Chapter 1

Theoretical background

Our goal, is to add the final steps in the theory chain which transforms the GRMHD simulations into interferometric observables. For this to be achieved and for the theory higher up in the chain to be maximally useful in data interpretation, realistic signal corruptions need to be considered. Hence, the purpose of this module is to further the sophistication of the interplay between theory and observation in the field. The signal corruptions which we have identified as the most prominent occurs in the troposphere, interstellar medium (ISM) and within the stations themselves. First I will review some EM wave fundamental and introduce scattering theory, which is applicable to both the radiative process occuring in the troposphere and ISM. Following the general introduction I will explore each specific case.

1.1 Radio Interferometry

1.1.1 Measurement Equation

1.1.2 mm-VLBI observables and data products

If the visibility phase is highly variable as in the case of a turbulent atmosphere, conventional calibration and imaging techniques have severely limited (if any) success. However information can still be extracted from the raw visibilities in the form of closure quantities [**?**] or polarisation ratios [**?**]. Closure phase, defined as the sum of 3 visibility phases of a triangle of stations $\{i, j, k\}$, is a probe of asymmetry in source structure,

$$\Phi_{ijk} = \phi_{ij} + \phi_{jk} + \phi_{ki}. \tag{1.1}$$

Because most signal corruptions are station based, the gain phase terms $\phi_{ij} = \phi^{\text{true}} + \phi_i^G - \phi_j^G$ for each antenna will cancel, yielding a more robust observable.

The uncertainty on the closure phase is model dependent [**?**] and is given as a function of the SNR s of each baseline

$$u(\Phi_{ijk}) = \frac{\sqrt{4 + (s_{ij}s_{jk})^2 + (s_{jk}s_{ki})^2 + (s_{ij}s_{ki})^2 + 2(s_{ij}^2 + s_{jk}^2 + s_{ki}^2)}}{s_{ij}s_{jk}s_{ki}}, \tag{1.2}$$

where s_{ij} is defined as

$$s_{ij} = |V_{ij}|\sqrt{\frac{\tau\Delta\nu}{SEFD_i SEFD_j}}, \tag{1.3}$$

where τ is the vector averaging timescale, $\Delta\nu$ is the bandwidth, $|V_{ij}|$ is the visibility amplitude and $SEFD$ is the system equivalent flux density. The result that closure phase is entirely immune to station based effects breaks down however when time averaging in the presence of baseline dependent effects like thermal noise as illustrated in section **??**.

1.1.3 Variability and the static source assumption

Implicit in our description of interferometry above, we assumed that the source remains approximately unchanged or static during the course of the observation. However, if this assumption does not hold (i.e. if the source is time-variable), the visibilities measured over the course of an observation can no longer be related to a single image and if they are, the resulting image would appear smeared out as it is averaged over many realisations. Note that I am using the term 'variability' in a general sense which refers to changes in any source observables. Most commonly variability is refers to changes in source flux (visibility amplitude) but I include changes in source structure and position (visibility phase) and source polarisation. Practically it is difficult to separate source and instrumental variability without accurate models and measurements for all non-source signal propagation effects. Although the static source assumption holds for most interferometric observations, the accretion flow and/or magnetic field structures around a SMBH can be variable on far shorter timescales. The primary mm-VLBI target, Sgr *, exhibits variability on timescales of minutes to hours in the radio (including in EHT observations), near-infrared (NIR), and X-ray bands [e.g. **? ? ? ? ? ?**]. This wealth of observational data has yielded several answers but the origin of the variability is still highly debated. To explain the observed delays between flares in different frequency bands, an expanding adiabatic plasma model (Marrone, 2008) has been presented however a recent flare observed with the EHT did not exhibit the increase in size expected from an expanding plasma outflow model [**?**]. Signatures of periodic variability at NIR and x-ray [**?**] have been used to argue for the presence of orbiting hotspots [**?**]. As the Innermost Stable Circular Orbit (ISCO) depends on spin of the BH, the spin can be constrain through the detection periodic orbital features. However a longer light curve in the NIR is more representative of a power-law scale variability [**?**]. These observations point to the possibility of multiple flaring mechanisms. An important mm-VLBI observational result is that variability in the polarisation domain is far more rapid than the total intensity (Johnson 2015b), indicating the presence of highly variable magnetic fields. In principle, the variability timescale can be comparable to the period of the Innermost Stable Circular Orbit (ISCO), which for Sgr A*, ranges from 4 minutes in the maximally rotating realisation to about half an hour for a non-rotating BH. The ISCO period for M87 is longer on the order of day scales. Considering light crossing times $\Delta t_{\rm cross}$, we can estimate the angular size θ of the emission region to be of order $\theta \sim \Delta t_{\rm cross} c/D_{\rm src}$, where c is the speed of light and $D_{\rm src}$ is the observer-source distance. Hence for Sgr * at a distance of 8.3 kpc

