PROCEEDINGS OF SCIENCE

Semi-parametric Air Shower Shape Reconstruction with Information Field Theory

Maximilian Straub,[a,*] **Torsten Enßlin,**[b,c] **Martin Erdmann**[a] **and Philipp Frank**[b]

[a] *RWTH Physics Institute III A, Otto-Blumenthal-Str., 52074 Aachen, Germany*

[b] *Max Planck Institute for Astrophysics, Karl-Schwarzschild-Str. 1, 85748 Garching bei München, Germany*

[c] *Ludwig Maximilian University of Munich, Geschwister-Scholl-Platz 1, 80539 München, Germany*

E-mail: straub@physik.rwth-aachen.de

Extensive air showers developing through the atmosphere emit radio signals, which can be measured by antennas on ground. Multiple antennas measure this radio footprint as voltage signals as a function of time. Information Field Theory (IFT) is a framework to address field reconstruction problems with many degrees of freedom, such as imaging, and is based on Bayesian statistics. With IFT, it is possible to approximate posterior distributions over field configuration spaces and thereby to effectively invert measurements. An inference process based on IFT is devised to reconstruct the four-dimensional distribution of currents in the atmosphere. Inside a volume along to the shower axis, is filled with non-parametric structures describing the mascroscopic currents that would produce the electric field at the antennas. In this contribution, we demonstrate this approach on a toy example to investigate technical limitations and challenges.

10th International Workshop on Acoustic and Radio EeV Neutrino Detection Activities - ARENA2024
11-14 June 2024
University of Chicago, USA

*Speaker

PoS(ARENA2024)055

1. Introduction

One of the most pressing issues in radio emission from cosmic-ray induced air showers is the direct reconstruction of the emission evolution. The aim is to reconstruct from measured antenna signals the temporal evolution of the radio emission and, in particular, the height-dependent magnitude of the energy flux.

In the already successful reconstructions of particle shower development via fluorescent light using telescopes, the increase in UV light to a shower maximum and subsequent decent can be seen. Such a development should also be present for radio signals. Simulations show the increase to maximum emission and the decrease, which follow the development of particle production in the shower with a slight delay. An important difference to fluorescent light is that the atmosphere is transparent for radio waves in the 30-80 MHz range, and therefore an unattenuated radio signal arrives at the antennas.

So far, for individual measured cosmic-ray events including radio signals, a number of simulations have been carried through. In this simulation dataset, a single simulated shower was then identified that is as similar as possible to the radio measurements. In this paper we will change the perspective and present an inference approach that in principle allows a direct reconstruction of the radio emission evolution in the atmosphere from the ground measurements in antennas.

2. Calculating Radio Emissions of Extensive Air Showers

In this section, we first recall the Lièнard-Wiechert potentials and specify them for the situation of a moving charged particle in an air shower [1]. Then we describe our simplified air shower model of an ordered multitude of charges, moving at relativistic velocities along a shower axis and allowing for lateral movements at non-relativistic velocities.

2.1 Lièнard-Wiechert Potential

The solution is modeled via the Lièнard-Wiechert four-potential A^μ which describes a relativistically moving point charge q along a trajectory. Here, the scalar part of A^μ denotes the charge q at position $\mathbf{r_s}$ that is moving with the normalized velocity $\boldsymbol{\beta} = \mathbf{v}/c$:

$$\varphi(\mathbf{x}_{\text{obs}}, t_{\text{obs}}) = \left(\frac{q}{n_{\text{eff}} |\mathbf{x}_{\text{obs}} - \mathbf{r_s}| - n_{\text{eff}} (\mathbf{x}_{\text{obs}} - \mathbf{r_s}) \cdot \boldsymbol{\beta}} \right)_{t_{\text{ret}}} \tag{1}$$

The three-vector potential $\mathbf{A}$ is aligned with the velocity $\boldsymbol{\beta}$:

$$\mathbf{A}(\mathbf{x}_{\text{obs}}, t_{\text{obs}}) = \left(\frac{q\,\boldsymbol{\beta}}{n_{\text{eff}} |\mathbf{x}_{\text{obs}} - \mathbf{r_s}| - n_{\text{eff}} (\mathbf{x}_{\text{obs}} - \mathbf{r_s}) \cdot \boldsymbol{\beta}} \right)_{t_{\text{ret}}} \tag{2}$$

