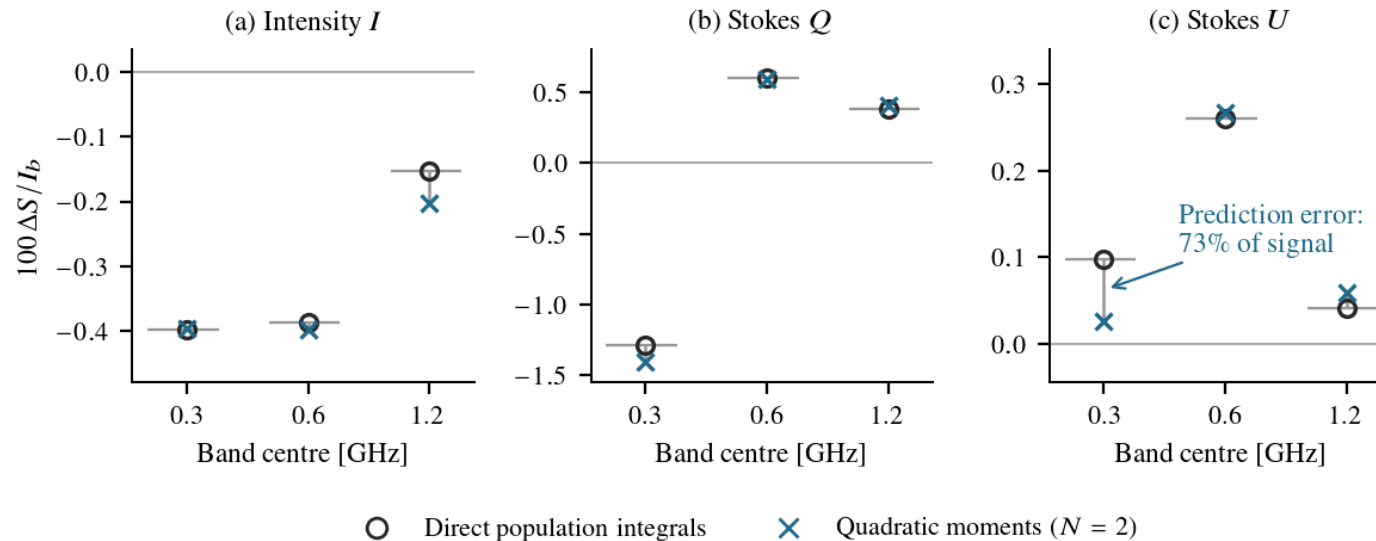
change from the linear reference: c_2 -1%, c_3 -3%, c_4 -25%

Two cumulants give one Gaussian answer for every screen; four can overshoot



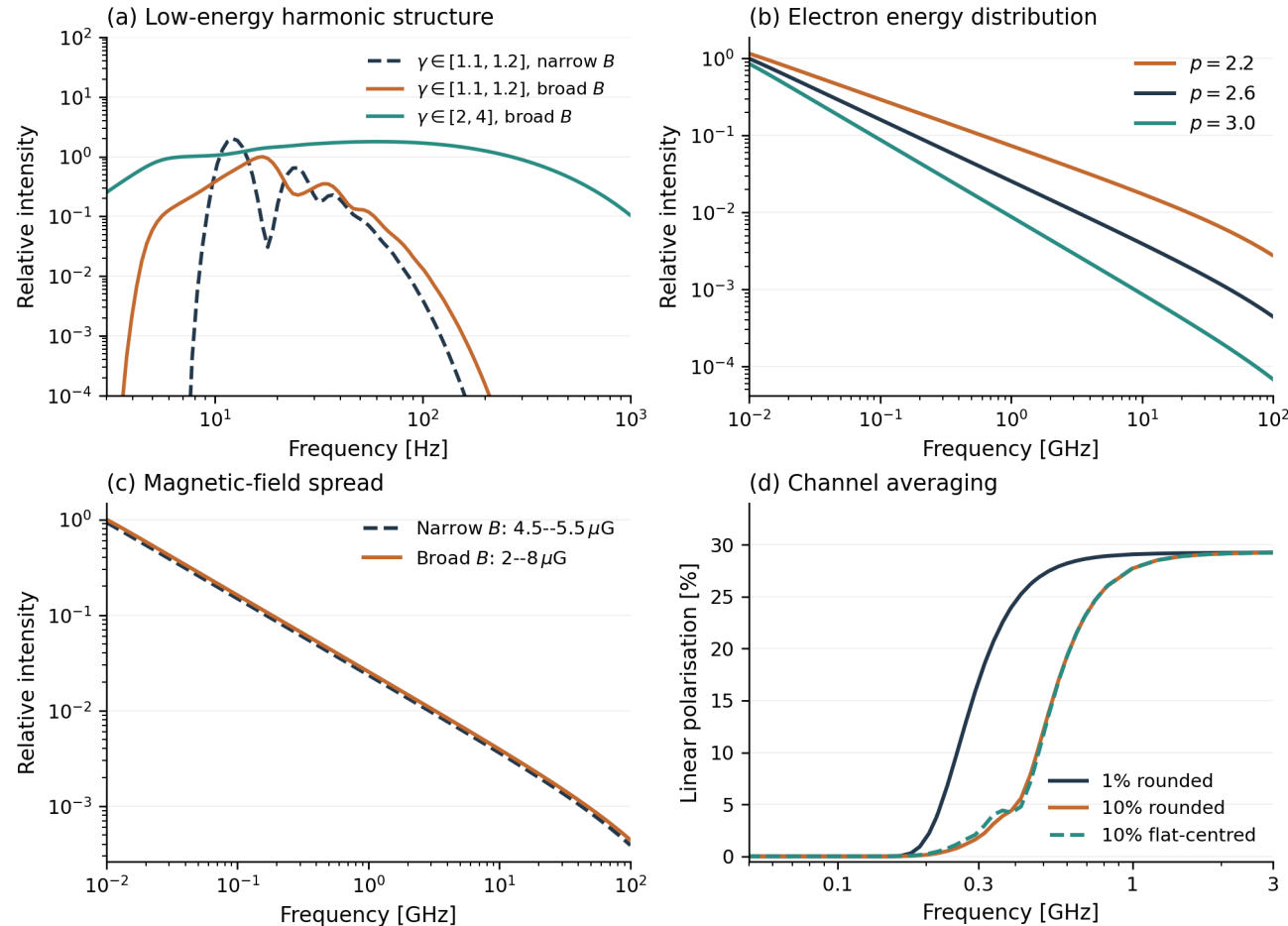
Same marginals, reversed correlations, different spectrum

Same marginals; reversed energy–field and depth–field correlations



Reversing the energy–field and depth–field pairings while keeping every one-dimensional marginal fixed changes I, Q, U by 0.2–1.5%. A model built from separate distributions would miss the change; the quadratic moments predict it, with a 73% error on the smallest U change at 0.3 GHz.

Energy index, field spread and channel width act on different parts of the spectrum



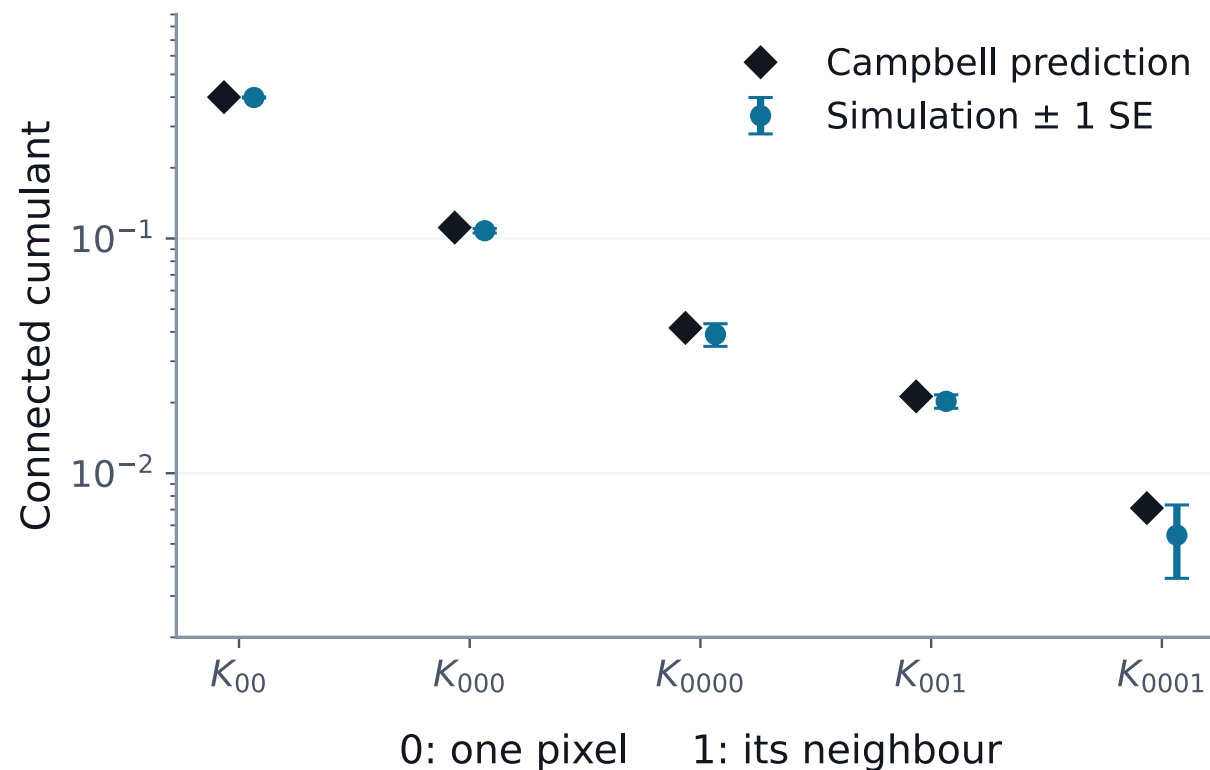
What the emission model assumes, and where each limit is assessed

Emission	helical orbit in a field uniform over the orbit; no radiation reaction; powers add incoherently; a supplied line-shape change is bounded only at fixed integrated power
Propagation	steady emission on one straight ray; cold-plasma Faraday rotation $\lambda^2\Psi$; absorption, scattering, refraction and conversion omitted (scalar absorption bounded if an optical depth is supplied)
Channels	smooth response on a declared finite band; shifting lines and rotation integrated across the channel before expansion
Population	correlated energy, field and depth allowed; all moments from one nonnegative population; fitted values need the same consistency check
Finite expansion	finite Legendre and Taylor orders about fixed references; conditional remainder bounds; population outside the declared ranges needs a separate estimate

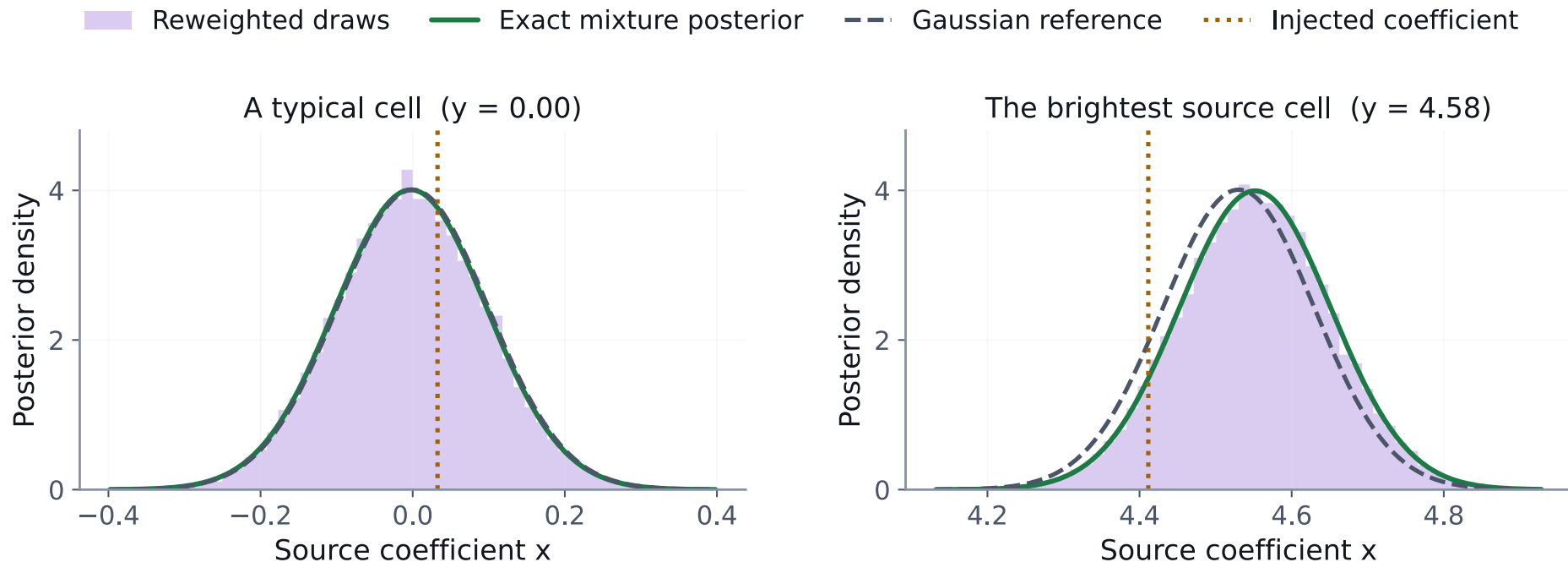
Known cumulants make the non-Gaussian test quantitative

$$K_{i_1 \dots i_r} = \lambda a^r \sum_j B_{i_1 j} \dots B_{i_r j}$$

For $r \geq 3$, Campbell's theorem: connected sky cumulants of the Poisson component; the Gaussian part contributes only at $r = 2$. $K_3 \propto \lambda^{-1/2}$, $K_4 \propto \lambda^{-1}$. Beam coupling predicts one-pixel and neighbouring-pixel statistics.



Same Gaussian draws, a non-Gaussian posterior



Each draw keeps its position; its weight $w_m(\lambda) = p_\lambda(x^{(m)}, d) / p_0(x^{(m)}, d)$ changes its contribution. Factorised weighting is valid here because the observation $d = B(x + \epsilon)$ with invertible B and source-coordinate white noise makes both reference and target factorise cell by cell; a general beam-correlated noise does not.

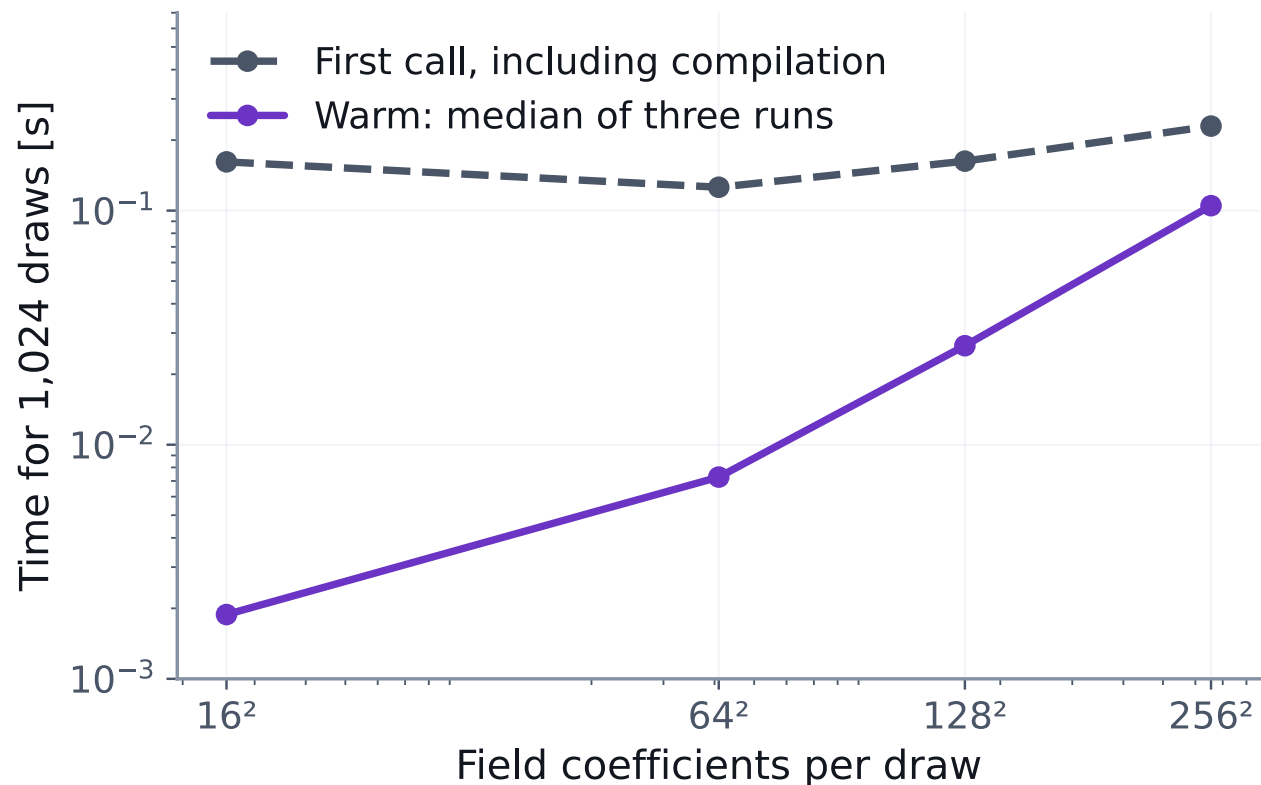
A tractable reference makes a large sample bank cheap

0.105 s

1,024 Gaussian field draws of 65,536 coefficients; warm JAX CPU, float32

$$p_0(x \mid y) = \mathcal{N}\left(\frac{y}{1.01}, \frac{0.01}{1.01} I\right)$$

closed-form diagonal GCR conditional in source coordinates, $y = B^{-1}d$



Is an exact block always better than NUTS?