(Gillessen, 2009), for a flare of duration $\Delta t_{\rm cross} = 10$ min, which corresponds to scales of $15 R_{\rm Sch}$, further evidence of emission areas close the event horizon. If a flare is dominated by a localised variable structure, several approaches [**?** **?** **?**] show that EHT can track flaring structures with ~ 5 μ-arcsec precision using closure quantities and polarimetric ratios which could help map the spacetime around the BH. Alternatively **?**] show that a gaussian weighting scheme can be applied to mitigate the effects of variability and measure the quiescent structure although approach would downweight the longest baselines. However all of these approaches assume only gaussian thermal noise, guassian-blurring in the ISM and no fringe-fitting errors.

1.2 Signal Corruptions

1.2.1 Scattering basics

Millimetre wavelength radiation originating at the Galactic Centre is repeatedly scattered along the signal path to the Earth-based observer. The first occurrence is due to electron plasma in the interstellar medium (ISM) (**?** , **?**), while the second is due to poorly-mixed water vapour in the Earth's troposphere (**?** , **?**). It is essential that the effects of the scattering phenomena are understood for a rigorous calibration and interpretation of data. Towards this end, simulation modules approximating scattering in both media are implemented in MeqSilhouette. As an introduction to the separate descriptions of each, we review a simple scattering model.

An electro-magnetic wave is scattered when it passes through a medium with refractive index inhomogeneities. Following **?**], this effect can be modeled as a thin screen, located between source and observer planes and orientated perpendicular to the line-of-sight. The screen, indexed by coordinate vector $\mathbf{x}$, adds a stochastic phase $\phi(\mathbf{x})$ to the incoming wave at each point on the screen, yielding a corrugated, outgoing wavefront. We define the Fresnel scale as $r_{\rm F} = \sqrt{\lambda D_{\rm os}/2\pi}$, where $D_{\rm os}$ is the observer-scatterer distance, or the distance where the geometrical path difference $\frac{2\pi}{\lambda}(D_{\rm os} - \sqrt{D_{\rm os}^2 + r_{\rm F}^2}) = \frac{1}{2}$ rad.

To determine the resultant electric field at a point in the plane of the observer, indexed by coordinate vector $\mathbf{X}$, one has to take into account all possible ray paths from the screen to $\mathbf{X}$. To illustrate the model, a calculation of the scalar electric field generated by a point source, $\psi(\mathbf{X})$ yields the Fresnel-Kirchoff integral [**?**]

$$\psi(\mathbf{X}) = C \int_{\rm screen} \exp\left[i\phi(\mathbf{x}) + i\frac{(\mathbf{x} - \mathbf{X})^2}{2 r_{\rm F}} \right] \mathbf{dx}, \tag{1.4}$$

where C is a numerical constant.

The statistics of $\phi(\mathbf{x})$ can be described by a power spectrum or equivalently the phase structure function,

$$D_\phi(\mathbf{x}, \mathbf{x}') = < [\phi(\mathbf{x} + \mathbf{x}') - \phi(\mathbf{x})]^2 >, \tag{1.5}$$

where $\mathbf{x}$ and $\mathbf{x}'$ represent two points on the screen and $<>$ denotes the ensemble average.

