A Mode-Tracking Formalism for Radio Interferometric Intensity Mapping

With Local Beam, Local Fringe and Global Sky

Zheng Zhang

April 2025

Abstract

Correlated noise in radio interferometric measurements poses significant challenges for statistical science extraction in intensity mapping experiments. This work addresses the critical need for a unified definition of sky modes across all data points to enable consistent analysis of correlations in both raw measurements and derived power spectra. While radio interferometric measurement equations (RIME) are inherently grounded in antenna-static coordinates, the nonlinear transformation between sky and antenna coordinates complicates the connection to universally defined sky modes. To reconcile ground-based observational practicality with global sky mode unification, we introduce a formalism based on three core components: local beam, local fringe, and global sky. The framework establishes exact analytical mappings between sky signals, observed data, and their covariances, with derivations presented in spherical harmonic space and translated to Cartesian Fourier modes via ℓ-to-$k_\perp$ conversion.

Contents

Chapter 1

Introduction

Correlated noise in the statistical measurements could be a problem in the science extraction of intensity mapping experiments. In order to analyse correlations between the radio interferometric measurements and between the power spectrum estimates, it is instructive to establish a universally consistent definition of sky modes across all data points.

Due to the nature of ground-based observations, the radio interferometric measurement equations (RIME) are indeed best understood from a ground-based (i.e., antenna-static) perspective. However, direct application of the Fourier transform understanding inherent in visibility measurements presents a challenge: the resulting sky Fourier modes are conjugate to antenna coordinates. The nonlinear transformation from sky coordinates to antenna coordinates makes it difficult to establish a direct analytical connection between the derived sky modes and universally defined sky modes (i.e. a common basis applicable to all data points).

To reconcile the practicality of ground-based interpretation with the unification of sky mode definitions, this note introduces a formalism involving three key components: local beam, local fringe, and global sky. The derivations are straightforward, with the emphasis on explicitly identifying potential simplifications and defining the concrete forms of linear mappings.

Chapter 2

Covariance of radio interferometric drift-scan measurements

2.1 Notations and Conventions

Coordinate

It is instructive to distinguish between the *sky-static* and *antenna-static* coordinate systems. For a field point (a physical identity rather than coordinates) on the celestial sphere:

- $\hat{n}_a$ denotes the coordinate in the antenna system (i.e., static to the antenna). In spherical coordinates, it is usually represented by the beam angles, (θ_a, ϕ_a).
- $\hat{n}_s$ is the celestial coordinate, which in a spherical system is (θ_s, ϕ_s).
- The transformation between them is LST dependent and is given by a rotation operator
$$\hat{n}_s(\hat{n}_a, t) = \mathcal{R}(t)\hat{n}_a.$$
For now, it is fine to keep the operation abstract. It will be treated explicitly later when we need to specify the representations of the rotations.

Beam and fringe

Below we present notations for antenna beams and fringes. We will also introduce the spherical harmonic notation for practical use with the rotation transformation.

- The primary beam is represented as

$$B_{ij}(\hat{n}_a,\nu) \equiv A_i(\hat{n}_a,\nu)\, A_j^\dagger(\hat{n}_a,\nu) = \sum_{\ell m} B_{\ell m}^{(ij)}(\nu) Y_{\ell m}(\hat{n}_a) \tag{2.1}$$

  where B is the intensity beam and A is the far-field E-beam, and i and j are antenna indices. $B_{\ell m}$ is the spherical harmonic coefficient, defined as the inner product of the field and the spherical harmonic.

- The fringe pattern is given by

$$F(\hat{n}_a, \boldsymbol{b}_{ij}, \nu) \equiv e^{-i\frac{2\pi\nu}{c}\boldsymbol{b}_{ij}\cdot\hat{n}_a} = \sum_{\ell m} F_{\ell m}(\boldsymbol{b}_{ij},\nu) Y_{\ell m}(\hat{n}_a) \tag{2.2}$$

We can further work out a more explicit form for the primary beam and fringe modes. For the primary beam modes $B_{\ell_1 m_1}^{(ij)}(\nu)$, only the $m_1 = 0$ modes survive if the primary beam is invariant under rotations around the antenna axis:

$$B_{\ell_1 m_1}^{(ij)}(\nu) = 0, \quad \text{if } m_1 \neq 0 \text{ (for rotationally symmetric beam)}. \tag{2.3}$$