Here $n_{\text{eff}}(\mathbf{x}_{\text{obs}}, \mathbf{x}')$ describes an effective refractive index between observer and point charge in space, which is obtained here by an averaging integral along a straight line between charge and observer. Approximating the trajectory as a straight line works well in air [2]. Both expressions are to be evaluated at the retarded time t_{ret} defined as

$$t_{\text{ret}} \equiv t_{\text{obs}} - \frac{n_{\text{eff}}(\mathbf{x}_{\text{obs}}, \mathbf{r_s})}{c} \left| \mathbf{x}_{\text{obs}} - \mathbf{r_s} \right| . \tag{3}$$

This is the electromagnetic foundation of calculating the emissions for air showers.

2.2 Numerical Calculation

To carry out the numerical computations, spacetime is partitioned into discrete elements: The air shower is modeled as a collection of point charges. The charges exist inside a co-moving volume that is moving with normalized velocity $\boldsymbol{\beta}_\mathrm{s}$ along the shower-axis. This volume is partitioned into a regular lattice of voxels, where each voxel is assigned a unique index α. The shower time, or *retarded time*, is discretized into bins $t_{\mathrm{ret},k}$. A voxel's location in the shower frame is fixed. For the observer, the voxels are moving and therefore are dependent on the shower time, denoted as $\mathbf{x}_\alpha(t_{\mathrm{ret},k}) \equiv \mathbf{x}_{\alpha,k}$.

The time for the observer is discretized into time bins $t_{\mathrm{obs},j}$, which is used to calculate the observed signal and in practice will be prescribed by antenna electronics. With function $T_\mathrm{obs}(\mathbf{x}_{\alpha,k}, \mathbf{x}_\mathrm{obs}) = t_\mathrm{ret} + \frac{n_\mathrm{eff}(\mathbf{x}_\mathrm{obs}, \mathbf{x}_\alpha(t_\mathrm{ret}))}{c} |\mathbf{x}_\mathrm{obs} - \mathbf{x}_\alpha(t_\mathrm{ret})|$ we calculate at what point in time the shower voxel α will influence the observer for a given shower time $t_{\mathrm{ret},k}$.

Given that each volume element contains a moving point-charge, we use the closed Lienard-Wiechert solutions. The shower-time-dependent charge and drift-velocity inside a given spacetime voxel (α, k) are denoted as $q_{\alpha,k}$ resp. $\boldsymbol{\beta}'_{\alpha,k}$. For each point in shower time, the algorithm may fit a different current distribution inside the volume.

With this discretized air shower, we calculate the four-potential of its radio emission via:

$$A^\mu(t_{\mathrm{obs},j}, \mathbf{x}_\mathrm{obs}) = \sum_{\alpha,k} \frac{q_{\alpha,k} \begin{pmatrix} 1 \\ \rho_k \boldsymbol{\beta}_\mathrm{s} + \boldsymbol{\beta}'_{\alpha,k} \end{pmatrix} \delta\left(T_\mathrm{obs}(\mathbf{x}_{\alpha,k}, \mathbf{x}_\mathrm{obs}) - t_{\mathrm{obs},j}\right) \left|\frac{\mathrm{d}T_\mathrm{obs}(\mathbf{x}_\alpha(t_\mathrm{ret}), \mathbf{x}_\mathrm{obs})}{\mathrm{d}t_\mathrm{ret}}\right|_{t_{\mathrm{ret},k}}}{n_\mathrm{eff}(\mathbf{x}_\mathrm{obs}, \mathbf{x}_{\alpha,k}) \left(1 - n_\mathrm{eff}(\mathbf{x}_\mathrm{obs}, \mathbf{x}_{\alpha,k}) \frac{(\mathbf{x}_\mathrm{obs} - \mathbf{x}_{\alpha,k}) \cdot \boldsymbol{\beta}_\mathrm{s}}{\|\mathbf{x}_\mathrm{obs} - \mathbf{x}_{\alpha,k}\|}\right) \left\|\mathbf{x}_\mathrm{obs} - \mathbf{x}_{\alpha,k}\right\|}. \quad (4)$$

We sum over all spatial indices α and all temporal indices k. Component A^0 consists only of the charge q of the point. A^i is made up of currents, which are split up into two components. The longitudinal current depends on the velocity of the shower front, but only the excess charges radiate, modeled with the density parameter ρ. Transverse to the shower axis, the current simply consists of transverse velocities and the charge.The discrete delta function selects the observer time bin $t_{\mathrm{obs},j}$ that will be affected by the emission. We weight the contribution of a space-time element with the Jacobian $|\frac{\mathrm{d}T}{\mathrm{d}t_\mathrm{ret}}|$. The denominator of the shower contains the *Doppler factor*, idem as the Lienard-Wiechert solution.