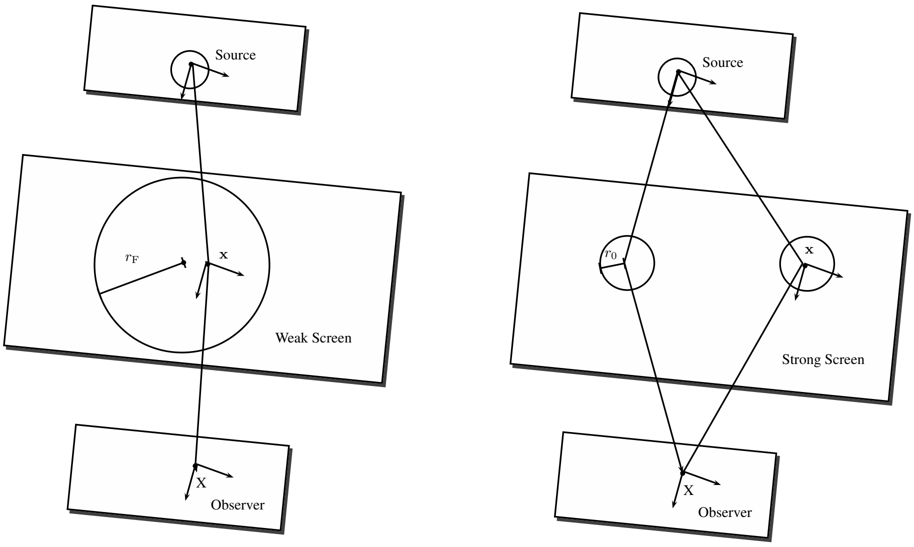

Figure 1.1: Illustration depicting the basics of scattering in the weak (left) and strong (right) regimes. In the weak regime, the signal is coherently propagated over an area, $A_{\rm weak} \approx \pi r_{\rm F}^2$ whereas in the strong regime, coherent propagation is split over many areas of size $A_{strong} \approx \pi r_0^2$.

There is evidence that D_ϕ can be reasonably approximated by a power law dependence on the absolute distance r between points on the screen [**?** **?**]

$$D_\phi(r) = (r/r_0)^\beta, \qquad r^2 = \mathbf{x}^2 - \mathbf{x}'^2 \tag{1.6}$$

where r_0 is the phase coherence length scale defined such that $D_\phi(r_0) = 1$ rad. Kolmogorov turbulence, which describes how kinetic energy injected at an outer length scale $r_{\rm out}$ cascades to increasingly smaller scales until finally dissipated at an inner length scale $r_{\rm in}$, predicts $\beta = 5/3$ in the domain $r_{\rm in} << r << r_{\rm out}$. This scaling has been demonstrated to be a reasonable approximation for the ISM over scales $r \sim 10^2$ km to > 1 AU [1], and also for the troposphere with $r < \Delta h$, where Δh is the thickness of the turbulent layer [**?**]. The specifics of the tropospheric model will be explored further in later sections.

The two length scales, $r_{\rm F}$ and r_0, define the nature of the scattering which is split into the strong and weak regimes, Fig. reffig:scatter. In weak scattering, $r_0 \gg r_{\rm F}$ and hence by equation **??**, $D_\phi(r_{\rm F}) \ll 1$. This implies that most of the radiative power measured on a point $\mathbf{X}$ will originate from a screen area $A_{\rm weak} \approx \pi r_{\rm F}^2$. Whereas in the regime of *strong scattering*, $r_0 \ll r_{\rm F}$ yielding $D_\phi(r_{\rm F}) \gg 1$. This results in coherent signal propagation onto the point $\mathbf{X}$ from multiple disconnected zones each of area $A_{strong} \approx \pi r_0^2$ [**?**]. Scattering in the troposphere and ISM fall into the regimes of weak and strong scattering respectively.