Sampler	Budget	max pull
factor sweep	600 + 900	8.46
factor sweep	2400 + 3600	4.10
pure NUTS	600 + 900	1.05
pure NUTS	2400 + 3600	0.97

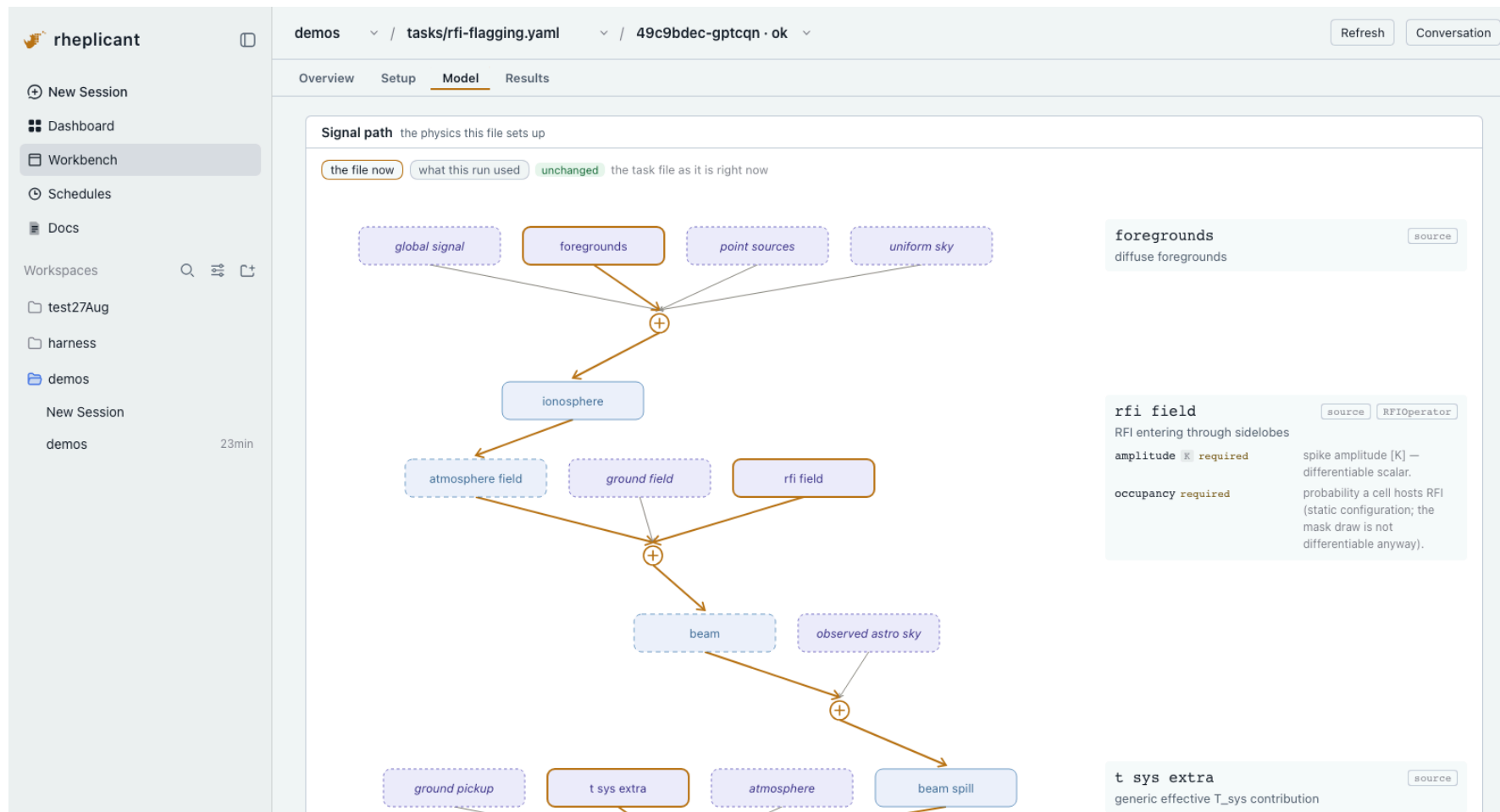
Worst replication of a seven-parameter hierarchical model; pull = (posterior mean – truth) / posterior SD. Each conditional draw in the sweep is exact, but alternation between strongly coupled blocks mixes slowly; NUTS with an adapted mass matrix moves along the coupling. The documented use of the sweep is an exact block too large for gradient sampling, such as a sky map or per-channel gains.

A gradient gives the response; identifiability is a rank test

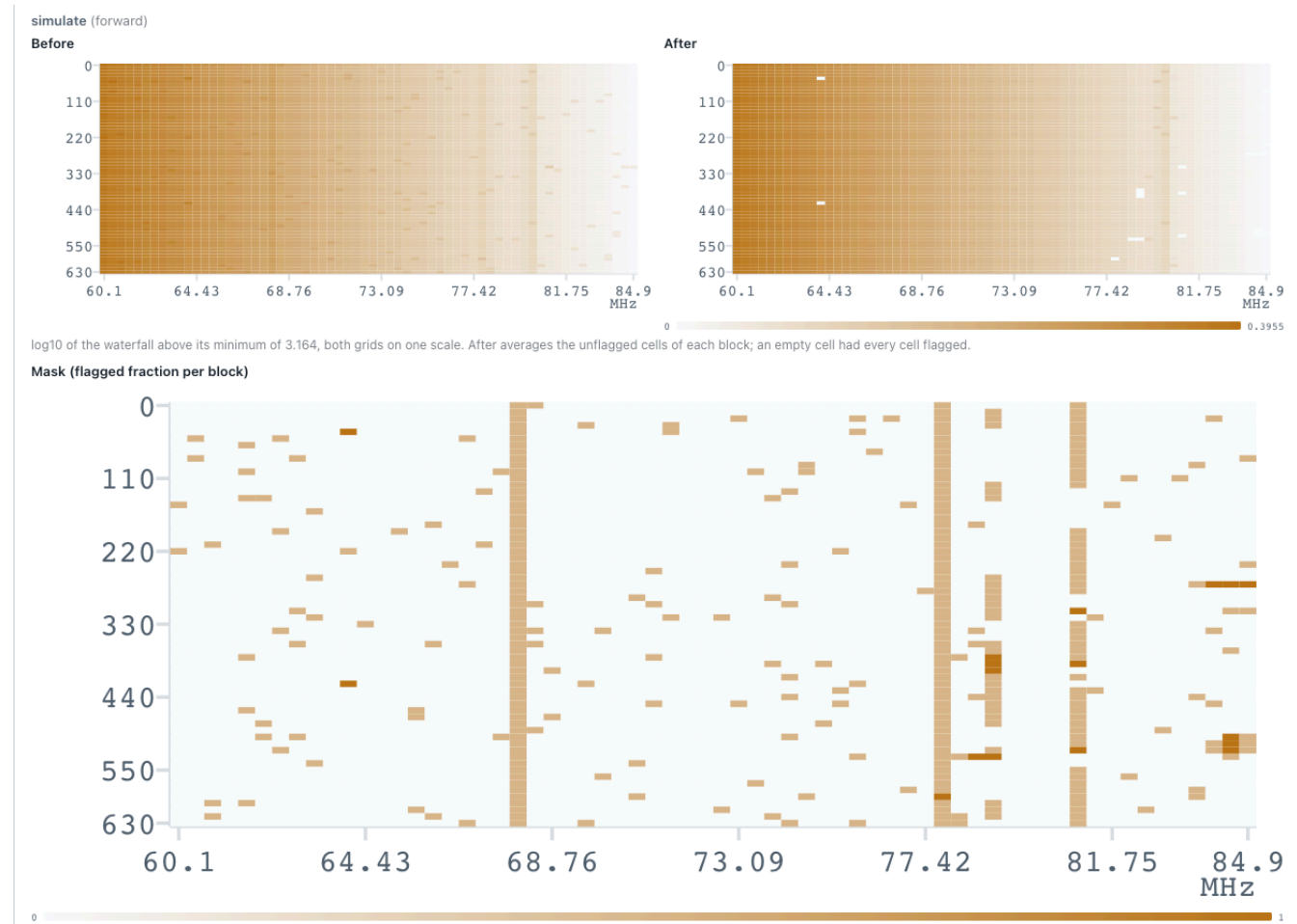
$$\text{rank} = \min(n_{\text{src}}, k) \times n_{\text{freq}}$$

One equation per switch position per channel; k free temperature families. Three sources make a three-family fit square and leave a four-family fit short by n_{freq} . Calibration on simulated data with 24 free noise-wave temperatures: Wiener mean by conjugate gradients, RMS errors 0.13 / 0.11 / 0.11 K on spreads of 40 / 60 / 8 K; condition number 40 with three sources, 4×10^6 with one, and the solver refuses. Adding two foreground parameters to the four-family twin gives nullity 2 of 34 and the plan refuses to run; three more loads and seven switch positions restore nullity 0.

The task file read as a signal path



Before and after flagging, on one colour scale



The chat that drove the four rounds

RFI flagging, round 3 of the session

continue

occupancy stable

kurtosis within

protected clean

broad applied

Flagged

4.57%

Residual excess kurtosis

0.5168

Parameters

sigma_threshold 4 · degree_freq/degree_time 6/3 · kernel_shapes 3x3, 1x9

Published as execution 20261007T231812Z-d12d4876-a75nsa of tasks/rfi-flagging.yaml; the run's grids are in its results folder.

RFI flagging, round 4 of the session

pass

occupancy stable

kurtosis within

protected clean

broad applied

Flagged

4.57%

Residual excess kurtosis

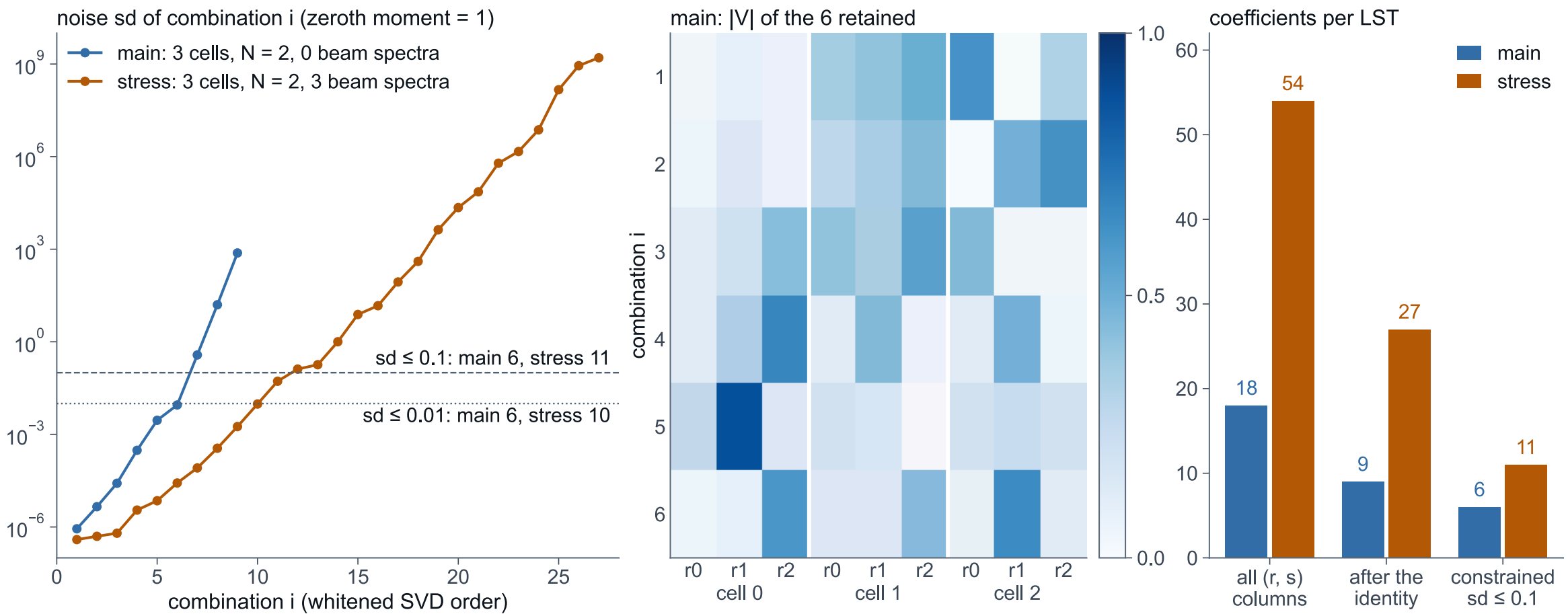
0.5168

Parameters

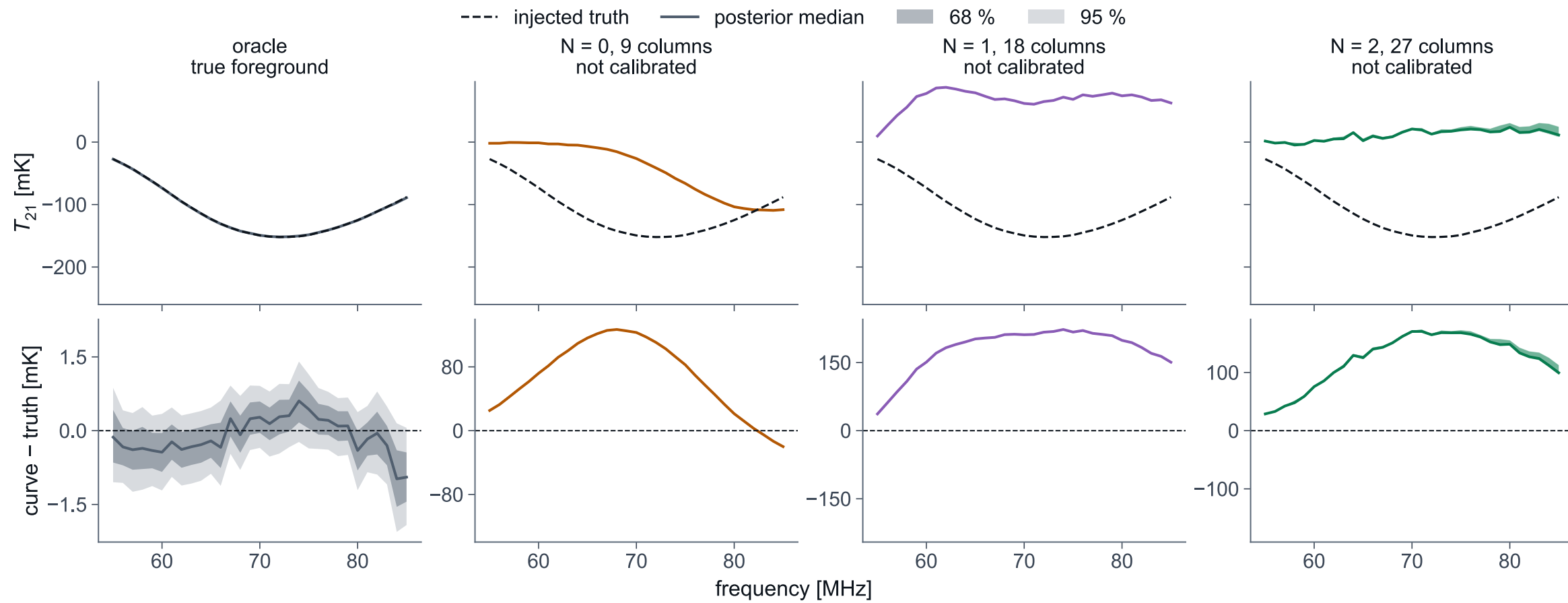
sigma_threshold 4 · degree_freq/degree_time 6/3 · kernel_shapes 3x3, 1x9

Published as execution 20261007T231816Z-49c9bdec-gptcqn of tasks/rfi-flagging.yaml; the run's grids are in its results folder.