From this, we can calculate the electric field via numerical derivatives:

$$E^i(t_{\mathrm{obs},j}, \mathbf{x}_\mathrm{obs}) = -\frac{\partial}{\partial t} A^i(t_{\mathrm{obs},j}, \mathbf{x}_\mathrm{obs}) - \frac{\partial}{\partial x^i} A^0(t_{\mathrm{obs},j}, \mathbf{x}_\mathrm{obs}) \quad (5)$$

For computational reasons, the current version of the inference algorithms only considers the time-derivative contribution to the electric field. This will be amended in future versions.

3. Inference

3.1 Information Field Theory

Information Field Theory (IFT) [5] is the application of information theory to fields. With IFT, we can summarize the available information on some field-like quantity using Bayesian probabilities.

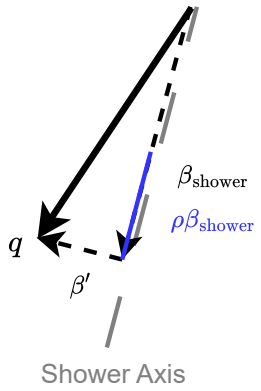


Figure 1: Separating the charge's movement into a transversal and longitudinal component. The longitudinal current is scaled by a density parameter, since we are only interested in an overdensity for that component, indicated in blue.

Specifically, it allows to reason about problems with infinite amounts of degrees of freedom, that is, continuous, physical fields, and their discretized counterparts. The degrees of freedom are constrained by probabilistic prior assumptions that contain physical knowledge about the problem. With these, it is possible to calculate field expectation values.

In Information Field Theory, probability distributions are expressed as *information Hamiltonians*, related via the Boltzmann distribution:

$$\mathcal{P}(s|d) = \frac{\mathcal{P}(d,s)}{\mathcal{P}(d)} \equiv \frac{e^{-\mathcal{H}(d,s)}}{\mathcal{Z}(d)} \tag{6}$$

This reformulation turns probabilities into additive quantities.

To write down a problem in this framework, we need to define the measurement equation:

$$d = R[s] + n \tag{7}$$

where d denotes data, s the signal of interest, n an additive noise term, and R some operator, that maps the signal into data space. This equation is directly linked to the likelihood. Assuming the noise follows Gaussian statistics with covariance N, the likelihood reads

$$\mathcal{P}(d|s) = \mathcal{G}(d - R[s], N) \tag{8}$$

which leads to the *likelihood Hamiltonian*

$$\mathcal{H}(d|s) = -\ln \mathcal{G}(d - R[s], N) = \frac{1}{2}(d - R[s])^\dagger N^{-1}(d - R[s]) + \frac{1}{2}\ln|2\pi N|, \tag{9}$$

which can be used in the inference process as explained below. Finally, one often used approach is *standardization*, that turns hierarchical models into flat models, which decouples input variables

and decreases computational complexity. This is also known as reparametrization trick [3]. In these standard coordinates, all parameters follow a priori a standard Gaussian distribution, centered around zero with unit width: $\xi \sim \mathcal{G}(\xi|0, \mathbb{1})$. These are transformed to uniformly distributed variables u, which are in turn transformed to the hierachical prior $\mathcal{P}(\theta)$ with

$$\theta = \mathcal{F}^{-1}_{\mathcal{P}(\theta)} \circ \mathcal{F}^{-1}_{\mathcal{G}(\xi,\mathbb{1})}(\xi) \equiv f(\xi). \tag{10}$$

$\mathcal{F}^{-1}(\ldots)$ denotes the applied transformations. With this, ξ are independent, and the hierarchical complexity is entirely contained within the transformation functions. The standardized Hamiltonian reads

$$\mathcal{H}(d, \xi) = \mathcal{H}(d|f(\xi)) + \frac{1}{2}\xi^\dagger \mathbb{1}\xi, \tag{11}$$

where terms independent of input parameters are dropped.