For the fringe modes, the plane wave $e^{-i\boldsymbol{u}\cdot\hat{n}_a}$, where $\boldsymbol{u}$ is a vector in the tangent plane at the pole of the antenna system[1], can be expanded in terms of the spherical harmonics using the plane wave expansion:

$$e^{-i\boldsymbol{u}\cdot\hat{n}_a} = \sum_{\ell} (-i)^\ell (2\ell+1) j_\ell(u) P_\ell(\hat{u}\cdot\hat{n}_a) \tag{2.4}$$

where $\boldsymbol{u} = \frac{2\pi\nu}{c}\boldsymbol{b}_{ij} = u\,\hat{u}$ and j_ℓ are the Bessel functions. The above equation effectively gives

$$F_{\ell m}(\boldsymbol{b}_{ij},\nu) = 4\pi(-i)^\ell j_\ell(u) Y_{\ell m}^*(\hat{u})\,. \tag{2.5}$$

Recall that in the antenna coordinate system, $\boldsymbol{b}$ (and thus $\boldsymbol{u}$) is parallel to the tangent plane at the zenith, which means that $\hat{u}$ is given by $(\theta_u = \frac{\pi}{2}, \phi_u)$. As a consequence, $Y_{\ell m}^*(\hat{u})$ is non-zero only if $\ell - m$ is even.

Sky

Since both the primary beam and the fringe are static to the antenna, I have tried to represent them in $\hat{n}_a$ to get rid of the time dependence. However, the stationary sky prefers $\hat{n}_s$. A handy representation could be to use the sky modes $T_{\ell m}$ defined

[1] This is the case when the baseline vector is perpendicular to the zenith direction.

with respect to $\hat{n}_s$, but transform the functional basis to the antenna system. Less abstractly, we can represent the sky as

$$\begin{aligned} T(\hat{n}_s,\nu) &= \sum_{\ell m} T_{\ell m}(\nu) Y_{\ell m}(\hat{n}_s) = \sum_{\ell m} T_{\ell m}(\nu) Y_{\ell m}\left(\mathcal{R}(t)\hat{n}_a\right) \\ &= \sum_{\ell m m'} T_{\ell m}(\nu)\, \mathcal{D}^{(\ell)}_{m',m}(\mathcal{R}^{-1}(t)) Y_{\ell m'}\left(\hat{n}_a\right) \end{aligned} \tag{2.6}$$

where $\mathcal{D}^{(\ell)}_{m',m}$ is the Wigner D-matrix.

For the drift scanning measurements, the sky representation could be simpler: we only need to connect the celestial coordinates with the antenna coordinates at a reference time (without loss of generality, we set $t_{\mathrm{ref}} = 0$ and denote $\mathcal{R}(t_{\mathrm{ref}}) = \mathcal{R}_0$). Then the sky can be rewritten as

$$T(\hat{n}_s,\nu) = \sum_{\ell m} T_{\ell m}(\nu)\, Y_{\ell m}\left(\mathcal{R}_0 \hat{n}_a\right) e^{imt}. \tag{2.7}$$

Note that for this simplification to work, the celestial pole should be parallel to the Earth's rotation pole, and t should be rendered in units of radians.

2.2 Radio Interferometric Measurement Equation

A best-knowledge model for measurement is

$$V(\boldsymbol{b}_{ij},\nu,t) = G_i(\nu,t) G_j^{\dagger}(\nu,t) \frac{\gamma(\nu)}{\Omega_{ij}(\nu)} V_{\mathrm{th}}(\boldsymbol{b}_{ij},\nu,t)\,(1+\hat{w}) \tag{2.8}$$

where $G_i(\nu,t)$ is the remaining gain error, $\gamma(\nu)$ is the down frequency taper function, $\Omega_{ij}(\nu)$ normalises the primary beam, and $\hat{w}$ is the white noise given by the radiometer equation, usually assumed to be Gaussian and stationary. V_{th} is the theoretical visibility given by

$$V_{\mathrm{th}}(\boldsymbol{b}_{ij},\nu,t) = \int \mathrm{d}^2\Omega\, A_i(\hat{n}_a,\nu)\, A_j^{\dagger}(\hat{n}_a,\nu)\, T(\hat{n}_s,\nu)\, e^{-i\frac{2\pi\nu}{c}\boldsymbol{b}_{ij}\cdot\hat{n}_a}. \tag{2.9}$$

We haven't specified the coordinate in which we do the integral. But of course we should do it in the antenna coordinate:

$$V_{\mathrm{th}}(\boldsymbol{b}_{ij},\nu,t) = \sum_{\ell m} T_{\ell m}(\nu) M^{\ell m}(\boldsymbol{b}_{ij},\nu) \tag{2.10}$$

$$M^{\ell m}(\boldsymbol{b}_{ij},\nu) \equiv e^{imt} \int \mathrm{d}^2\hat{n}_a\, B_{ij}(\hat{n}_a,\nu)\, F(\hat{n}_a,\boldsymbol{b}_{ij},\nu)\, Y_{\ell m}\left(\mathcal{R}_0\hat{n}_a\right) \tag{2.11}$$