Six constrained combinations per LST, out of eighteen columns

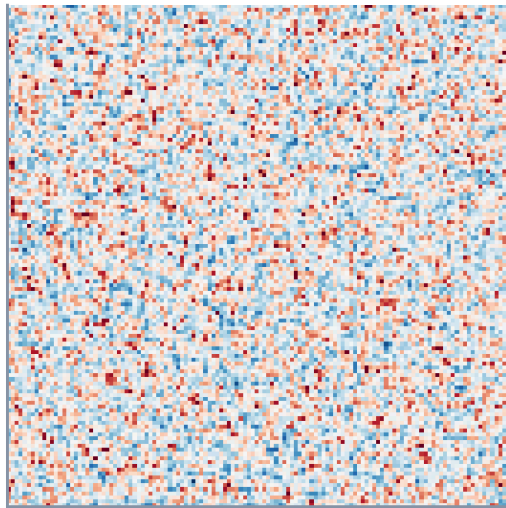


Where no per-LST spectral basis passes



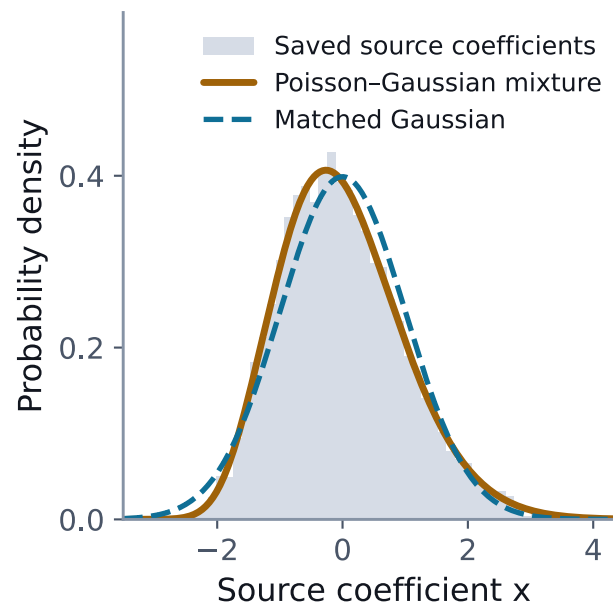
Same power spectrum, different source populations

Beam-smoothed Poisson sky



128 × 128 pixels

Same mean and variance

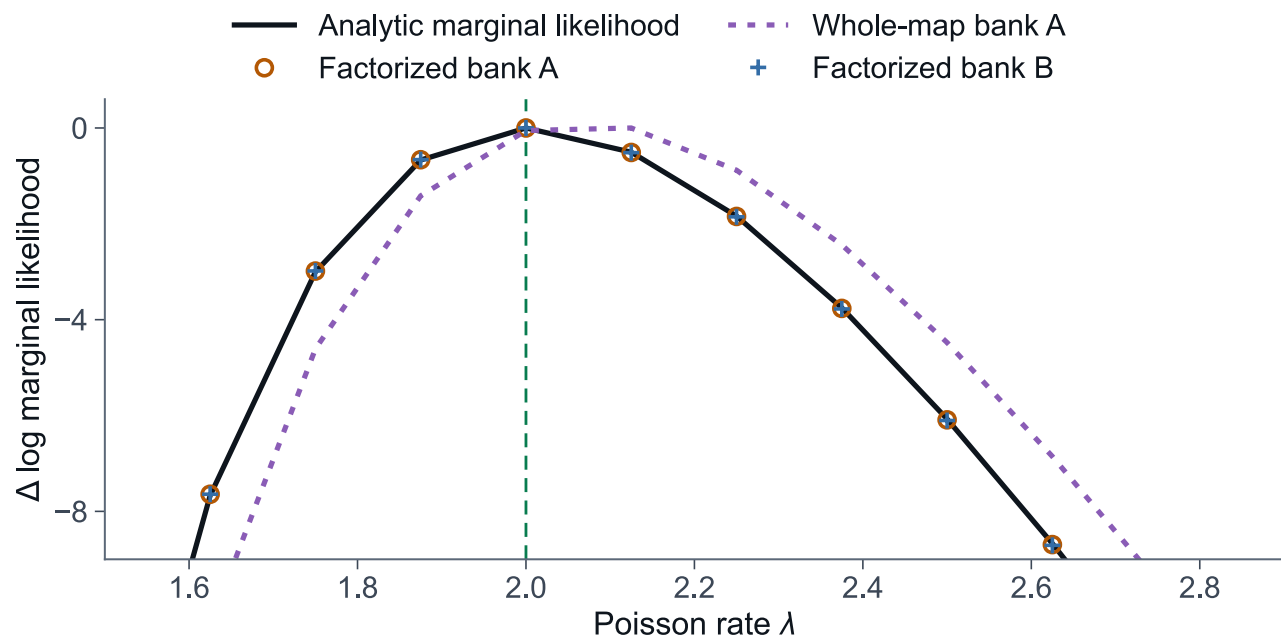


$$x_j = a(N_j - \lambda) + g_j, \quad N_j \sim \text{Pois}(\lambda)$$

With $a = \sqrt{0.8/\lambda}$, for every candidate λ :
 $\mathbb{E}[s] = 0$, $\text{Cov}(s) = BB^\top$. Rare bright sources or many faint ones: the covariance cannot tell. The **higher cumulants** can, and Campbell's theorem gives them in closed form.

Part one's shot-noise drive as a sky to be inferred: recover λ from a 16,384-coefficient map.

Recover the source rate, and check the weights



16,384 latent coefficients · rate MAP 2.0014 · analytic 2.0015 · injected $\lambda = 2$

$$w_m(\lambda) = \frac{p_\lambda(x^{(m)}, d)}{p_0(x^{(m)}, d)}$$

$$\frac{Z_\lambda}{Z_0} \simeq \frac{1}{M} \sum_m w_m(\lambda)$$

Draw the Gaussian reference bank **once**; every candidate λ only changes the weights. 1,024 draws of 65,536 coefficients cost 0.105 s.

Whole-map weights collapse: ESS 6.2 of 2,048. The exact source-coordinate independence of this model factorises the estimator; minimum local ESS 1,614 of 2,048, and two independent banks agree with the analytic curve.

The amplitude still rises with the cutoff; the observable is fixed

Held fixed in the number

Source	single plane, nominal $z_s = 5$
Path	Born; scalar-only, $\Phi = \Psi$
Driver	linear P , tree-level SPT bispectrum
Cutoff	vertex $\ell_{\text{max}} = 15360$
Readout	raw $\langle \kappa_1 \kappa_2 \rangle$ moment, linearized map

What moves it

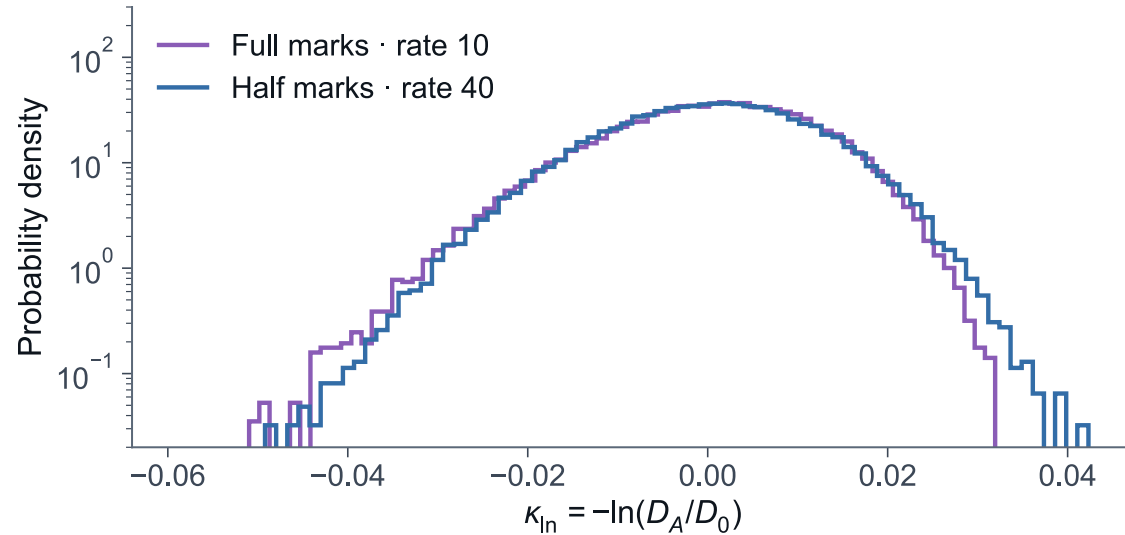
UV sensitivity. A collapsed configuration is a coincident-point moment. On the post-fix cutoff ladder the last doubling of ℓ_{max} still adds 17% at 0.5' and about 9–11% in the band at tree level, and 103% at 0.5' with BiHalofit.

No cutoff-independent nonlinear amplitude is quoted: slide 21's 5–10% is the $\ell_{\text{max}} = 15360$ value, still rising.

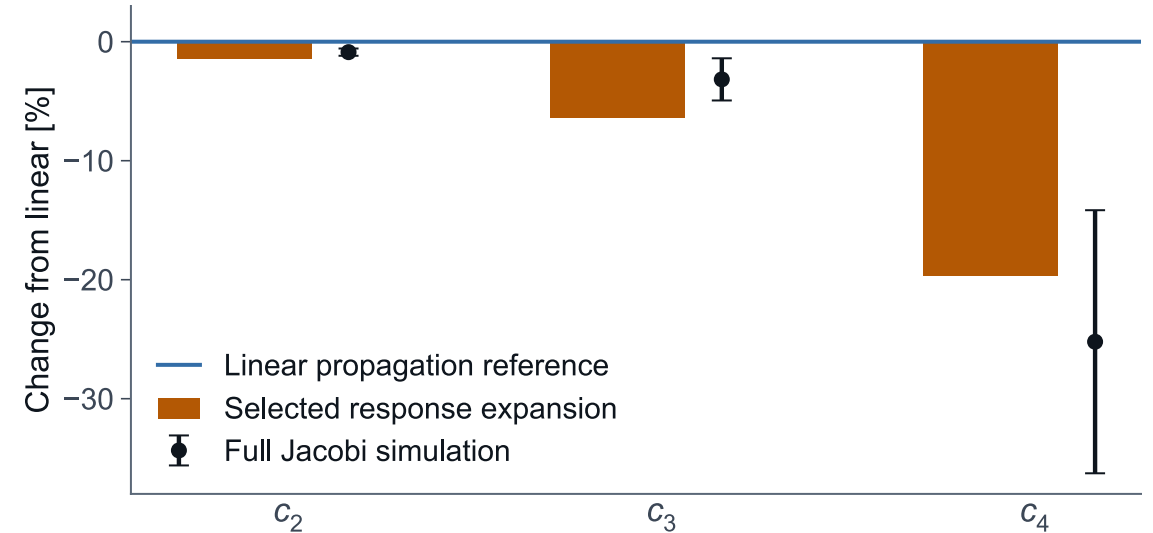
Spin signature. Linear scalar lensing is pure E in every realisation, so $\Delta C_\ell^{BB, FK} = 2 \text{Re} \langle B^{(1)} B^{(2)*} \rangle = 0$. FK changes E only; the first allowed B-mode is Gaussian FF at field order 4.

Readout. Jacobi κ_A , reduced shear and the source sample are external insertions still to be folded: the other stations of the complete view.

The one-point law reads what the covariance cannot



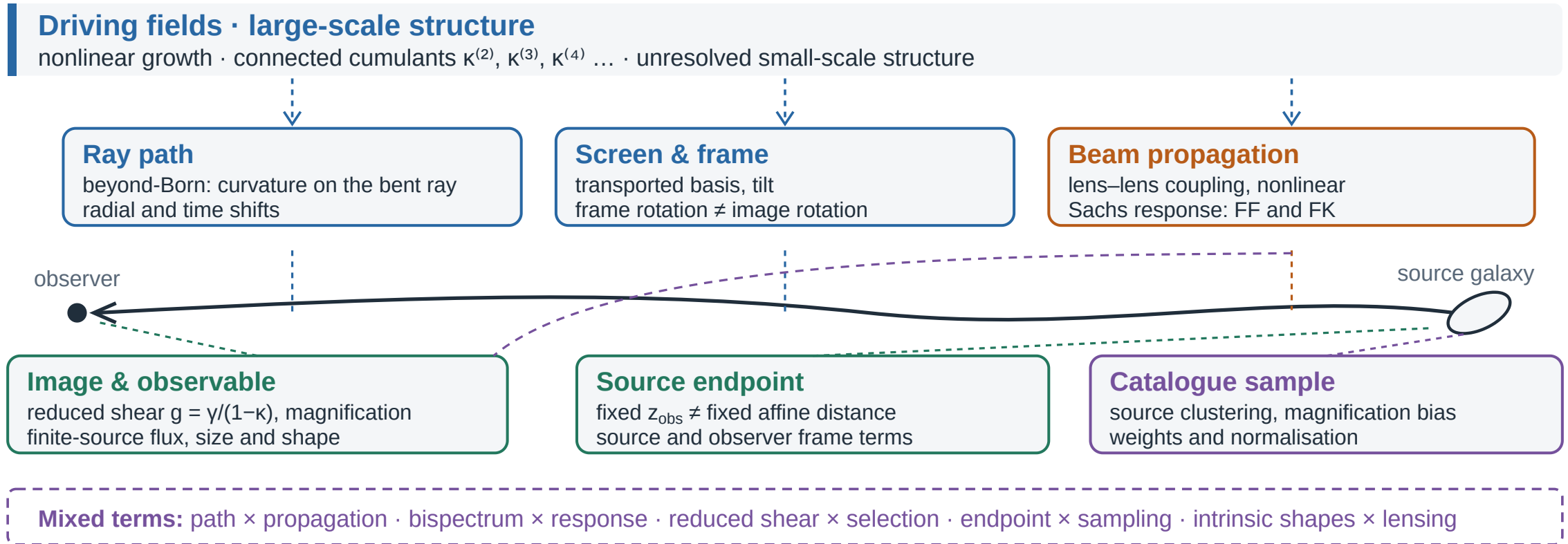
same driver covariance, half the marks $\times$ four times the rate:
skewness $0.37 \rightarrow 0.18$



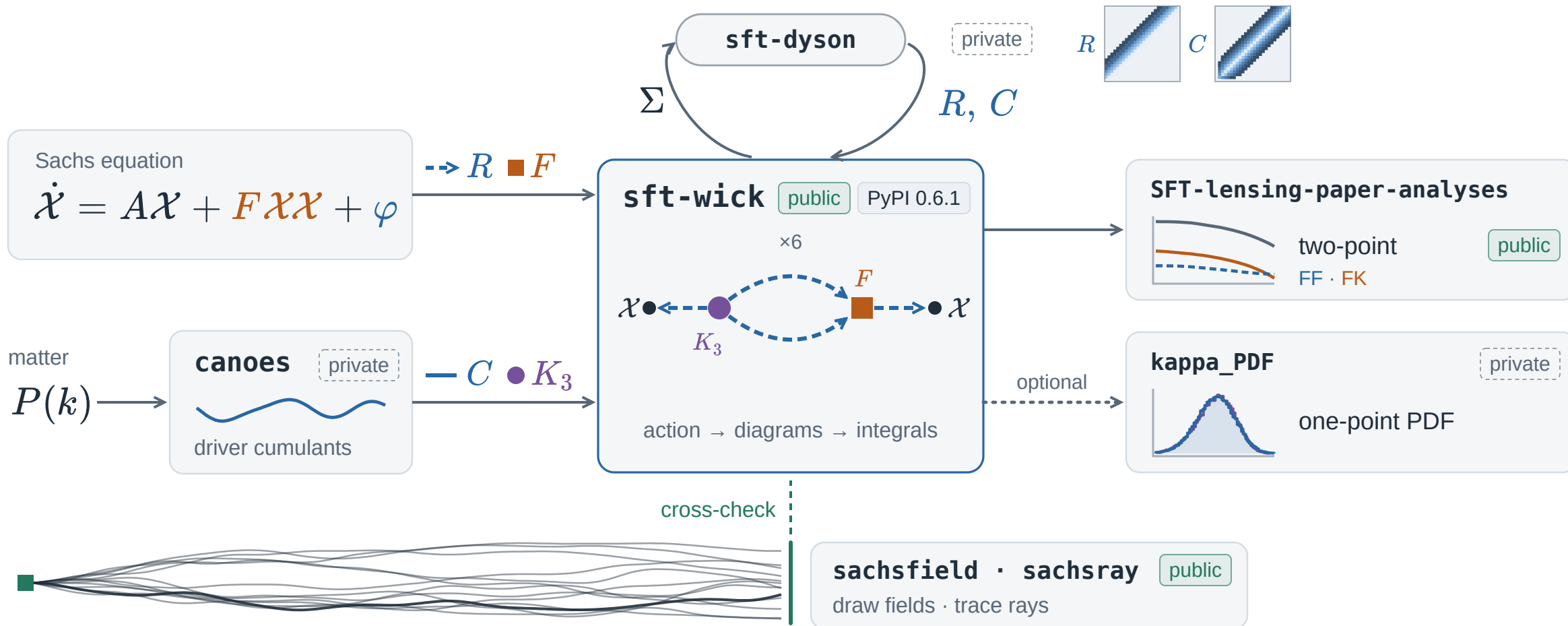
same driver, linear vs full Jacobi propagation:
 c_2 -1%, c_3 -3%, c_4 -25%

Same power spectrum, different laws: the skewness halves. Nonlinear propagation moves the fourth cumulant by a quarter while the variance moves 1%.