3.2 JAX

JAX is an open source machine learning framework [4]. For the user it is very similar to the widely-used numpy package [6]. Beyond ordinary numpy functionality, it offers three key features that make it particularly useful for this inference problem:

- Most JAX code can be automatically differentiated, yielding exact derivatives and gradients. This is referred to as autograd.
- Python code can be just-in-time (JIT) compiled under certain conditions, e.g., no side effects, that is subroutines must be true functions. Just-in-time compilation can then adaptively optimize the code and vastly decrease computation times, often over multiple magnitudes.
- JAX is able to use Nvidia's cuda backend [8], allowing seamless execution on GPUs. This can decrease computation time further.

Next to autograd and JIT compilation, JAX offers the convenience function `vmap`, which maps functions in parallel over arrays. This automatically vectorizes calculations written for individual elements. The function `pmap` is similar, but instead of vectorization, it parallelizes execution over multiple computation cores.

As such, JAX is essential for this inference, by providing both gradients for the optimization procedure, as well as time-efficient procedures to work with the large arrays describing the shower geometry.

3.3 Shower Model

During inference, the algorithm will generate possible shower geometries and evolutions, whose radio emission is simulated and measured as described above. The shower model needs to be expressed in a way that is accessible to the fitting procedure. This is done as follows:

$$q_{\alpha,k} = \exp(\phi_{\alpha,k}); \quad \boldsymbol{\beta}'_{\alpha,k} = \frac{\boldsymbol{\chi}_{\alpha,k}}{\sqrt{1 + \|\boldsymbol{\chi}_{\alpha,k}\|^2}} \tag{12}$$

describe charges and drift velocities for each spacetime element (α, k) with some voxel-dependent values $\phi_{\alpha,k}$ and $\boldsymbol{\chi}_{\alpha,k}$. To produce only positive charges per voxel, we plug the corresponding values

into an exponential function. This way, a possible degeneracy is broken where current direction and charge sign cancel each other. The resulting vector describing the drift velocities is forced to have a magnitude of strictly less than 1, $\|\boldsymbol{\beta}'\| < 1$. To impose structure on the shower geometry, we build spatial and temporal correlations into the prior model by generating ϕ and $\boldsymbol{\chi}$ with Gaussian processes:

$$\phi \hookleftarrow \mathcal{G}(\phi, \Phi); \quad \boldsymbol{\chi} \hookleftarrow \mathcal{G}(\boldsymbol{\chi}, X) \tag{13}$$

with Φ and X as covariance matrices of the Gaussian processes. Since the correlation structures of the shower are unknown, these matrices are to be inferred from data. For this, they are assumed to be diagonal in Fourier space and modeled via power spectra. Thus, we get

$$q_{\alpha,k} = \exp(\phi_{\alpha,k}) = \exp(\phi(\xi_{\alpha,k}, P_\Phi(\xi_\Phi))) \equiv q(\xi) \tag{14}$$

where ξ are standard normal distributed variables, consisting of per-pixel per-time values $\xi_{\alpha,k}$ and the spatio-temporal correlation power spectrum $P_\Phi(\xi_\Phi)$. In particular, the power spectra are modeled non-parametrically, they are generated by another Gaussian process each. The Gaussian processes producing the spectra are tuned by a handful of hyperparameters. The entire shower is then described by a set of random numbers $\xi = (\xi_{\alpha,k}, \xi_\Phi, \xi_{\alpha',k'}, \xi_X)$.

3.4 Approximating Posteriors

The posterior distribution describing the shower shape s given observed data D can be expressed using the model parameters ξ, which together with the generative functions $q(\xi)$ and $\boldsymbol{\beta}'(\xi)$ provide full reconstruction statistics. This inference uses Metric Gaussian Variational Inference (MGVI) [7], provided by the nifty package [10]. For computational simplicity, we use MGVI instead of the more expressive successor geoVI [9] in this demonstration. MGVI allows the inference for problems with many degrees of freedom, such as this four-dimensional imaging task. The posterior distribution is approximated by a multivariate Gaussian:

$$\mathcal{P}(s|D) \equiv \mathcal{P}(\xi|D) \approx \mathcal{G}\left(\xi - \bar{\xi}, \Xi\right) \tag{15}$$