To evolve the screen in time, we assume a frozen screen i.e. that the velocity of the individual turbulent eddies is dominated by the bulk motion of scattering medium [e.g. **?**]. This allows us to treat the screen as frozen but advected over the observer by a constant motion. Hence time variability can

Table 1.1: A re-analysis of VLBI observations of Sgr A* by **?**] has yielded revised estimates of the parameters associated with the Gaussian scattering kernel. An accurate estimation is needed for accurate extrapolation to 1.3 mm. Note that the position angle is measured East of North.

major axis FWHM (mas/cm^{-}2)	1.32	0.04
minor axis FWHM (mas/cm^{-}2)	0.82	0.21
position angle ($^\circ$)	77.8	9.7

be easily incorporated by the relative motion between source, scattering screen and observer.

1.2.2 Interstellar medium scattering

Electron density inhomogeneities in the interstellar medium (ISM) plasma scatter the radio emission from the Galactic centre. Radio interferometric observations of Sgr A* have characterised the basic properties of the intervening plasma matrial, however extensive developments in scattering theory and simulations have proved essential to the interpretation of more subtle scattering phenomena. This section begins with the early observational results which studied the Gaussian blurring effect of the scattering; we then expand on the scattering theory introduced in Sec. **??** to review the latest theoretical developments which explore the presence of scattering-induced substructure; finally we review recent observational results which account for scattering substructure in their data interpretation.

The dominant observational effect of the scattering for $\lambda \gtrsim$ cm is to convolve the intrinsic source structure with an elliptical Gaussian. The size of the Guassian exhibits a λ^2 scaling dependence over several orders of magnitude [Fig. **??** **?** **?** **?** **?**],which is consistent with the wavelength dependence of the refractive index of a plasma. In order to determine the parameters of the scattering kernel, i.e. major axis, minor axis and position angle, one has to observe at wavelengths where the angular size of scattering ellipse is much larger than the expected source size. A Very Long Baseline Array (VLBA) + Green Bank Telescope (GBT) campaign [**?**] estimated the size at 1.31×0.64 mas cm^{-2}, oriented 78° east of north.

An accurate extrapolation of scattering kernel to 1.3 mm is important for the EHT scattering-mitigation strategy [**?**] which aims to deblur the scattered image through a deconvolution procedure. However as this extrapolation is over at least an order of magnitude, any small systematic error in the original measurement can significantly effect the 1.3 mm extrapolated parameters. A recent review of VLBI observations of Sgr A* [**?**] has noted that there are significant inconsistencies between different measurements. The authors used a Bayesian methodology to re-analyse the datasets resulting in increased uncertainties as shown in **??**. The minor axis has a much larger uncertainty than the major axis due to the limited north-south coverage of the VLBA.

The Gaussian blurring effect can explained by the simple scattering model introduced in Sec. **??**. Recall, that in the strong scattering regime light is propagated from coherent patches with linear size $\sim r_0$. Each patch will emit light coherently into a single-slit diffraction cone of angular size $\theta_{\rm scatt} \sim \lambda/r_0$.

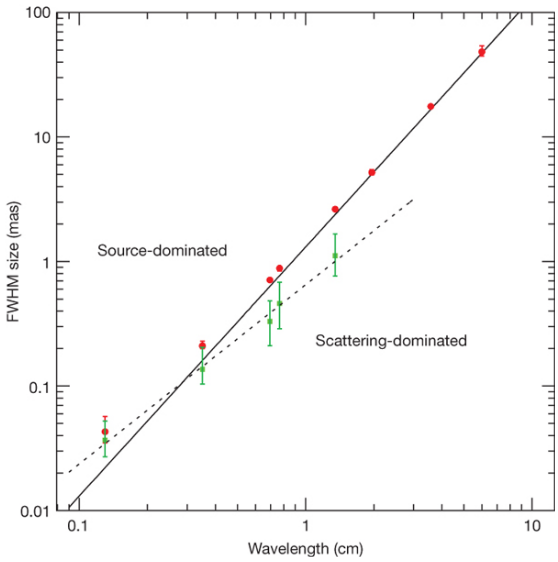

Figure 1.2: The λ^2 dependence of scattering kernel size is shown by the solid line. This has been derived from measurements made at $\lambda > 17$ cm [**?**]. The dotted line shows the derived intrinsic source size which scales as $\lambda^1.44$. This was derived from measurements in the wavelength range, 2 cm $< \lambda <$ 1.3 mm. The red circles show major-axis observed sizes of Sgr A* and the green points show the derived intrinsic major-axis size. This plot was reproduced from **?**].