which can be further expanded as

$$
\begin{aligned}
M^{\ell m}(\boldsymbol{b}_{ij},\nu) &= e^{imt}\sum_{\ell_1 m_1}\sum_{\ell_2 m_2} B^{(ij)}_{\ell_1 m_1}(\nu) F_{\ell_2 m_2}(\boldsymbol{b}_{ij},\nu)\int d^2\Omega\, Y_{\ell_1 m_1}(\hat{n}_a) Y_{\ell_2 m_2}(\hat{n}_a) Y_{\ell m}(\mathcal{R}_0\hat{n}_a) \\
&= e^{imt}\sum_{\ell_1 m_1}\sum_{\ell_2 m_2}\sum_{m'} B^{(ij)}_{\ell_1 m_1}(\nu) F_{\ell_2 m_2}(\boldsymbol{b}_{ij},\nu)\, \mathcal{D}^{(\ell)}_{m',m}(\mathcal{R}_0^{-1})\, \mathcal{G}(\ell_1,\ell_2,\ell;m_1,m_2,m')
\end{aligned}
\tag{2.12}
$$

where we have applied the definition of the Gaunt's integral

$$
\begin{aligned}
\int d^2\Omega\, Y_{\ell_1 m_1} Y_{\ell_2 m_2} Y_{\ell_3 m_3} &= \left(\frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4}\right)^{\frac{1}{2}} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \\
&\equiv \mathcal{G}(\ell_1,\ell_2,\ell_3;m_1,m_2,m_3).
\end{aligned}
\tag{2.13}
$$

and remember that the Gaunt integral $\mathcal{G}$ is non-zero only if the selection rules are satisfied:

- $m + m_1 + m_2 = 0$,
- $\ell + \ell_1 + \ell_2$ must be even, due to parity,
- ℓ, ℓ_1, ℓ_2 satisfy the triangle condition: $|\ell_i - \ell_j| \leq \ell_k \leq \ell_i + \ell_j$.

The visibility equation can now be rewritten in the way that all components are happy with their mode definitions:

$$
V_{\mathrm{th}}(\boldsymbol{b}_{ij},\nu,t) = \sum_{\ell m} \mathbf{K}_{\ell m}(\boldsymbol{b}_{ij},\nu) T_{\ell m}(\nu) e^{imt} \tag{2.14}
$$

$$
\mathbf{K}_{\ell m}(\boldsymbol{b}_{ij},\nu) = \sum_{\ell_1 m_1}\sum_{\ell_2 m_2}\sum_{m'} B^{(ij)}_{\ell_1 m_1}(\nu) F_{\ell_2 m_2}(\boldsymbol{b}_{ij},\nu)\, \mathcal{D}^{(\ell)}_{m',m}(\mathcal{R}_0^{-1})\, \mathcal{G}(\ell_1,\ell_2,\ell;m_1,m_2,m') \tag{2.15}
$$

2.3 Data Covariance

We assume m-homogeneous statistics, where all $T_{\ell m}$ are independent realisations of the ℓ-dependent statistics,

$$
\langle T_{\ell m}(\nu) T^*_{\ell' m'}(\nu')\rangle = \delta_{\ell\ell'}\delta_{mm'} C_\ell(\nu,\nu'). \tag{2.16}
$$

Then the data covariance is straightforward to calculate:

$$\langle V_{\rm th}(\boldsymbol{b},\nu,t)V^*_{\rm th}(\boldsymbol{b}',\nu',t')\rangle = \sum_\ell \mathcal{K}_\ell(\boldsymbol{b},\nu,t;\boldsymbol{b}',\nu',t')\,C_\ell(\nu,\nu') \tag{2.17}$$

$$\mathcal{K}_\ell(\boldsymbol{b},\nu,t;\boldsymbol{b}',\nu',t') \equiv \sum_m \mathbf{K}_{\ell m}(\boldsymbol{b},\nu)\mathbf{K}^*_{\ell m}(\boldsymbol{b}',\nu')e^{im(t-t')} \tag{2.18}$$

Below are some important simplified cases:

1. If $t=t'$:

$$\mathcal{K}_\ell(\boldsymbol{b},\nu,t;\boldsymbol{b}',\nu',t) = \sum_{\ell_1 m_1}\sum_{\ell_2 m_2} B_{\ell_1 m_1}(\nu){B'}^*_{\ell_1 m_1}(\nu')F_{\ell_2 m_2}(\boldsymbol{b},\nu)F^*_{\ell_2 m_2}(\boldsymbol{b}',\nu') \cdot \frac{(2\ell_1+1)(2\ell_2+1)(2\ell+1)}{4}\begin{pmatrix}\ell_1 & \ell_2 & \ell\\ 0 & 0 & 0\end{pmatrix}^2, \tag{2.19}$$

   where we have applied the unitarity of the Wigner D-matrix and a $3j$-symbol orthogonality:

$$\sum_{m_3}\begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ m_1 & m_2 & m_3\end{pmatrix}\begin{pmatrix}\ell_1' & \ell_2' & \ell_3\\ m_1' & m_2' & m_3\end{pmatrix} = \frac{1}{2\ell_3+1}\delta_{\ell_1\ell_1'}\delta_{\ell_2\ell_2'}\delta_{m_1m_1'}\delta_{m_2m_2'}. \tag{2.20}$$