where the covariance matrix Ξ is the Fisher information metric. The posterior is found by minimizing the Kullback-Leibler divergence $D_{\mathrm{KL}}\left(\mathcal{G}\left(\xi - \bar{\xi}, \Xi\right) \| \mathcal{P}(\xi|D)\right)$. The posterior distribution itself can be expressed by the likelihood and prior distributions via Bayes' theorem. Using the Hamiltonian language of IFT, this leads to the expression

$$\begin{aligned} D_{\mathrm{KL}}\left(\mathcal{G}\left(\xi - \bar{\xi}, \Xi\right) \| \mathcal{P}(\xi|D)\right) &= \langle \mathcal{H}(\xi|D) \rangle_{\mathcal{G}\left(\xi - \bar{\xi}, \Xi\right)} - \langle \mathcal{H}(\xi - \bar{\xi}, \Xi) \rangle_{\mathcal{G}\left(\xi - \bar{\xi}, \Xi\right)} \\ &= \langle \mathcal{H}(D|\xi) + \mathcal{H}(\xi) \rangle_{\mathcal{G}\left(\xi - \bar{\xi}, \Xi\right)} - \langle \mathcal{H}(\xi - \bar{\xi}, \Xi) \rangle_{\mathcal{G}\left(\xi - \bar{\xi}, \Xi\right)}. \end{aligned} \tag{16}$$

The evidence $\mathcal{P}(D)$ can be ignored in the minimization procedure, since it does not depend on variational parameters $\bar{\xi}$ and Ξ describing the approximate posterior distribution.

3.5 Inference of Toy Example

To test the inference procedure, a random shower geometry is generated by the model and then reconstructed by the algorithm. First we draw model parameters ξ_{test}, producing the shower shape s_{test}, consisting in turn of q_{test} and $\boldsymbol{\beta}'_{\mathrm{test}}$. The generated shower is put through the forwards

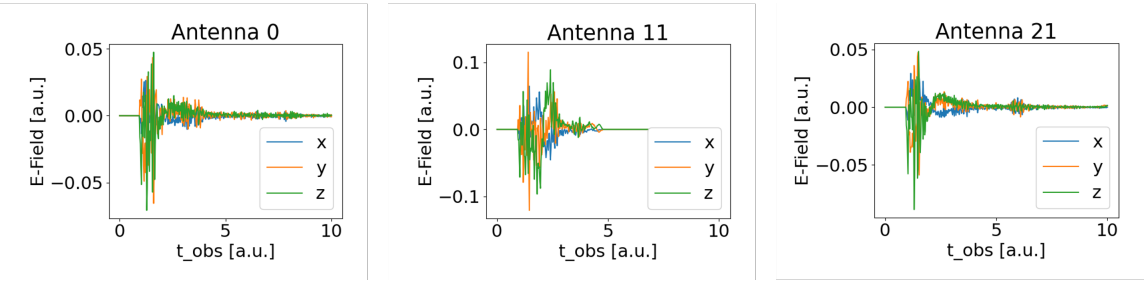
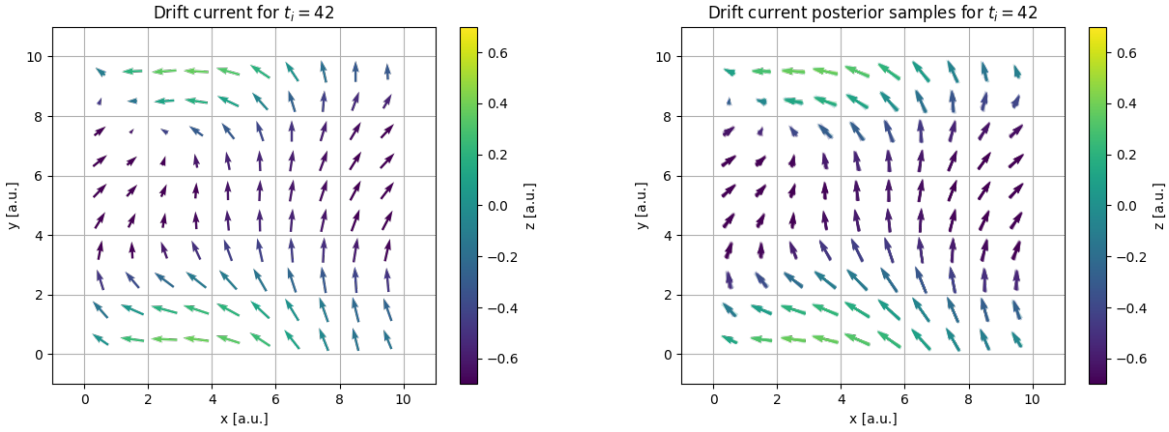