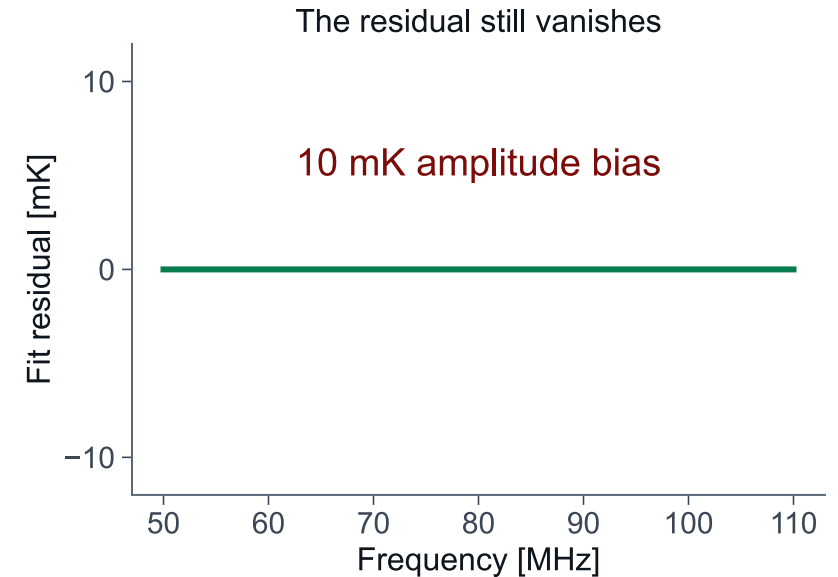
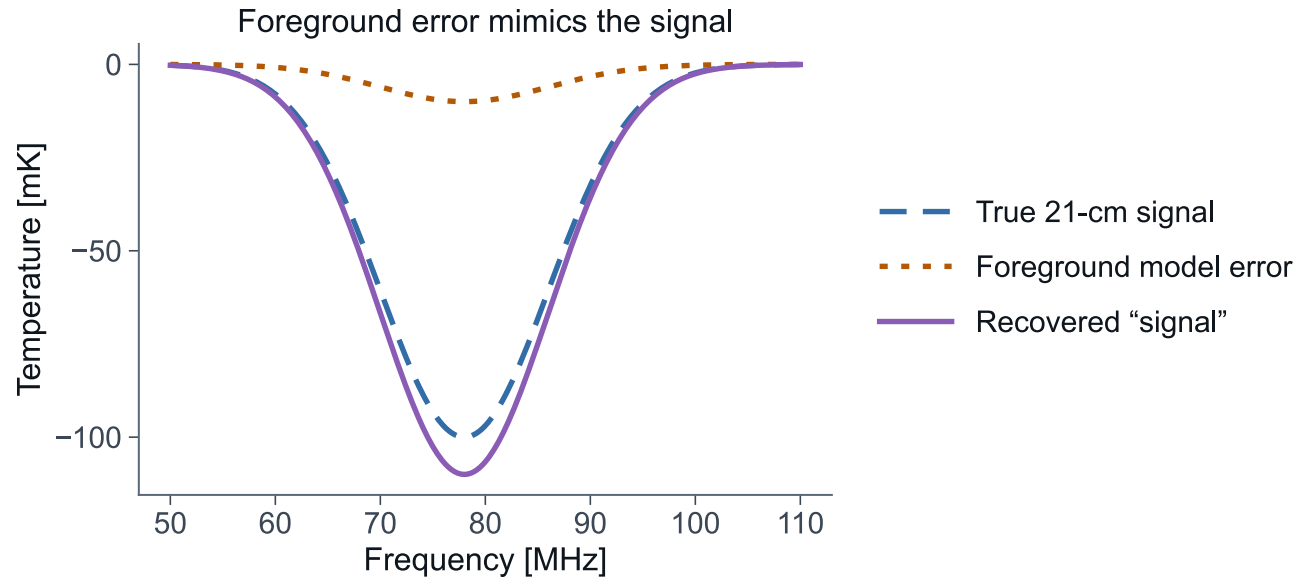
Corrections enter at every station along the ray



Each link of the calculation has its own package



Judge the model error in the space of the science parameter



10 mK → 0

signal-shaped foreground error:
10 mK amplitude bias, zero
residual

A finite-moment error is small or large only relative to the signal it mimics. Propagate it through the observing response and the estimator, not the raw residual.

Demo C's plan: a tolerance the arithmetic cannot reach

```

block 0 {fg_coeff}    GCR exact    structurally certified;
        numerical check 3 scales x 3 at-points (max 0.00e+00)
        kappa=1e+14 -> tol=1e-17 ... tol UNATTAINABLE at this
        dtype: the condition bound times the machine epsilon
        is 0.0222, already past the 0.001 target ... Run inside
        `with jax.enable_x64(True):` ... or strengthen the prior
block 1 {u}           NUTS
        outside the exact block ['fg_coeff']: sampled by NUTS,
        and the exact block is solved CONDITIONAL on them,
        once per sweep
execution: HMCgibbs(inner=NUTS, gibbs_sites=['fg_coeff'])

```

The graph: c , 96 LST $\times$ 9 foreground coefficients, Gaussian prior; u , 7 box-mapped 21-cm parameters through an emulator; waterfall $\sim \mathcal{N}(c \Phi^T + T_{21}(u), 10 \text{ mK})$. **Nothing was declared:** the affine block was found in the traced program, then checked numerically.

Gaussian conditional

solve, draw (GCR) or integrate out analytically; the plan also says when the arithmetic cannot deliver the tolerance

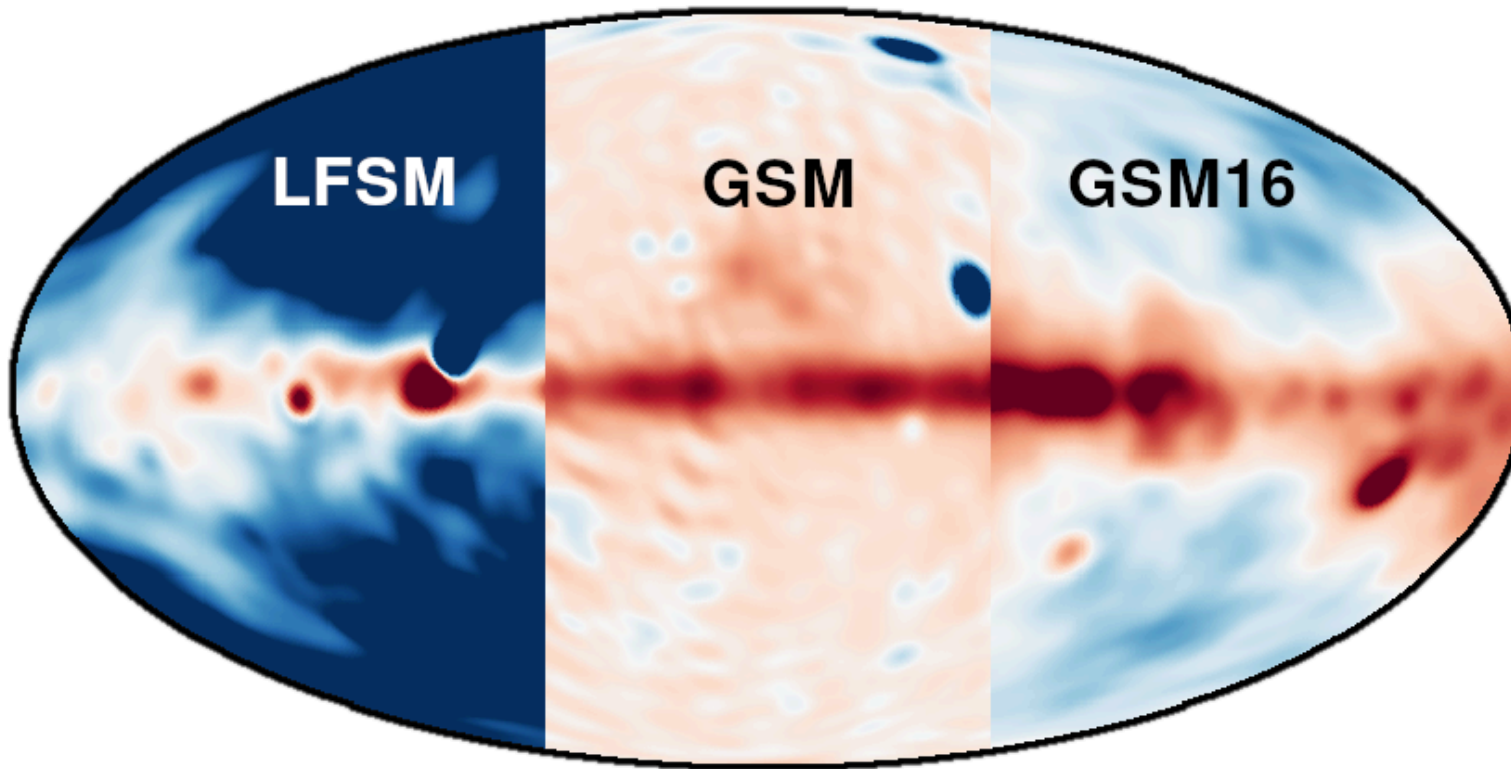
Useful local structure

log-space, iterative GLS or Gauss–Newton proposal, corrected to the full target by Metropolis–Hastings

General dependence

NUTS; a gradient route returns a convergence *certificate* (Newton decrement) or refuses with a reason

Sky models disagree at the 10 % level



LFSM · GSM · GSM16, one sky, three models

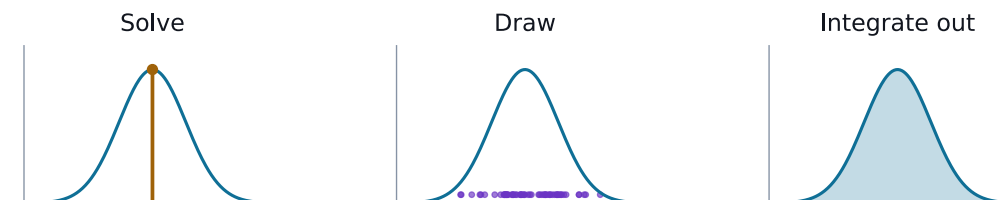
> 10 %

disagreement between the models, as quoted in the RHINO overview

A global 21-cm model has to be good to about one part in a hundred thousand for the model error to stay below the signal.

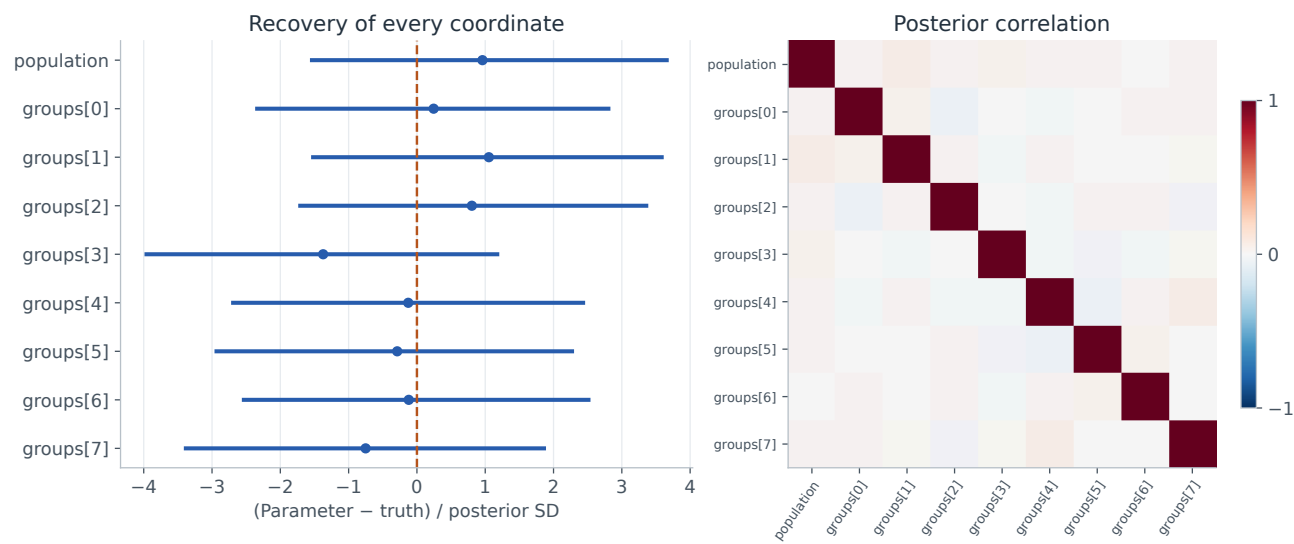
Five routes a block can take

method	printed as	meaning
gcr	GCR exact	exact conditional Gaussian draw; the noise does not move with the block
gcr+snis	GCR + SNIS	whole graph exact, noise moves with it: GLS fixed point, corrected by self-normalised importance sampling
gcr+mh	GCR + iterative GLS + MH	exact block beside a sampled remainder with moving noise: the exact draw is an MH independence proposal
log-gcr	log-gcr	the exact draw on a log-transformed block whose prediction is exponential in an affine function
nuts	NUTS	NumPyro's No-U-Turn sampler



One Gaussian conditional supports three operations: solve, draw, integrate out. Which structure takes which route: linear-Gaussian with an eligible covariance, a supported log-linear form, conditionally Gaussian affine groups inside a chain, an eliminable exact block with a sampled residual, a general differentiable model.

Example 03: chains and provenance



Hierarchical-model recovery · recorded all-NUTS run

chains × draws	2 × 2,000 after 1,000 warmup, seed 0
divergences	0
population R-hat	1.000004
population ESS	8,817
population mean	1.008 ± 0.217, truth 0.8
run record	plan, fingerprints, seed, dtype, devices, backend, budget, termination, approximation class
refusal	failed premise, grounds, scope, remedies, optional fallback

RHINO's open-source pipeline

rhino-daq

PFB spectrometer and Arduino control driver

momentRFI

RFI flagger: iterative clipping and 2-D surface fitting

rhino-cal

noise-wave fitting, Kirkham+ method

rhino-monitor

live satellite, weather and plane monitor

rhino-sparams

lab measurements

rhino-beams

beam simulations

husky

Bayesian sky model, simple foreground fits

limTOD

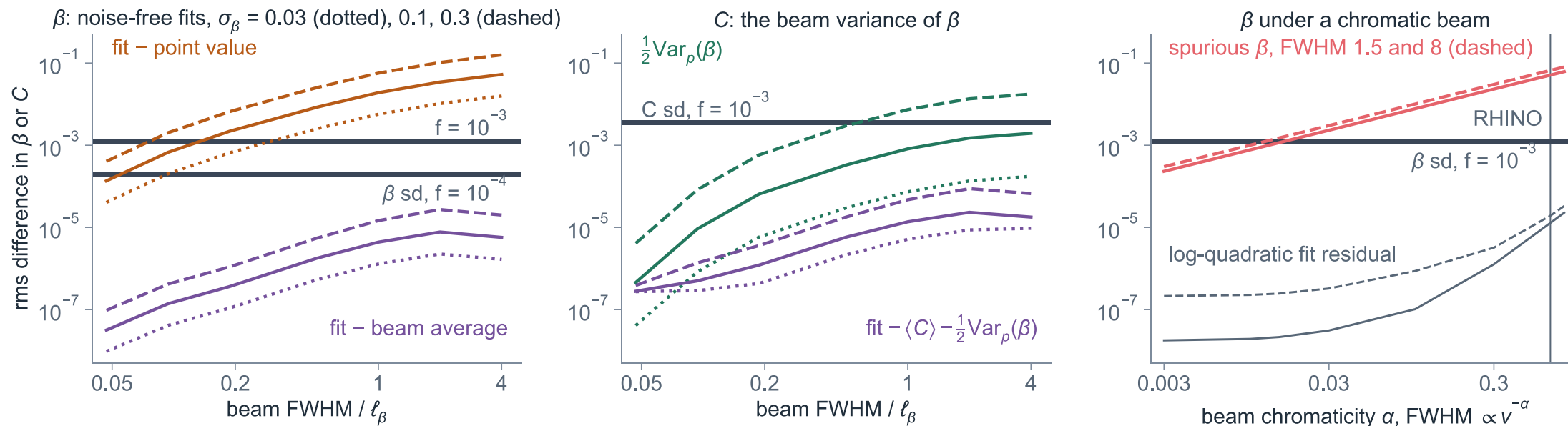
Bayesian map-maker

rheplicant

digital twin, live simulation; inference on bayesmith

A log fit through a beam returns brightness-weighted moments

$$\ln d_p = \ln \tilde{T}_p - \langle \beta \rangle_p L + \left(\langle C \rangle_p + \frac{1}{2} \text{Var}_p \beta \right) L^2 - \left(\frac{1}{6} K_{3,p}(\beta) + \text{Cov}_p(\beta, C) \right) L^3 + O(L^4)$$

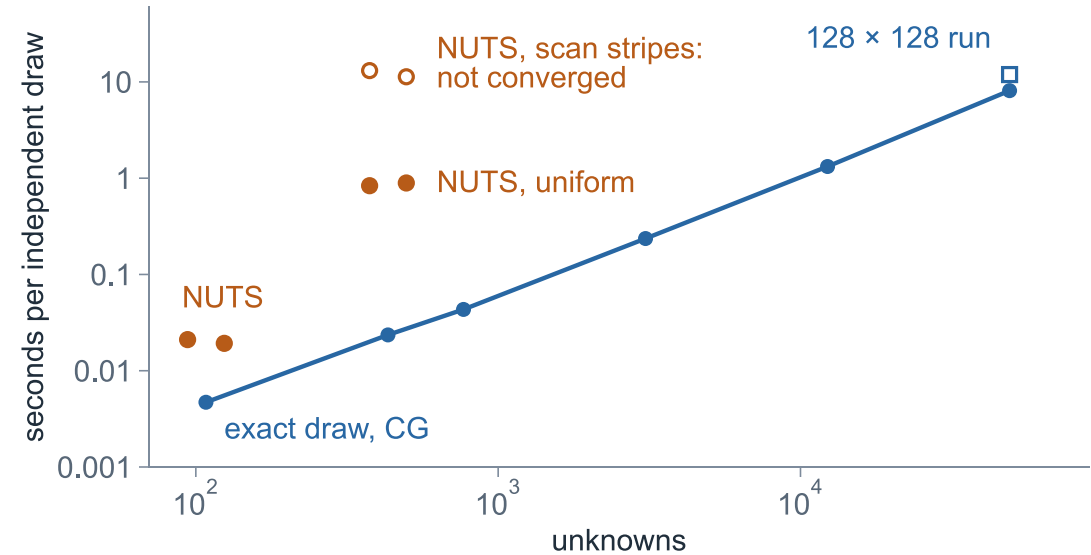


moments weighted by $B_{pq} T_q(70 \text{ MHz})$ · noise-free fits of the simulator sky through Gaussian beams · sd bands from the beam-smoothed prior

Fit – moment: 2.7×10^{-5} (β) and 8.8×10^{-5} (C) rms; $\frac{1}{2} \text{Var} \geq 0$; fit residual $\leq 8.7 \times 10^{-6}$; cubic term against its prediction: correlation 0.9996. A pivot at 60 MHz moves β by 0.022 rms. Chromatic beam: spurious β grows by 0.0077 rms per 0.1 in α . Smoothing every channel to the widest beam restores the moments without noise; with noise it is untested.