An observer will hence be illuminated by many patches spanning θ_{scatt}, yielding a blurred and broadened image, with projected size on the screen equal to the *refractive scale*

$$r_{\mathrm{ref}} = \theta_{\mathrm{scatt}} D_{\mathrm{os}} = r_{\mathrm{F}}^2 / r_0.$$

r_{ref} is the third fundamental length scale in the strong scattering regime and is associated with the refractive timescale,

$$t_{\mathrm{ref}} = r_{\mathrm{ref}} / v.$$

We can calculate r_0 given the FWHM of θ_{scatt} through the more precise relation

$$\theta_{\mathrm{scatt}} = \frac{2\sqrt{2\ln 2}}{2\pi} \lambda / r_0 (M+1) \tag{1.7}$$

where $M = D_{\mathrm{os}}/R$ is the magnification and R is the source-screen distance. The magnification factor is a correction to the model introduced in Sec. **??** when $R \sim \infty$ no longer holds and should be used when calculating distances in the observer plane [**?**]. The location of the scattering medium was originally thought to be quite close to Sgr A*. However, observations of a newly discovered pulsar, SGR J1745-29, indicate that the scattering screen is located at a distance $D_{\mathrm{os}} = 5.8 \pm 0.3$ kpc, within the Scutum spiral arm. Using Eq. **??** and the parameters given in table **??**, we find that the major axis of the coherence length at 1.3 mm, $r_0 \approx 3136.67$ km.

As the VLBI moves to higher frequencies, focus has shifted away from the well-studied Gaussian convolution effect of ISM scattering and onto the presence of stochastic scattering-induced substructure. To understand this phenomenon, we must first develop our conceptual framework.

Strong scattering can be further subdivided into *snapshot*, *average* and *ensemble-average* regimes [**?** **?**]. To understand the different regimes, remember that for each point on the source, the observer sees emission from coherent patches of size $\sim r_0$ spanning $\sim r_{\mathrm{ref}}$. The diffraction cones from each of the patches will interfere, resulting in a multi-slit *diffractive scintillation* pattern.

In the *snapshot regime*, a compact source is observed with a narrow bandwidth and over a short time integration. This yields a single realisation of the diffractive scintillation pattern. By averaging over many snapshots, diffractive scintillation is quenched. This occurs if the source size θ_{src} is much larger than the diffractive scale $\theta_{\mathrm{src}} \gg r_0/R$; if the fractional bandwidth $\delta\nu/\nu$ is much larger than the decorrelation bandwidth $\delta\nu/\nu \gg \delta\nu_{\mathrm{dc}}/\nu \approx (r_0/r_{\mathrm{F}})^2$ [**?**]; or if the integration time t_{int} is much larger the diffractive timescale $t_{\mathrm{int}} \gg t_0 = r_0/v$, where v is the relative velocity between screen, source and observer. This regime is hence only accessible through observations of compact objects like pulsars. On a side note, observations in this regime can be used to probe the source with angular resolution given by the $\sim \lambda/r_{\mathrm{ref}}$ [e.g. **?**]. This is because the scattering screen is essentially a lens of size $\approx r_{\mathrm{ref}}$.

In the *average regime*, diffractive scintillation has been averaged over, however there still exists scintillation over scales comparable to the size of the scattered image of a point source $\sim r_{\mathrm{ref}}$, termed *refractive scintillation*. Phase fluctuations on this scale acts like a weak lens to focus or defocus the λ/r_0 scale diffraction cones in the direction of the observer. For a point source this would lead to weak flux variations in the total flux [**?**]. We will show later that refractive scintillation leads to the presence of substructure for a resolved scatter-broadened source. In contrast to diffractive scintillation, refractive scintillation is much more difficult to average over. Typically the refractive time scale $t_{\rm ref} = r_{\rm ref}/v$ is on the order of weeks to months for Sgr A*; the fractional decorrelation bandwidth is on the order of unity $\delta\nu_{\rm dc}/\nu \sim 1$; and the source has to be much larger than the image of a scattered point source $\theta_{\rm src} \gg \theta_{\rm scatt}$.