Figure 2: Electric fields at selected antennas in the example reconstruction.

process and, with added Gaussian noise n, yields the toy data D_{test}. Several toy time-traces are shown in Figure 2. Given this data, the algorithm is used to find valid q and $\boldsymbol{\beta}'$ configurations. Figure 3a shows the drift currents $\mathbf{j}_{\text{drift}} = q_{\text{test}}\boldsymbol{\beta}'_{\text{test}}$ for one slice of the shower cube at a given time t_{ret}, Figure 3b shows the drift current posterior samples. The individual samples are plotted on top of each other to visualize the directional scatter. Overall, the structure is remarkably similar. While there are small translational differences, the currents follow the same directions.



(a) True drift currents for one horizontal slice of the shower volume for one point in time.

(b) Posterior samples of drift currents for the same slice.

Figure 3: Drift currents in the toy inference problem. The color indicates the z component of the vectors.

4. Summary

In this paper we have demonstrated a working prototype of an imaging algorithm for extensive air showers. Using Information Field Theory, we devised a procedure to fit many space-time elements at the same time, reconstructing macroscopic current distributions in the atmosphere from measured electric fields. With IFT's variational inference and the computational framework JAX, it was possible to test this approach on toy examples. The inference algorithm offers statistical insights by calculating an approximate posterior. With this, we were able to show that we can calculate the posterior mean and width of the reconstructed shower shape, completely characterizing a simplified toy shower.

Next, we want to expand this method to improve the physical assumptions inside the emission calculation. After making sure the inference is self-consistent, we plan to apply this to physical scenarios and explore the imaging capabilities of the presented method. Open questions include flexibility of the framework, interpretation of statistics, and degeneracy of the solution space. Eventually, we are planning to image air showers from measured radio signals.

Acknowledgements

Maximilian Straub acknowledges funding through the German Federal Ministry of Education and Research for the project ErUM-IFT: Informationsfeldtheorie für Experimente an Großforschungsanlagen (Förderkennzeichen: 05D23PA1)

Philipp Frank acknowledges funding through the German Federal Ministry of Education and Research for the project ErUM-IFT: Informationsfeldtheorie für Experimente an Großforschungsanlagen (Förderkennzeichen: 05D23EO1).

References

1. J. D. Jackson, *Classical electrodynamics* (Wiley, New York, NY, 3rd ed. 1999), ISBN: 9780471309321, (http://cdsweb.cern.ch/record/490457).
2. K. Werner, O. Scholten, *Astroparticle Physics* **29**, 393–411, ISSN: 09276505, arXiv: 0712.2517 [astro-ph, physics:hep-ph] (July 2008).
3. D. P. Kingma, T. Salimans, M. Welling, arXiv: 1506.02557 [cs, stat], (http://arxiv.org/abs/1506.02557) (Dec. 20, 2015).
4. J. Bradbury *et al.*, *JAX: composable transformations of Python+NumPy programs*, version 0.3.13, 2018, (http://github.com/google/jax).
5. T. A. Enßlin, *Annalen der Physik* **531**, 1800127, ISSN: 0003-3804, 1521-3889, arXiv: 1804.03350 [astro-ph] (Mar. 2019).
6. C. R. Harris *et al.*, *Nature* **585**, 357–362, (https://doi.org/10.1038/s41586-020-2649-2) (Sept. 2020).
7. J. Knollmüller, T. A. Enßlin, arXiv: 1901.11033 [astro-ph, physics:physics, stat] (Jan. 2020).
8. NVIDIA, P. Vingelmann, F. H. Fitzek, *CUDA, release: 10.2.89*, 2020, (https://developer.nvidia.com/cuda-toolkit).
9. P. Frank, R. Leike, T. A. Enßlin, *Entropy* **23**, 853, ISSN: 1099-4300, arXiv: 2105.10470[astro-ph,stat], (2024; http://arxiv.org/abs/2105.10470) (July 2, 2021).
10. G. Edenhofer *et al.*, arXiv: 2402.16683 [astro-ph, stat] (Feb. 2024).