The block is exact; converged NUTS means differ by ≤ 0.15 sd

check	unknowns	result
Fourier modes, 1 and 4 visits	12,288	sd ratio 0.988–1.002
dense oracle	768	mean $\leq 4.8 \times 10^{-5}$ sd; sd ratio 0.990, 1.003
dense Cholesky, t eliminated	12,288	mean $\leq 3.8 \times 10^{-5}$ sd; sd ratio 0.995, 0.998
NUTS, $f = 0.001, 0.03, 0.0599$	108	two chains: means within 0.10, 0.14, 0.15 sd; sd ratio 0.976–1.001
log step, $f = 0.06, 0.07$	108	refused (fractional_too_large)
NUTS, uniform	432	within 0.12 sd; R-hat ≤ 1.01 ; 0.84–0.89 s per effective draw
NUTS, scan stripes	432	not converged: R-hat up to 1.092 (β); ESS down to 13.6 (C)
128 × 128, end to end	49,152	11.9 s per draw; z rms 1.08, 1.04



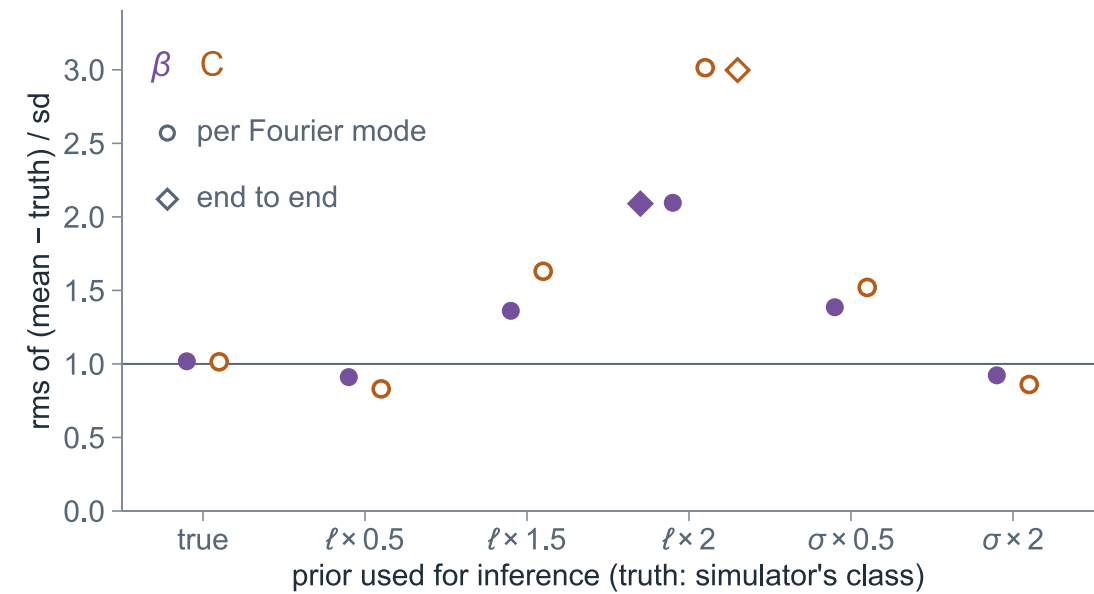
beam-smoothed prior · NUTS: dense mass, 1000 + 1000 steps
NUTS: two chains per size, drawn at $\div 1.15$ and $\times 1.15$ · square: end to end

Certified CG runs: relative error $\leq 5.0 \times 10^{-4}$. Operator records at $f = 10^{-3}$ – 0.0599 : $\kappa \leq 6.81 \times 10^7$, ≤ 0.007 sd per pixel up to 12,288 unknowns; at 49,152, $\kappa \leq 6.96 \times 10^7$, 0.0115–0.0127 sd. Bounds analytic, checked by Lanczos. Calibration and chromatic runs: mean relative error only. At $f = 3 \times 10^{-4}$: $\kappa \leq 7.21 \times 10^8$, no pixel bound recorded. $f \leq 10^{-4}$: not certified.

Curvature leans on the prior; an over-smooth prior over-claims

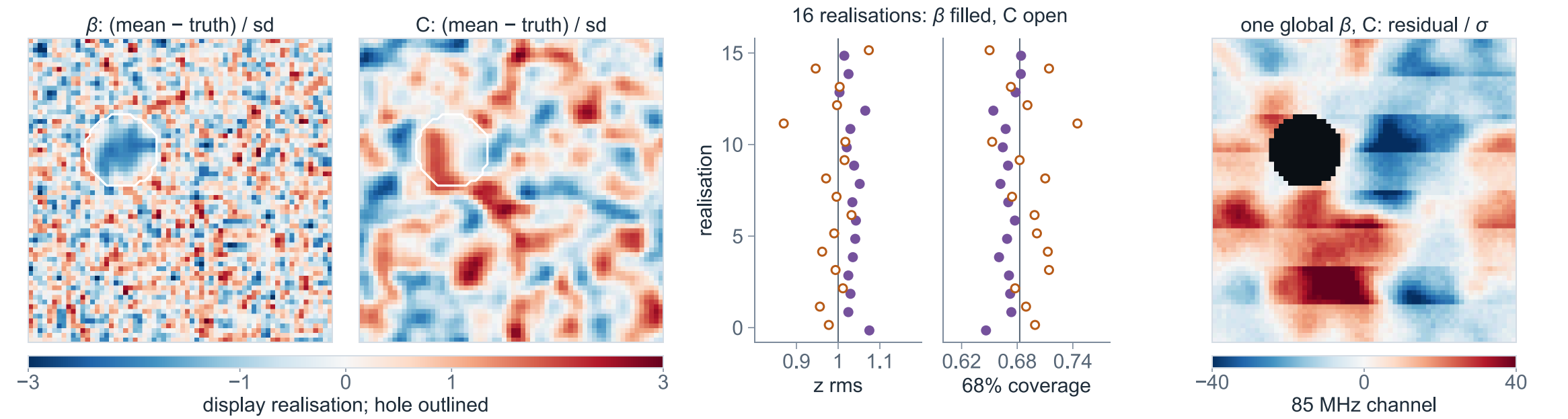
per-pixel sd, one visit	β	C
data alone, one pointing	2.6×10^{-3}	0.025
joint posterior (median)	1.2×10^{-3}	0.0036
prior	68% coverage, β / C	
beam-smoothed (used)	67% / 68%	
point field	74% / 71%	

Correlation length 2× too long: z rms 2.10 (β) and 3.01 (C) per Fourier mode, 2.09 and 3.00 end to end. Not done: inferring the correlation length and amplitude.



beam-smoothed prior class · $f = 10^{-3}$ · dots: per Fourier mode, uniform one visit · diamonds: end to end, stripes and hole

Errors against beam averages match the sd in 16 realisations



beam-smoothed prior · 16 realisations redraw the sky and the noise; posterior sd from the display realisation

	z rms (median), β / C	68% coverage (mean), β / C
observed pointings	1.03 / 0.99	67% / 69%
hole	0.88 / 0.88	66% / 71%

Infer the beam and the sky spectrum together

Synthetic ring drift scan · 96 sky pixels · 32 LST bins · 16 channels · 60–85 MHz

$$T(\phi, \nu) = A T_0(\phi) (\nu/70)^{-\beta}$$

$$w(\nu) = w_{70} (\nu/70)^{-0.8}$$

Known sky template, Gaussian beam.

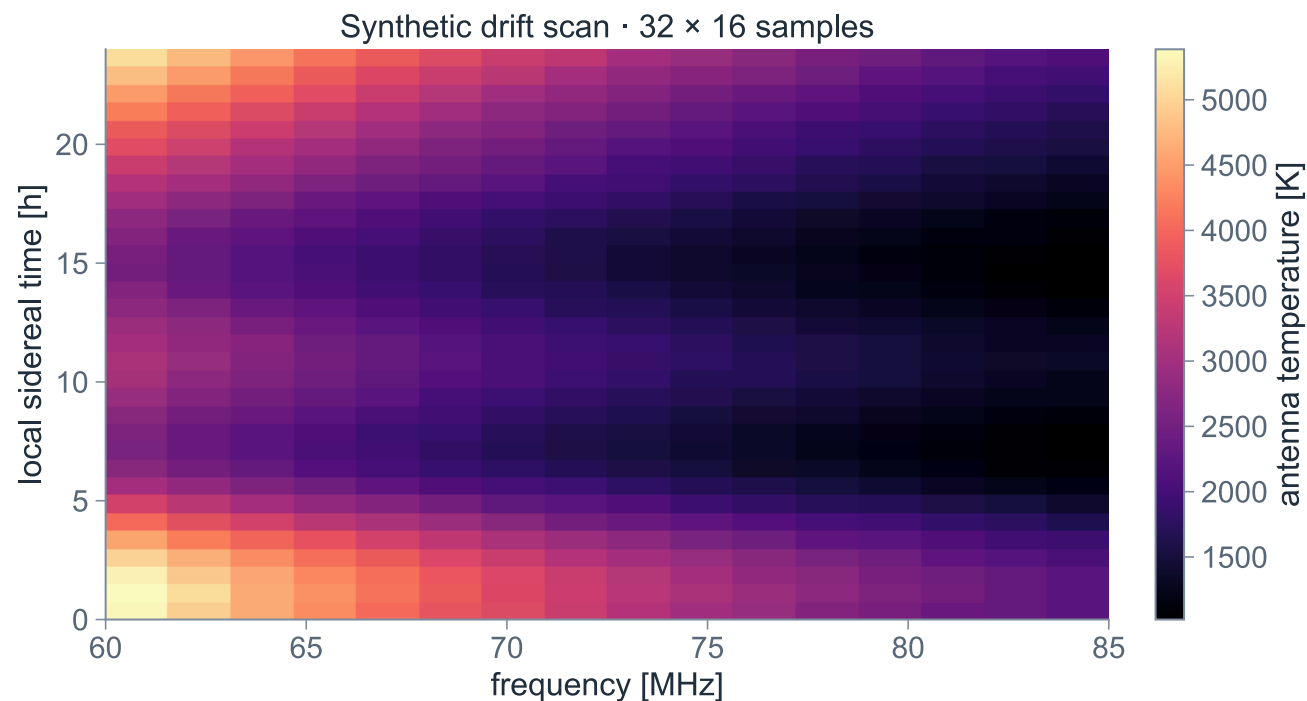
Infer the beam FWHM w_{70} and β .

One common $A \sim \mathcal{N}(1, 0.2^2)$ integrated out exactly.

Inject $w_{70} = 40^\circ$, $\beta = 2.55$, $A = 1$.

20 K independent measurement noise.

One forward model: sky maps $\rightarrow$ chromatic beam $\rightarrow$ time–frequency data.



The full drift scan constrains both unknowns

Same synthetic data and forward model · numbers: 801×801 grid, contours: 401×401 · shared amplitude integrated out

$$w_{70} = (40.25 \pm 0.22)^\circ$$

$$\beta = 2.5483 \pm 0.0035$$

Posterior means $\pm$ standard deviations.

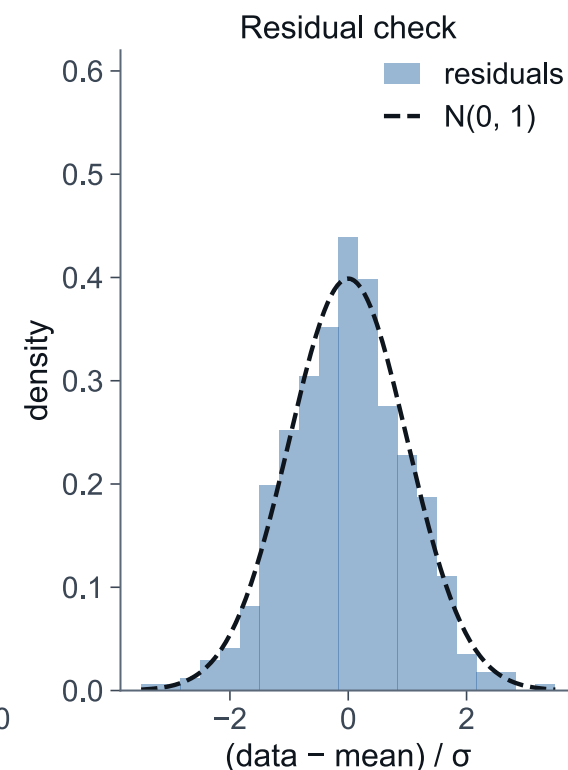
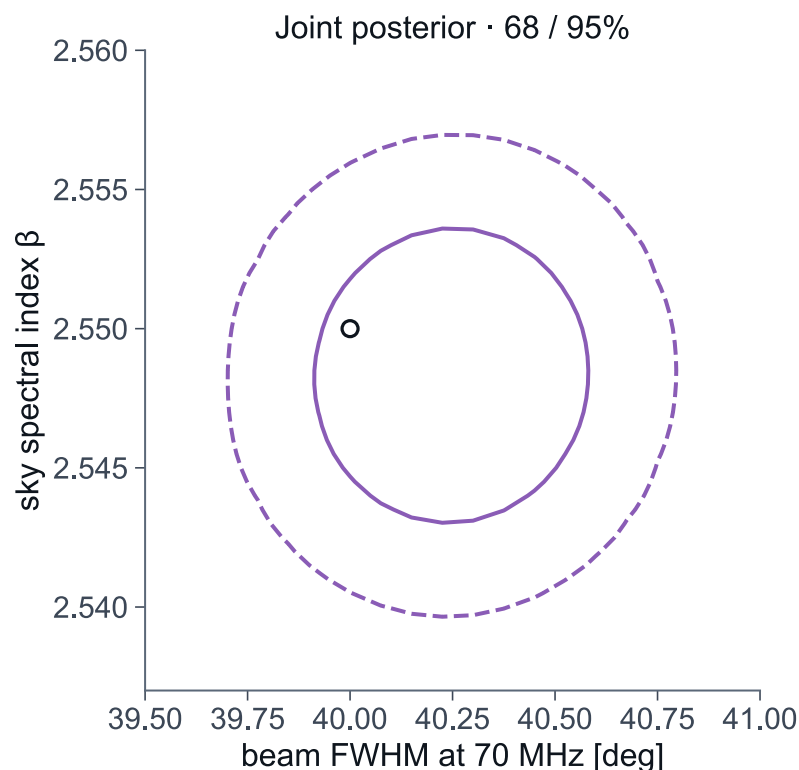
Ring: injected values; contours: 68% solid, 95% dashed.

Joint correlation $\rho = 0.020$.

Residual rms / noise = 0.991.

Check the entire posterior, then the predicted observation.

$401 \rightarrow 801$ grid: stable means and intervals.



Time dependence separates beam width from spectral slope

Controlled comparison · same truth, noise law, priors and common amplitude

One LST at 0 h

FWHM sd 8.6° , as in the 25° – 55° prior.