In the *ensemble-average regime*, both diffractive and refractive scintillation have been averaged over. It is in this regime when the scattering is equivalent to Gaussian convolution which is deterministic and not time variable.

A recent theoretical work [1] has derived a useful approximation of the resolved scattered image $I_{\rm ss}$ in the average regime,

$$I_{\rm ss}(\mathbf{x}) \approx I_{\rm src}\left(\mathbf{x} + r_{\rm F}^2 \nabla\phi(\mathbf{x})\right), \tag{1.8}$$

where ∇ is the directional derivative. Here we have used the same two-dimensional coordinate system, indexed by $\mathbf{x}$ to describe the source, screen and observer planes which are considered to be aligned along the vertical axis. The scattered image $I_{\rm ss}$ is approximated by a 'reshuffling' of the source image $I_{\rm src}$. As $|\nabla\phi| \sim 1/r_0$, the magnitude of the translation of points on $I_{\rm src} \sim r_{\rm ref} \sim 10$ μ-arcsec in the case of Sgr A*.

Even though $\phi(\mathbf{x})$ is only coherent to $\sim r_0$, the directional phase derivative $\nabla\phi(\mathbf{x})$ remains spatially coherent over much larger scales. Following **(author?)** [1], the autocovariance of phase derivative can be related to the structure function

$$\begin{aligned} \langle[\partial_x\phi(\mathbf{x_0})][\partial_x\phi(\mathbf{x_0}+\mathbf{x})]\rangle &= -\partial_x^2\langle\phi(\mathbf{x_0})\phi(\mathbf{x_0}+\mathbf{x})\rangle & (1.9)\\ &= \partial_x^2 D_\phi(\mathbf{x}). & (1.10)\end{aligned}$$

as $\langle\phi(\mathbf{x})^2\rangle = 0$

Using a typical structure function [**?**],We are interested in the case $r_{\rm in} \gg r_0$, the structure function becomes [**?**]

$$D_\phi = \begin{cases} \left(\frac{r}{r_0}\right)^2 & \text{if } r \ll r_{\rm in},\\ \frac{2}{\beta}\left(\frac{r_{\rm in}}{r_0}\right)^{2-\beta}\left(\frac{r}{r_0}\right)^\beta & \text{if } r \gg r_{\rm in}.\end{cases}$$

Note that the structure function is quadratic at small scales as fluctuations are smooth andthen kolmogorov then constant (Tatarskii, 1971). Hence We are interested in the case $r/ggr_{\rm in}$

$$\partial_x^2 D_\phi(\mathbf{x}) = \left(\frac{r_{\rm in}}{r_0}\right)^{2-\beta} 2(\beta-1)\frac{r^{\beta-2}}{r_0^\beta} \tag{1.11}$$

Therefore in the Kolmogorov regime $\beta = 5/3$, the coherence of image shift relative to the refractive scale $\propto (r/r_0)^{-1/3}$. Inner scale extends coherence and outer scale cuts it off.

If $\nabla\phi(\mathbf{x})$ was incoherent between patches of size $\sim r_0$, this would effectively be convolution with Gaussian of size $r_{\rm ref}$.)