2. If $t=t'$ and $\boldsymbol{b}=\boldsymbol{b}'$, we can further reduce the sum over m_2 (actually we only need $\boldsymbol{b}\parallel\boldsymbol{b}'$ to reduce this sum):

$$\mathcal{K}_\ell(\boldsymbol{b},\nu,t;\boldsymbol{b},\nu',t) = \sum_{\ell_1 m_1}\sum_{\ell_2} B_{\ell_1 m_1}(\nu){B'}^*_{\ell_1 m_1}(\nu')\,\mathcal{F}_{\ell_2}(\boldsymbol{b},\nu,\nu') \cdot \frac{(2\ell_1+1)(2\ell_2+1)(2\ell+1)}{4}\begin{pmatrix}\ell_1 & \ell_2 & \ell\\ 0 & 0 & 0\end{pmatrix}^2, \tag{2.21}$$

   where $\mathcal{F}_{\ell_2}(\boldsymbol{b},\nu,\nu') = 4\pi(2\ell+1)j_\ell(u)j_\ell(u')$ and, again, u is the wave number given by the baseline and frequency.

3. If $t=t'$ and $\boldsymbol{b}\parallel\boldsymbol{b}'$ and the beam is rotationally symmetric, we can further reduce the sum over m_1:

$$\mathcal{K}_\ell(\boldsymbol{b},\nu,t;\boldsymbol{b},\nu',t) = \sum_{\ell_1}\sum_{\ell_2} B_{\ell_1,0}(\nu){B'}^*_{\ell_1,0}(\nu')\,\mathcal{F}_{\ell_2}(\boldsymbol{b},\nu,\nu') \cdot \frac{(2\ell_1+1)(2\ell_2+1)(2\ell+1)}{4}\begin{pmatrix}\ell_1 & \ell_2 & \ell\\ 0 & 0 & 0\end{pmatrix}^2. \tag{2.22}$$

2.4 Linear maps of signals and covariances

The linear mapping in eq. (2.17)

$$\mathcal{K}_\ell : C_\ell(\nu, \nu') \mapsto \text{data covariance} \tag{2.23}$$

can be used for defining a quadratic estimator of the delay power spectrum, which is the ν-space Fourier conjugate of $C_\ell(\nu, \nu')$

The linear mapping in eq. (2.10)

$$\mathcal{M}_{\ell m} : T_{\ell m} \mapsto \text{visibility data} \tag{2.24}$$

can be used for various purposes. It can also be thought of as a window function where the peak-ℓ is determined by the baseline. In the context of quadratic power spectrum estimation, it can be used to estimate the degree of the statistical independence of different estimates, understanding the mixing of different m-modes, given ℓ.

2.5 m-mode formalism for driftscan measurements

For drift-scan measurements, we see that the time dependence in the visibilities is simply a phase term, and the rest are the m modes in the above formalism:

$$V = \sum_{m=0}^{\ell_{\max}} V_m \, e^{imt} \qquad\qquad V_m(\boldsymbol{b}, \nu) = \sum_\ell \mathbf{K}_{\ell m}(\boldsymbol{b}, \nu) T_{\ell m}(\nu)$$

where $\mathbf{K}_{\ell m}$ is defined in eq (2.15), which can be precomputed. Then the linear system can be represented using these m-mode visibilities as basic components:

$$\vec{V} = \mathbf{U}\vec{V}_m, \tag{2.25}$$

where $\mathbf{U}$ represents the transformations in terms of e^{imt}. And the parameter space covariance of the m modes is independent of time:

$$\langle V_m(\boldsymbol{b}, \nu) V^*_{m'}(\boldsymbol{b}', \nu') \rangle = \sum_\ell \mathbf{K}_{\ell m}(\boldsymbol{b}, \nu) \mathbf{K}_{\ell m'}(\boldsymbol{b}', \nu')^* C_\ell(\nu, \nu'), \tag{2.26}$$

where the r.h.s is something like $\mathbf{KCK}^\dagger$ for each ℓ-block. We can further represent frequency angular spectrum as the Fourier transform of the delay spectrum:

$$\mathbf{E} = \sum_\ell (\mathbf{KFPF}^\dagger\mathbf{K}^\dagger)_\ell, \tag{2.27}$$

where $\mathbf{E}$ is the covariance matrix of the EoR m-mode visibilities, and $\mathbf{P}$ is a diagonal matrix which represents the delay spectrum.