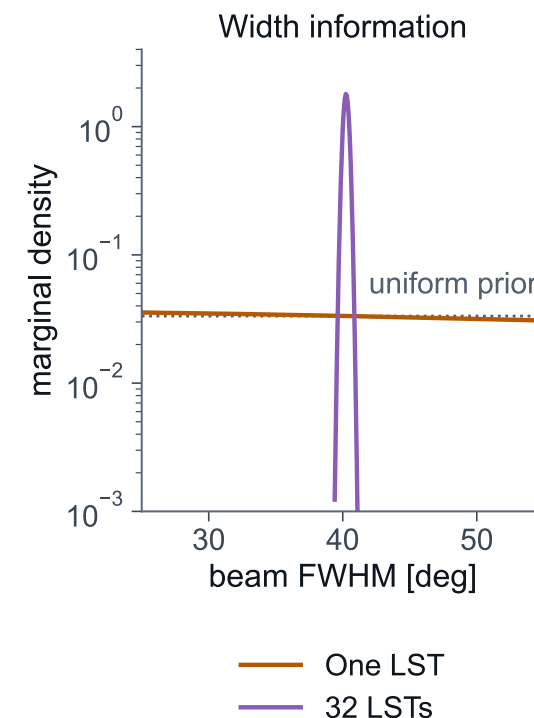
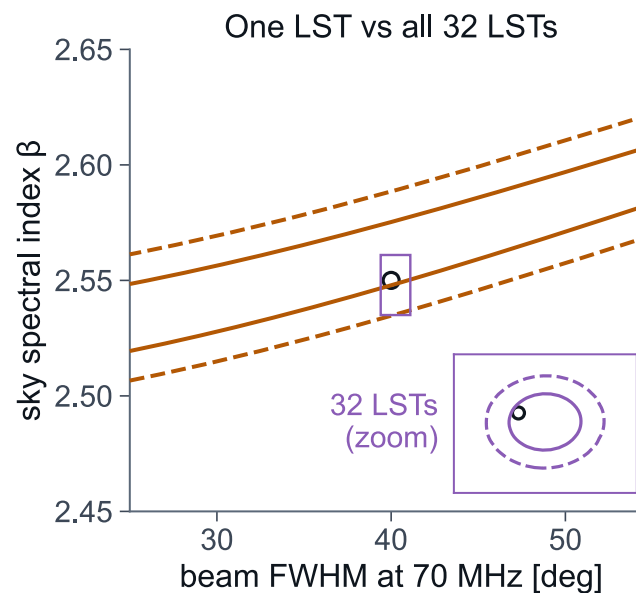
No FWHM information; β follows FWHM along a ridge.

All 32 LSTs

FWHM sd 0.22° ; β sd 0.0035 .

The sky rotates through the beam.

Keep the angular modulation in the data: it constrains the beam while the spectrum constrains β .



Gaussian structure makes inference tractable

bayesmith · exploit Gaussian blocks; correct to the full target

Solve, draw, integrate

GLS, constrained realisations
Exact draws in log coordinates
Analytic elimination

Build a Gaussian proposal

Iterative GLS
Bias-corrected log-linearisation
Gauss–Newton linearisation

Recover the full target

Metropolis–Hastings correction
Importance weights
NUTS for remaining blocks

A non-Gaussian posterior can still be reached with Gaussian proposals

Draw Gaussian fields, then weigh the non-Gaussian law

Observation: $d = Rx + \epsilon$, Gaussian noise; population law $p_\eta(x)$

1 · One Gaussian reference posterior

$x_i \sim q(x \mid d) \propto p(d \mid x) p_0(x)$ · fast GCR draws

2 · Evaluate the complete field's density ratio

$w_i(\eta) = p_\eta(x_i) / p_0(x_i)$

3 · Infer the population, then draw the field

$p(\eta \mid d) \propto p(\eta) \frac{1}{N} \sum_i w_i(\eta)$

Draw η ; choose a whole x_i with its weights at that η .

Fixed likelihood here. Check weight ESS and independent banks; the finite bank is an approximation.

Infer a bimodal population with Gaussian field draws

$$d_j = x_j + \epsilon_j, \quad \sigma_\epsilon = 0.35$$

$$p(x_j \mid a) = (1 - a) \mathcal{N}(-1, 0.6^2) + a \mathcal{N}(1, 0.6^2)$$

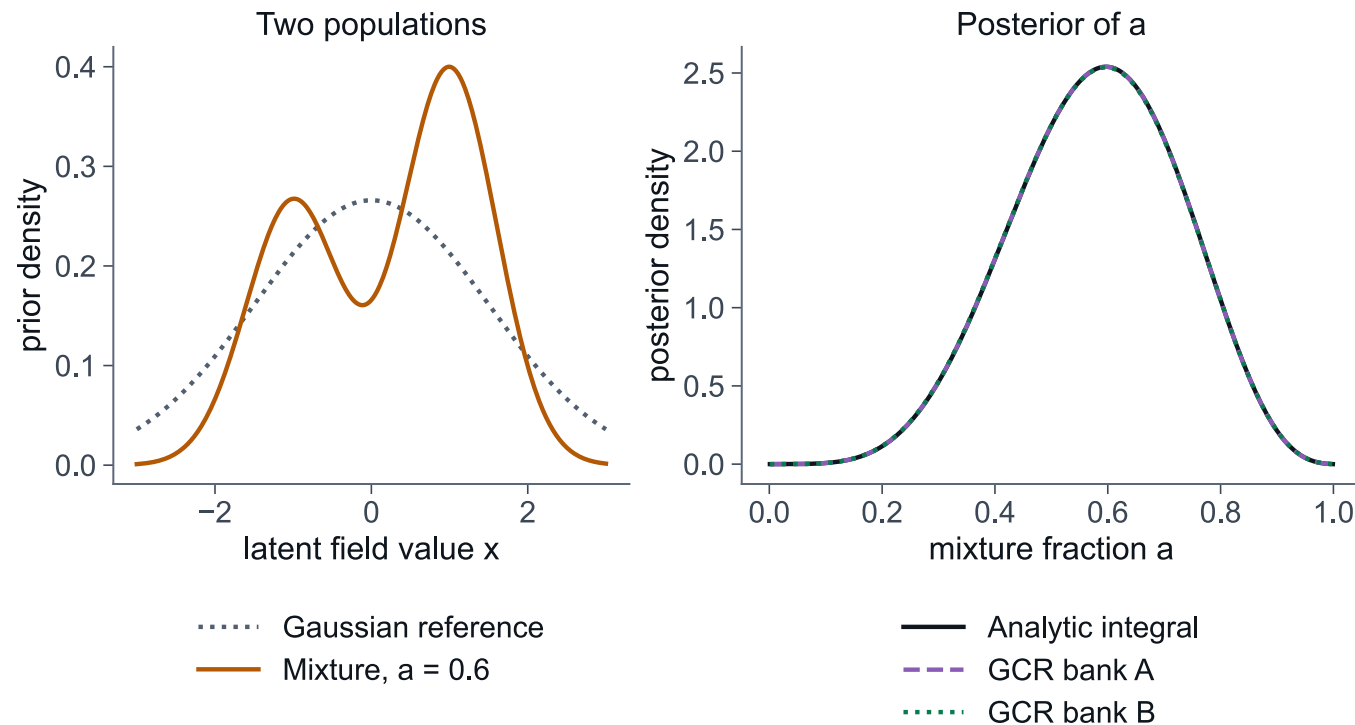
Eight observed objects. Infer the mixture fraction $a \sim \text{Beta}(2, 2)$.

Reference: $p_0(x_j) = \mathcal{N}(0, 1.5^2)$.

Two banks × 12,000 whole-field draws.

a : 95% interval [0.28, 0.85]

min weight ESS 5,387 of 12,000



The Gaussian covariance stays fixed; infer the skewness

Synthetic 12×12 grid · prior $\lambda \sim U(0.5, 8)$

$$n_j \sim \text{Poisson}(\lambda)$$

$$x_j = \sqrt{0.8/\lambda}(n_j - \lambda) + g_j$$

$g_j \sim \mathcal{N}(0, 0.2)$; observe $d = B(x + \epsilon)$.

Known invertible toy beam; ϵ white before it.

$K_1 = 0$, $K_2 = 1$ at every λ .

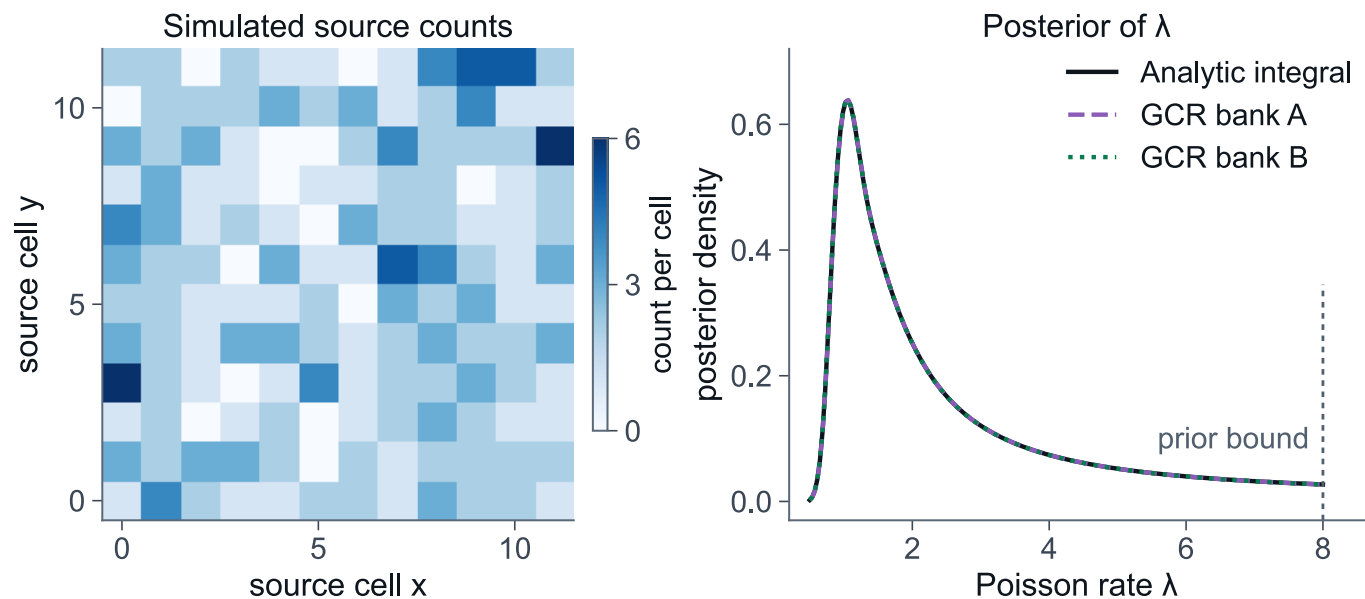
$K_3 = 0.8^{3/2}/\sqrt{\lambda}$, $K_4 = 0.8^2/\lambda$.

$\lambda > 0.78$ at 97.5%, so $K_3 < 0.81$

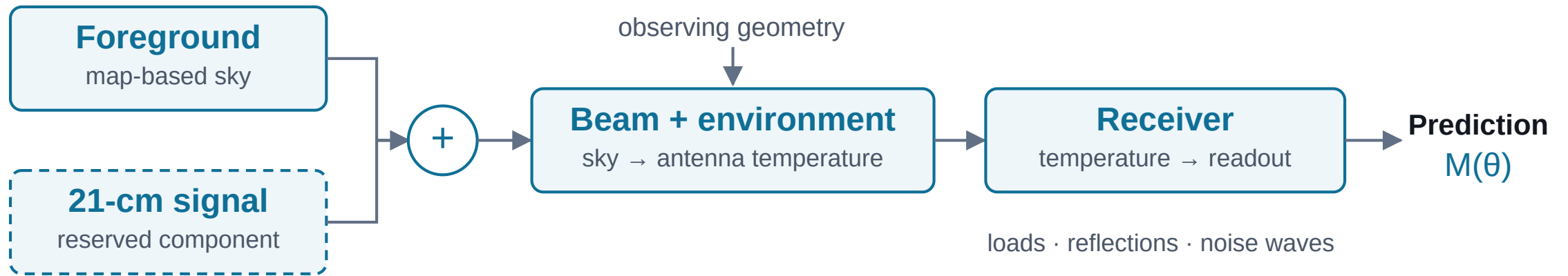
Upper end: the prior bound $\lambda \leq 8$

Injected $\lambda = 2$

min weight ESS 2,730 of 8,192



One model graph holds a sky and an instrument



Forward simulation

Draw parameters → generate a measurement.

Bayesian inference

Condition on the measurement → infer parameters. Same arrows, same physical calculation.

Logs turn a curved power-law sky into one Gaussian block

Synthetic periodic patch · 64×64 pointings · 16 channels, 60–85 MHz · narrow achromatic beam, not RHINO's

Simulated with rheplicant: 256^2 sky

$$\ln T_q = t_q - \beta_q L + C_q L^2$$

$$d_p = \sum_q B_{pq} T_q (1 + f_p w)$$

$L = \ln(\nu/70 \text{ MHz})$; beam 1.5 pointings;

$f_p = 10^{-3}$ for one visit, 5×10^{-4} for four;
one hole.

Inferred: one spectrum per pointing

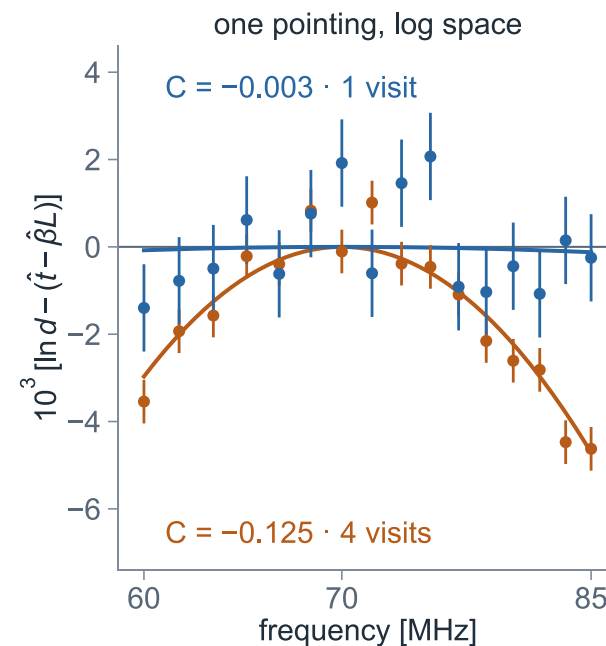
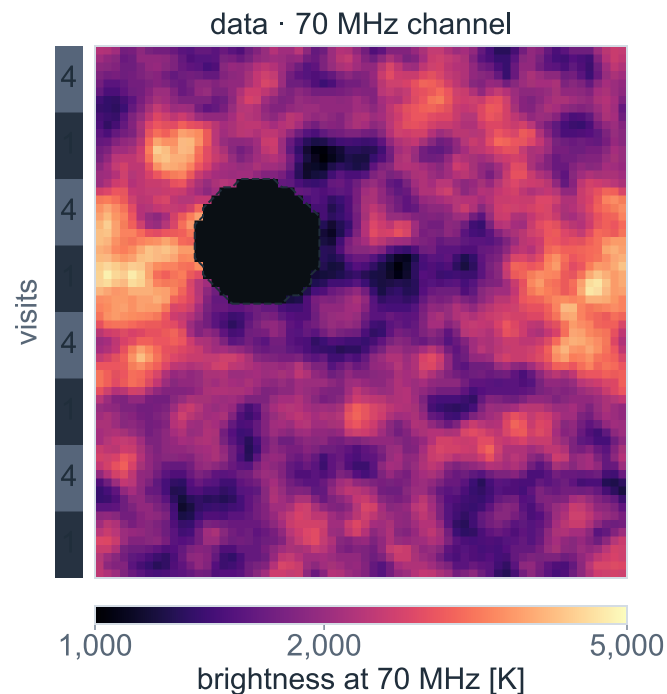
$$\ln d_p = t_p - \beta_p L + C_p L^2 + f_p w$$

Linear in t , β , C ; noise f_p in every channel.