A recent observation of Sgr A* at 3.5 mm by the VLBA+LMT [see Fig. reffig:substructure2 **?**] show that the closure phase measured is consistent with

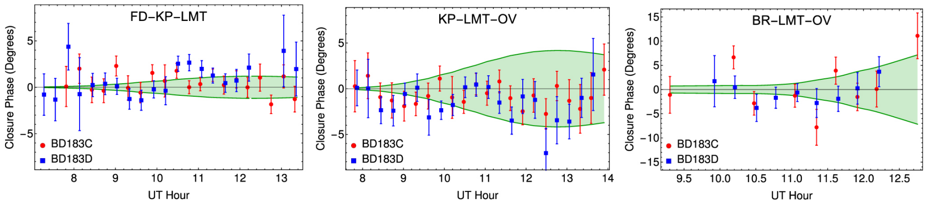

Figure 1.3: Closure phases recorded in a VLBA + LMT observation of Sgr A* at $\lambda = 3.5$ mm [?]. The data points are shown as red circles and blue squares and are only distinguished by the calibrator used. The green envelopes show the 1σ closure phase prediction induced by scattering-induced substructure. Reproduced from ?]

refractive scintillation. Another observation at 1.3 cm shows flux modulation due to scattering substructure $\sim$ 10 mJy [?] and other predictions show $\sim$ 60 mJy for long East-West baselines and $\sim$ 25 mJy for long North-South baselines [1], assuming a Gaussian source of $FWHM = 40$ μ-arcsec.

Distinguishing intrinsic source structure and variability and ISM variability is an interesting challenge. Observations at mm-wavelengths have revealed deviations from the λ^2 scattering scaling law, see Fig. ??. This is interpreted as due to the presence of intrinsic source structure and has been fitted with a power-law with an exponent of 1.34 ± 0.01 [?]. This has enabled the constraint of various theoretical models [?], excluding advection-dominated accretion flows (ADAF) [?] and Bondi-Hoyle accretion [?]. However observations extending over month timescales are required to properly sample the larger scale inhomogeneities and even with multiple epoch observations, it can be difficult to distinguish source and scattering characteristics [?].

Knowledge of the scattering characteristics can allow the two to be decoupled without sampling a refractive ensemble. It therefore provides a robust and rapid mechanism for quantifying refractive effects.

1.2.3 Troposphere

The coherence and intensity of millimetre wavelength electromagnetic waves are most severely deteriorated in the lowest atmospheric layer, the troposphere which extends up to an altitude of $7 - 10$ km above sea level and down to a temperature $T \sim 218$ K [?]. The troposphere is composed of a number of different components including primary gases N_2 and O_2, trace gases e.g. water vapour and CO_2, as well as particulates of water droplets and dust. The rest of this section will explore the tropospheric corruption for the mm-VLBI case.

Propagation fundamentals

Consider a quasi-monochromatic wave passing through a linear medium,

$$E_\nu(x,t) = E_0 \exp^{i(kn_\nu x - 2\pi\nu t)}, \qquad (1.12)$$

where $k = 2\pi\nu/c$ is the propagation constant in free space and $n = n_\mathrm{R} + jn_\mathrm{I}$ is the complex index of refraction. Note that we will occasionally omit the frequency dependence of n and related quantities to simplify the notation. If n_I is nonzero, the electric flux I will decay exponentially

$$I = EE^* = E_0^2 \exp(-\tau), \tag{1.13}$$

where τ is called the opacity or optical depth and is related to the absorption coefficient $\kappa = 4\pi\nu n_I/c$ via $d\tau = \kappa dx$. If $n_\mathrm{R} > 1$ the phase velocity of light will decrease, $v_\mathrm{p} = c/n_\mathrm{R}$, which results in a time delay. The time delay, δt and opacity τ can be calculated simultaneously,

$$\delta t + i\tau/4\pi\nu = 1/c \int_{path} d\mathbf{s}\ (n_\nu(\mathbf{s}) - 1). \tag{1.14}$$

Opacity and time delays are often viewed independently, however the electric field is real and causal and this imposes restrictions on the complex refractive index. Specificially n_R and n_I contain the same information and can be interchanged via the Kramers-Kronig relations.

Absorption is accompanied by emission and for a medium in local thermodynamic equilibrium, Kirchoff's law states that

$$\frac{\epsilon_\nu}{\kappa_\nu} = B_\nu(T), \tag{1.15}$$

where $\epsilon_\nu = dI_\nu/dx$ is the emission coefficient and $B_\nu(T)$ is the Planck function. Hence the absorbing molecules are also emitters, increasing system noise. Therefore opacity, time delay and atmospheric noise are interrelated and should be simulated consistently. On a side note this opens up the possibilities for phase calibration using measurements of sky emission via Water Vapour Radiometry (WVR).