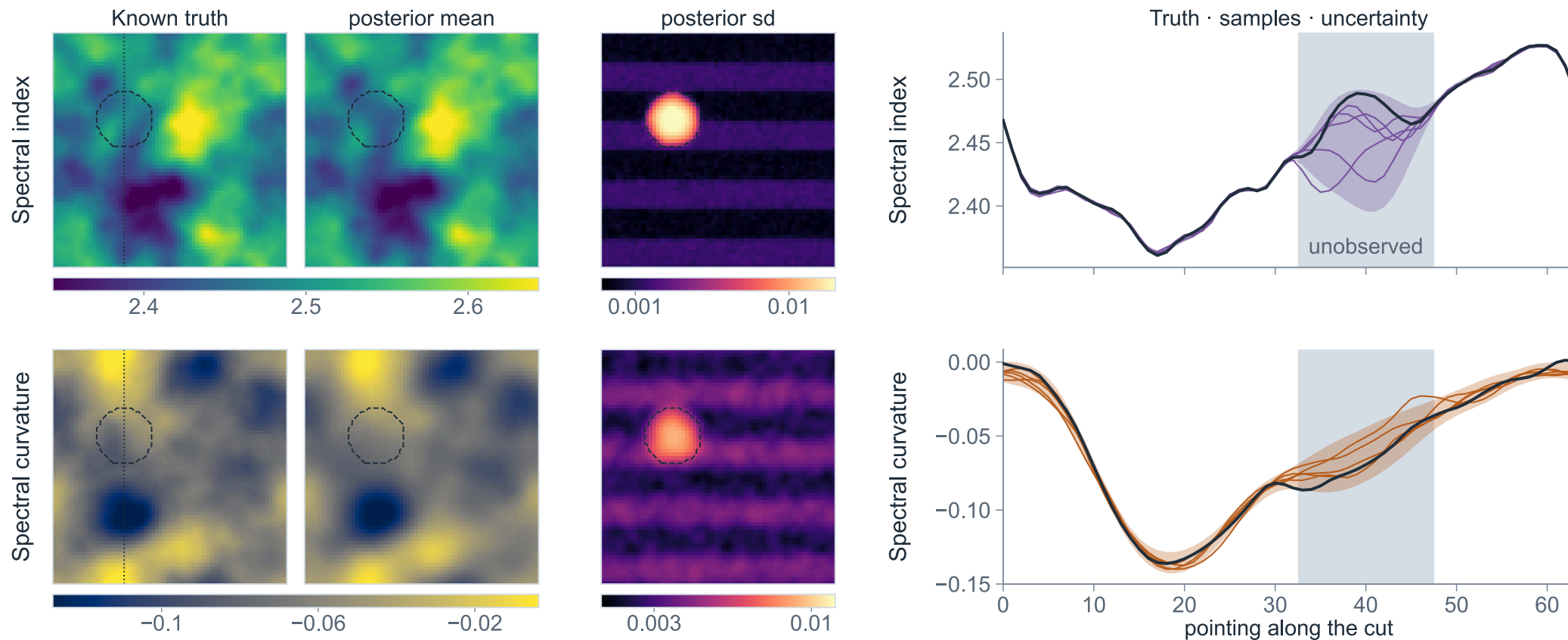
Gaussian-field priors on β and C : 12,288
unknowns, 62,608 data, one Gaussian
block, exact independent draws.

one global β , C : $\chi^2/\text{dof} = 123$

β , C per pointing: $\chi^2/\text{dof} = 0.99$



Exact draws map β and curvature; their errors follow the scan

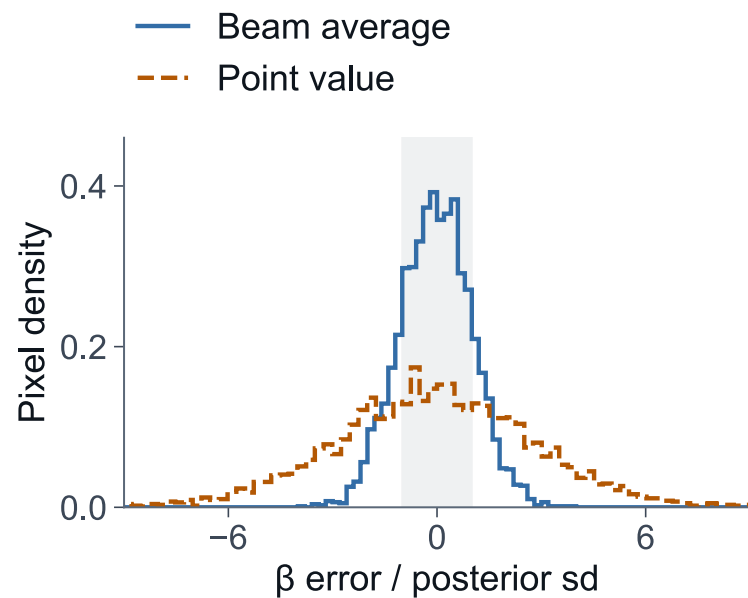


12,288 unknowns: 1.3 s per independent draw; CG error ≤ 0.01 sd in every pixel of every draw

432 unknowns, same scan: exact draws 0.023 s; NUTS R-hat 1.092 (β), min ESS 14 (C)

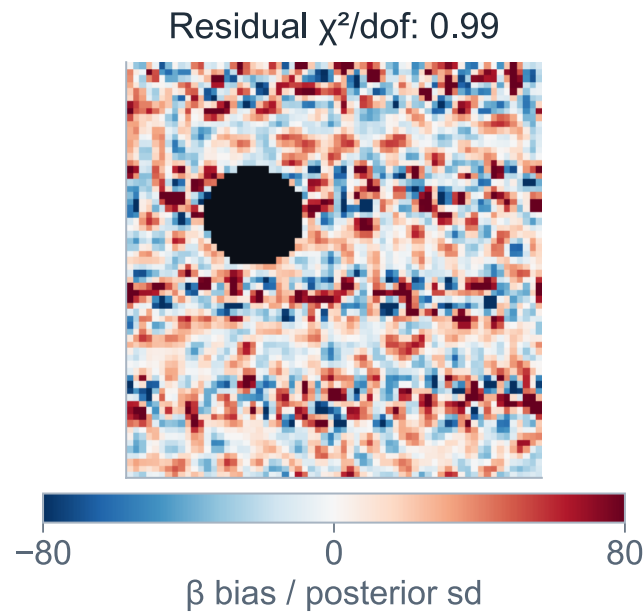
A good fit can still give the wrong map.

Beam-averaged sky



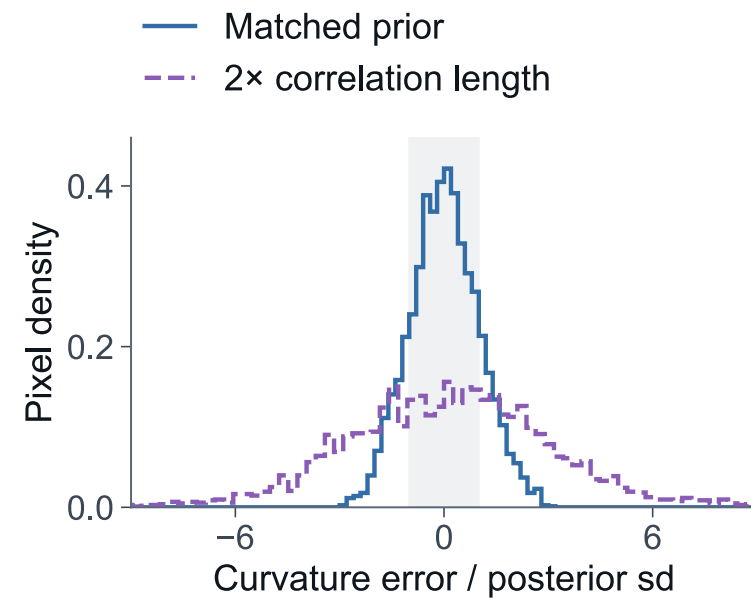
Know which sky the map estimates.

Chromatic response



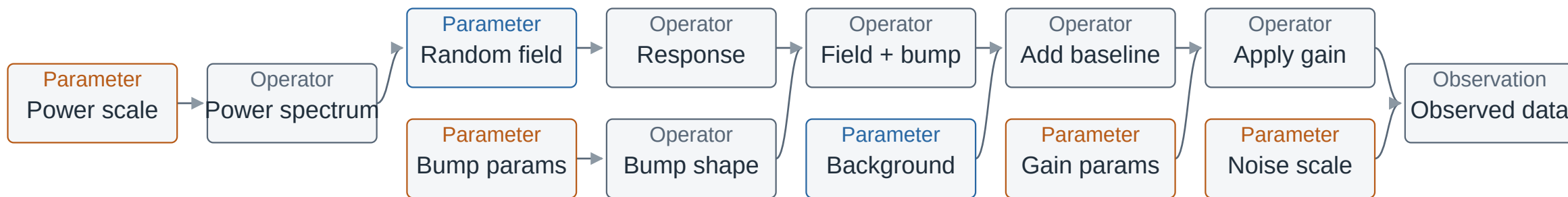
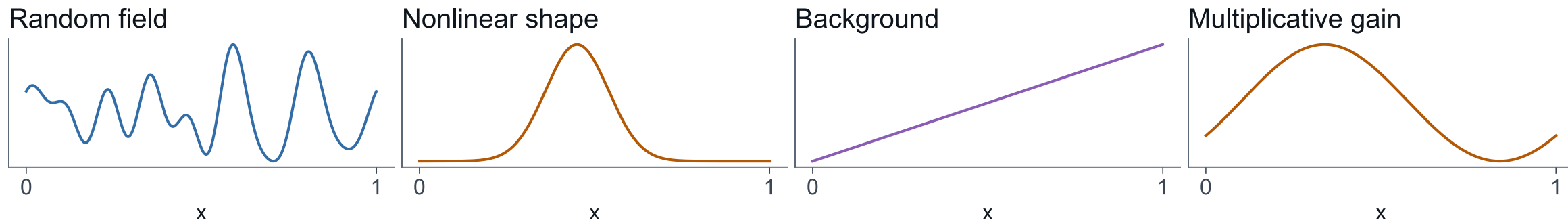
A good residual can hide a biased map.

Prior smoothness



Smoothness changes the uncertainty.

A real model. An automatically generated sweep.



Traced synthetic model · known design inputs omitted · coloured parameters identify the two updates

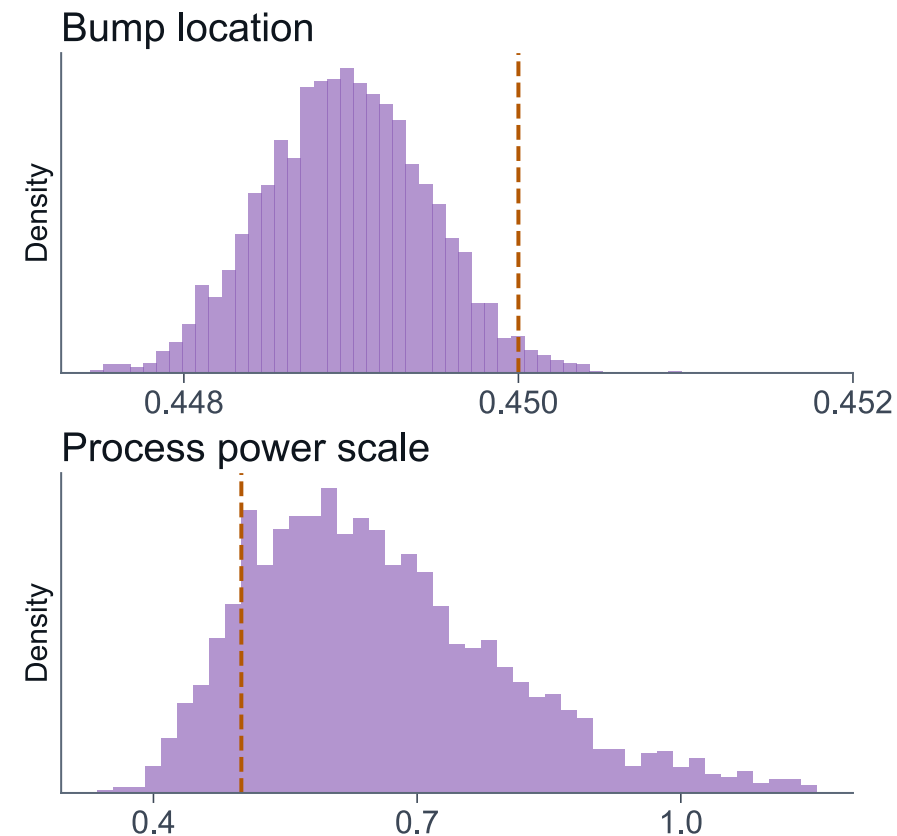
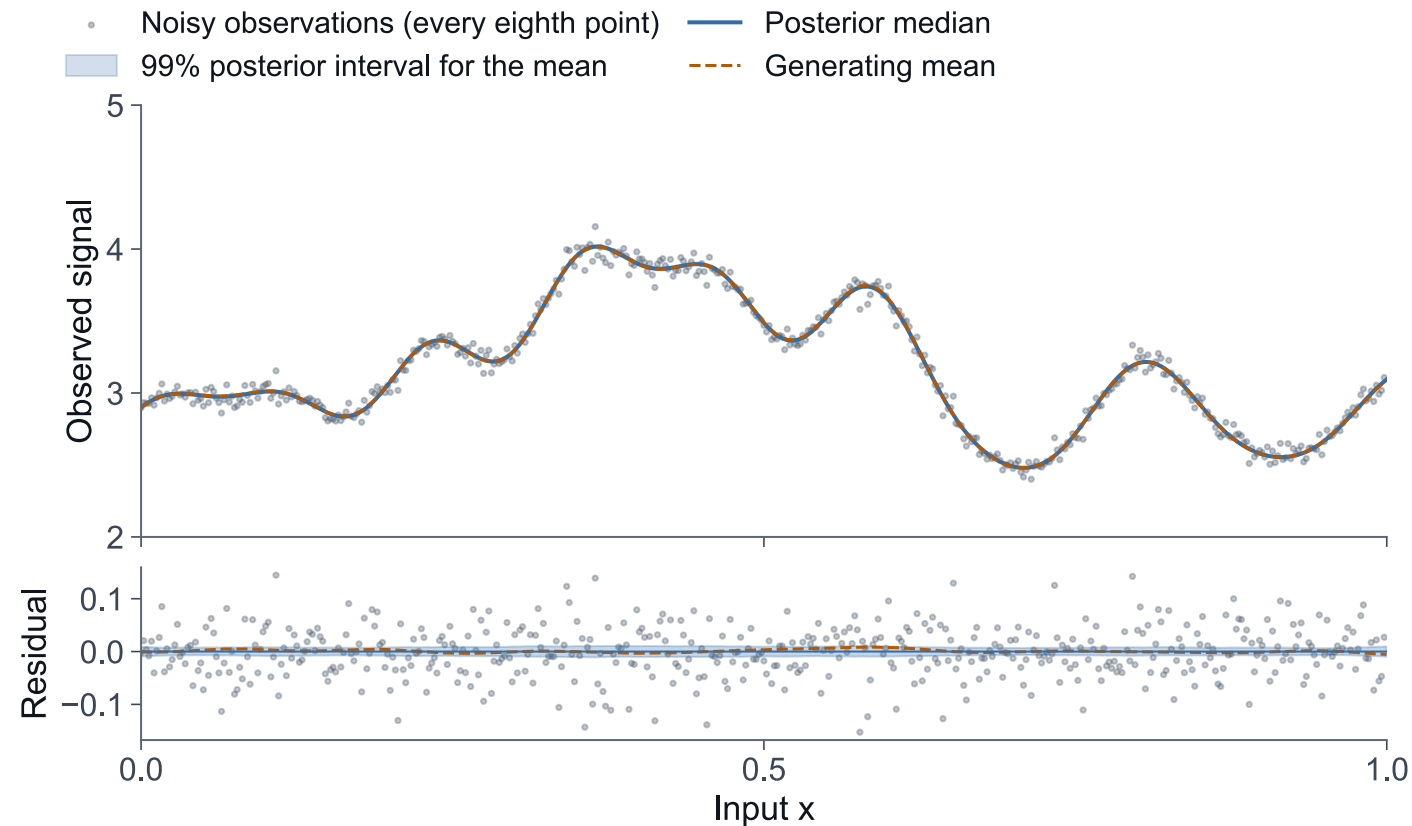
Random field + background

Moving-noise GLS proposal + MH

Power + shape + gain + noise

NumPyro NUTS · conditional on the field

Approximate the update. Keep the model.



Structured proposal + original-target MH · nonlinear remainder sampled by NumPyro

Examples and further directions

RHINO: an observing system

Sky model and simulated survey

Reuse at larger map sizes

Learning across observing nights

Native observation graph

Whole-map posterior draws

Inspectable tools for agents

Further directions

RHINO: an observing system to model



Jodrell Bank · photo P. Bull

Sky-averaged 21-cm signal

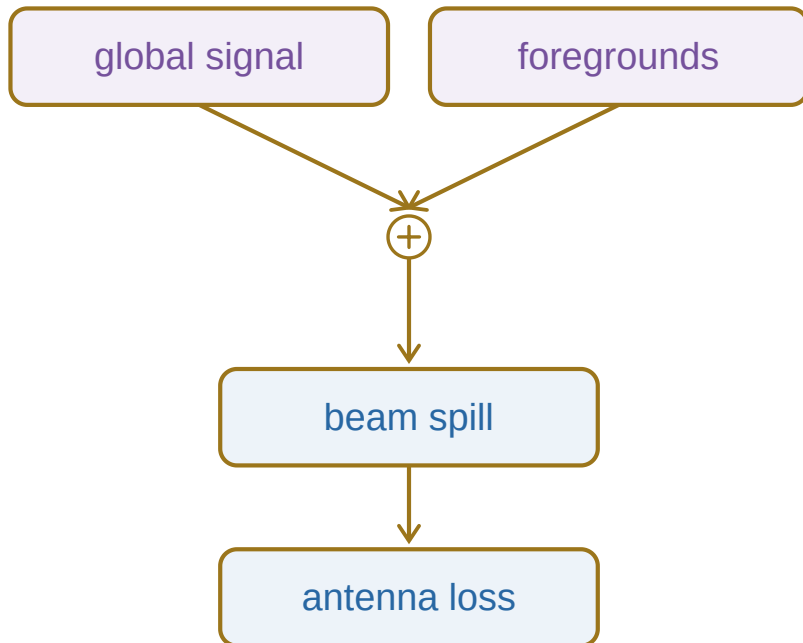
Sky and instrument response

Reference measurements

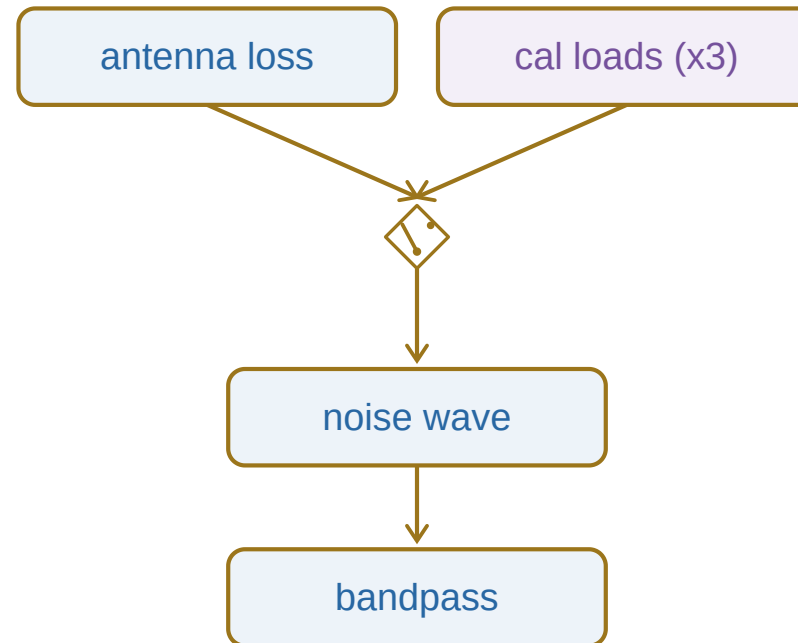
Calibration uncertainty inside the model

Rheplicant draws the twin it assembles.

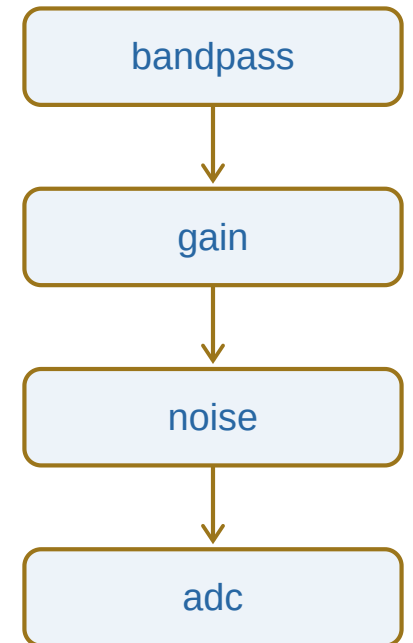
Sky → antenna



Switch → receiver

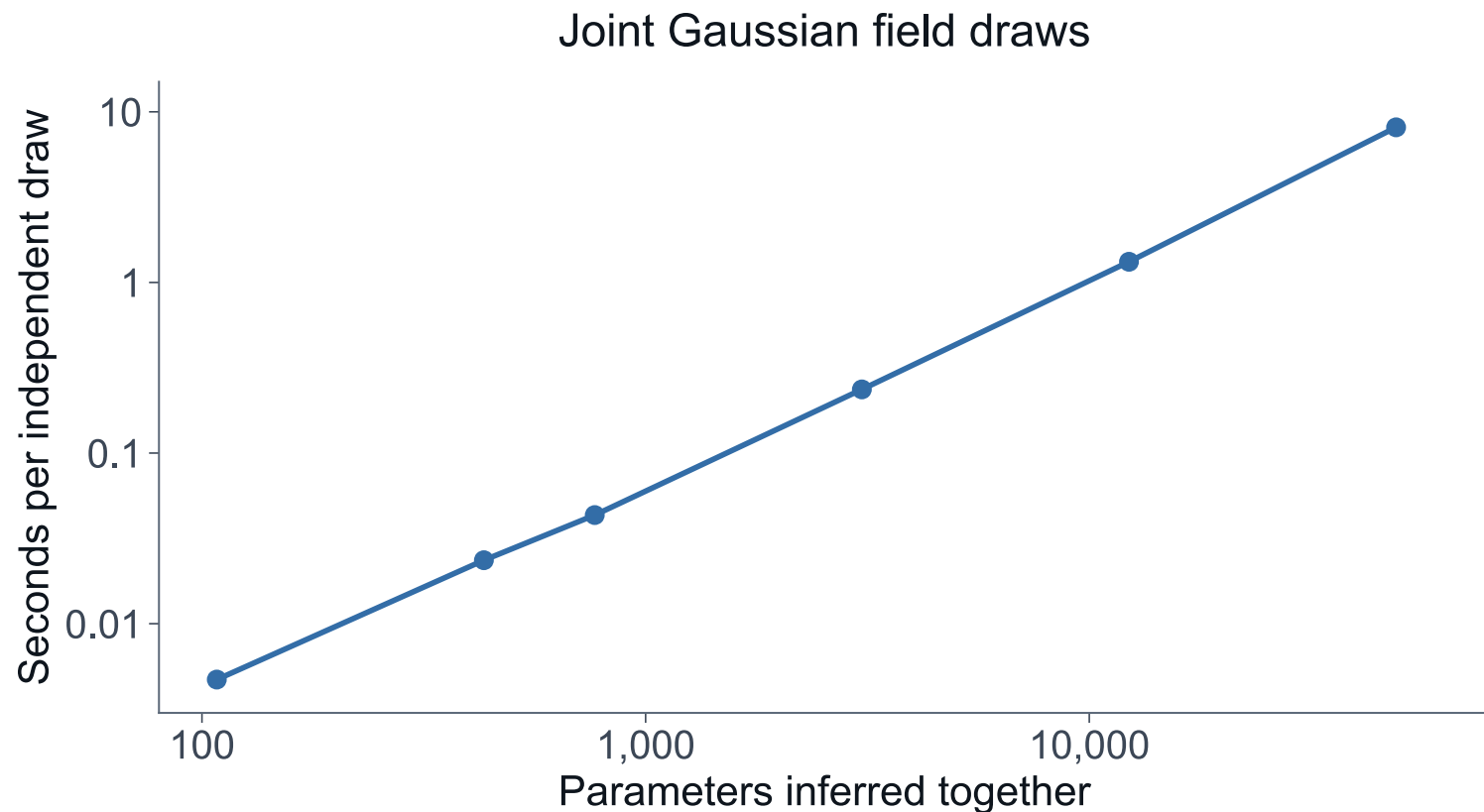


Readout



Native tour assembly · $\oplus$ adds branches · the diamond switches between antenna and reference loads · loads $\times 3$

Larger maps. The same inference machinery.



One joint field update

Keep the map's spatial structure.

Independent map draws

Reuse the Gaussian update.

Numerical accuracy checked

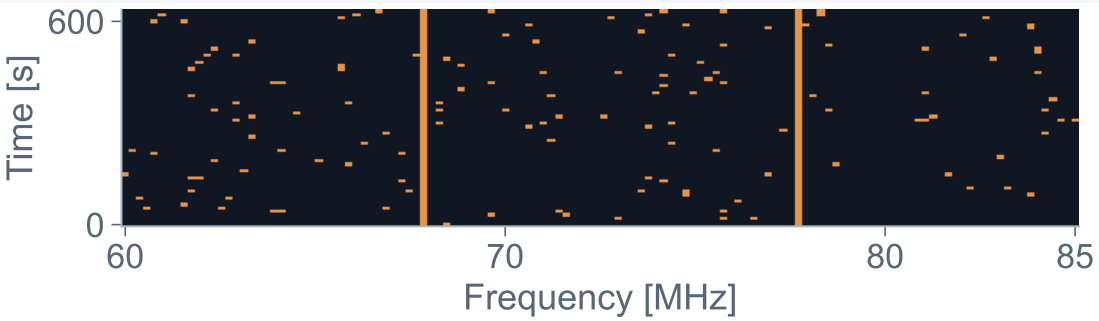
Check the solve as the field grows.

Reuse the method as the field grows.

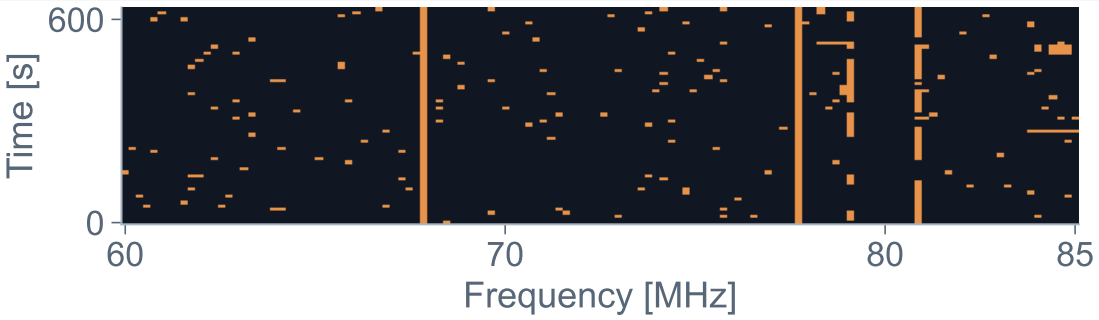
Tools for agents. Results you can inspect.

rheplicant-agent / Workbench Flagging rounds · recorded tool executions

First tool call



Refined result · checks repeated



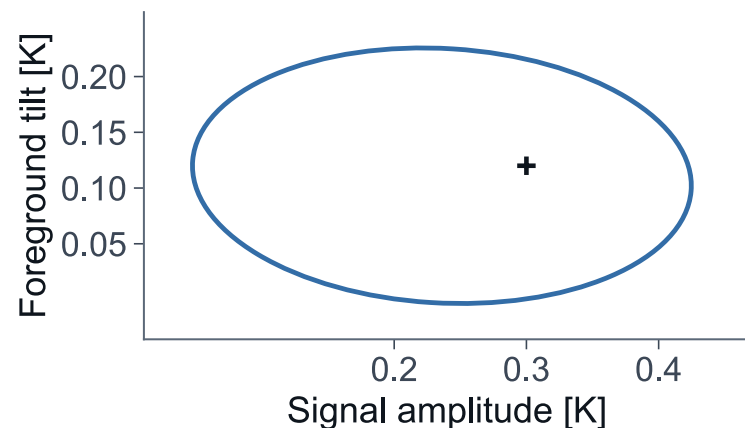
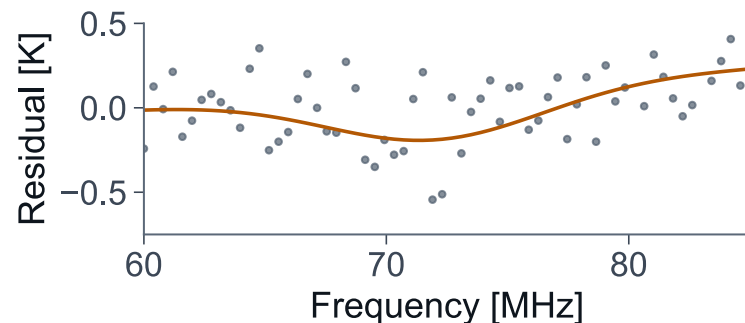
Four fixed checks · every call

Call	Stable result	Residual check	Protected data	Complete procedure	Verdict
1 First pass	—	✓	✓	—	continue
2 Refine	✓	✓	✓	—	continue
3 Broaden	—	✓	✓	✓	continue
4 Repeat	✓	✓	✓	✓	pass

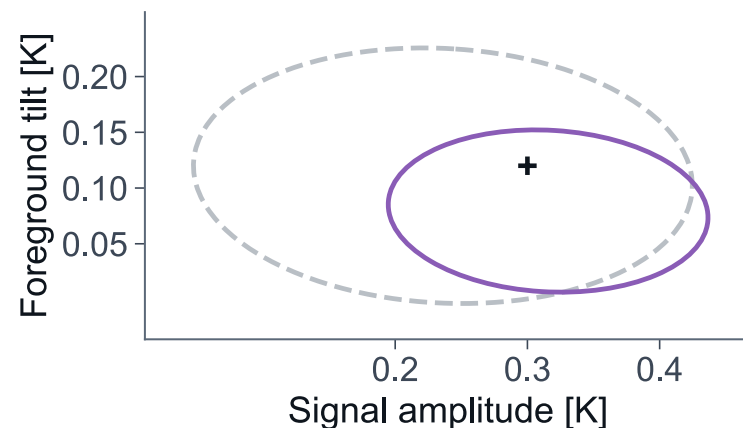
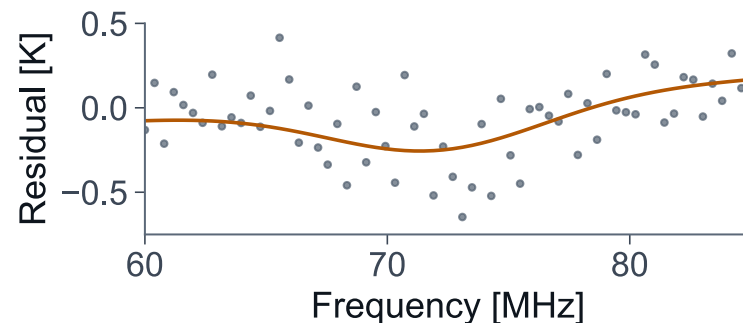
Stop when every fixed check passes.

Keep learning as observations arrive.

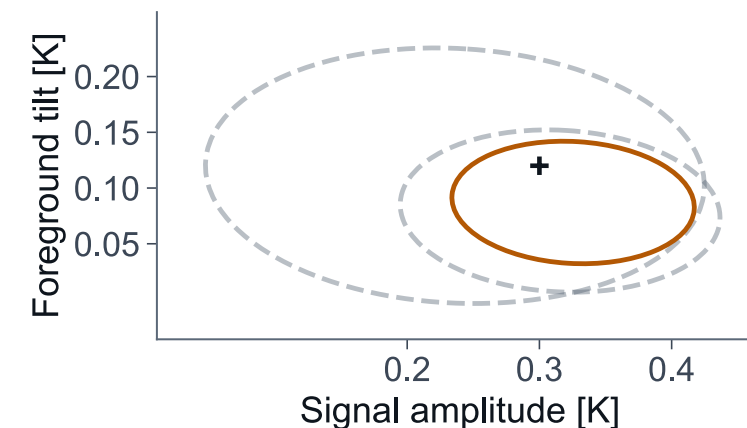
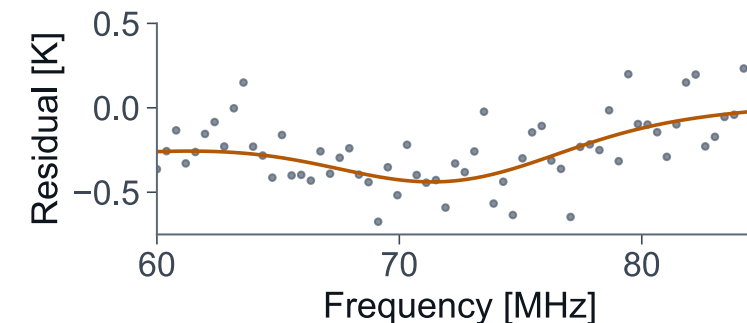
Night 1



+ Night 2



+ Night 3



Top: new data (dots), generating mean (orange) · bottom: 95% joint posterior · +: truth · dashed: earlier nights

Stored information reproduces the full-data posterior.

What comes next

Lensing

fold the Jacobi map, reduced shear and the source sample as external insertions; converged nonlinear bispectrum; ray-tracing validation

Synchrotron

the moment foreground inside the twin with Faraday depth and polarisation; coverage over noise realisations; the chromatic stress band

Inference

RHINO calibration on real data; live agent sessions on real nights; streamed campaign evidence; certificates for Gibbs sweeps; a preconditioner for long integrations; chromatic beams inside the forward model

↘ Technical backup

Selected sources and implementation

Radio: Zhang & Chluba, *A statistical framework for synchrotron emission*, in prep.; SyncMoments 0.4.0. Bull et al. (2025), RASTI, RHINO. rheplicant 0.9 and bayesmith 0.10 on PyPI; Zhang (2025) MomentEmu, arXiv:2507.02179.

Gibbs sampling: Wandelt, Larson & Lakshminarayanan (2004), PRD 70, 083511; Eriksen et al. (2004), ApJS 155, 227; Eriksen et al. (2008), ApJ 676, 10; BeyondPlanck Collaboration (2023), A&A 675, A1; Galloway et al. (2023), A&A 675, A3; Kennedy et al. (2023), ApJS 266, 23; Murphy et al. (2026), MNRAS 551, 1503; Dutta et al. (2026), arXiv:2604.26885; Zhang, Bull, Santos & Nasirudin (2026), RASTI, hydra-tod. **Foreground targets:** Pritchard & Loeb (2012); Fixsen et al. (2011); Dowell & Taylor (2018); Chluba (2016); Abitbol et al. (2017); BICEP/Keck (2021); Irfan et al. (2023). Gibbs schematic after K. Glasscock, RHINO overview (P. Bull).

Formalism: Zhang, Bull, Clarkson & Nicola, *Statistical Field Theory for Weak Gravitational Lensing*, arXiv:2606.19472, and its analysis package. **Software:** sft-wick, arXiv:2606.19480; kappa_PDF for the controlled optical PDF models; Unify-WL for the correction panorama.

Stochastic optics: [Fleury, Larena & Uzan \(2015\)](#). **Precision shear modelling:** [Euclid Collaboration \(2024\)](#). **Source sampling and finite images:** [lensing bias](#), [extended sources](#).

PDFs: [Patton et al.](#); [Clerkin et al.](#); [Barthelemy et al.](#); [Zhong et al.](#)

Provenance and hashes: [assets/source_manifest.json](#) · [generated_manifest.json](#) · [reused_manifest.json](#) · [SOURCES.md](#)