Atmospheric corruptions in the (sub-)mm regime

An analysis of the absorption spectrum in the GHz range (Fig. **??**), shows that it is dominated by transitions of H_2O and O_2 as well as a pseudo-continuum opacity which increases with frequency. The pseudo-continuum opacity is due to the cumulative effect of the far wings of a multitude of broadened water vapour lines above 1 THz [**?**]. At 230 GHz the absorption is typically $5 - 10\%$ at the best sites, during good weather.

In contrast to the dry atmospheric components, water vapour mixes poorly and its time-variable spatial distribution induces rapid fluctuations in the time delays $\delta t(\nu)$ above each station. The phase error for a baseline (1,2) where antenna 1 is the reference will be

$$\delta\phi(t,\nu) = (\delta t_2(t,\nu) - \delta t_1(t,\nu))/\nu. \tag{1.16}$$

The water vapour column density is measured as the depth of the column when converted to the liquid phase and is referred to as the precipitable water vapour (PWV). PWV is directly proportional to the time delay and hence the phase delay,

$$\delta\phi \approx \frac{12.6\pi}{\lambda} \times w, \tag{1.17}$$

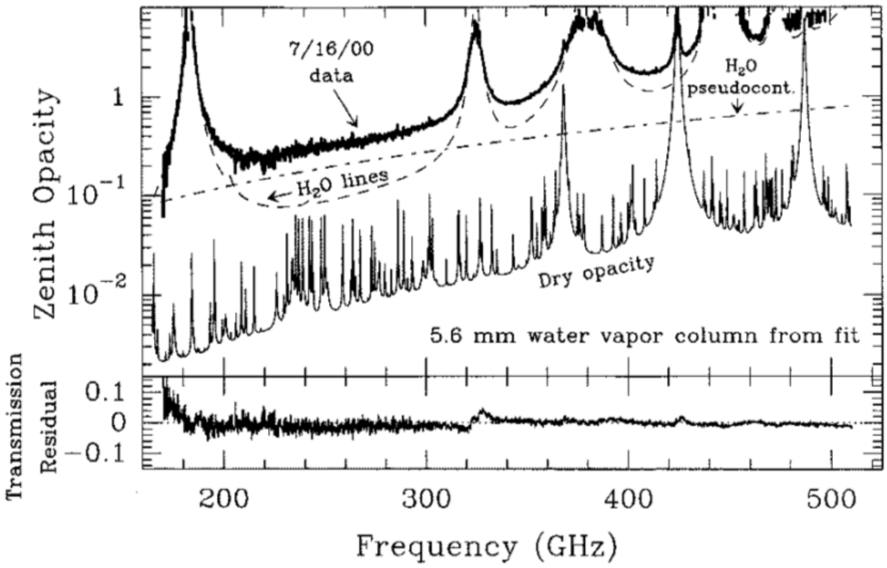

Figure 1.4: Recorded zenith absorption spectrum in the 160 − 520 GHz range, taken on Mauna Kea at an altitude of ∼ 4000 m. The data has been fit to a sum of H_20 lines, an H_20 pseudo-continuum and dry absorption lines. The model has been generated using the ATM code, with the bottom panel showing the residuals. Here 'dry' refers to all atmospheric constituents except H_20. Reproduced from **?**]

where w is the depth of the PWV column [**?**] and an atmospheric temperature $T = 270$ K has been assumed. This relationship between phase and water vapour content has been experimentally verified [**?**]. At 230 GHz, the change in PWV needed to offset the phase by 1 rad is $\Delta w \approx 0.03$ mm.

This sensitive dependence of phase coherence on atmospheric stability is aggravated by typically low antenna observation elevation angles as the atmospheric path length is increased; uncorrelated atmospheric variations between stations as correlated atmospheric variations fall away; and observing with a sparse VLBI array as this leads to less redundancy for calibration.

Radiative transfer

The problem of radiative transfer through a static atmosphere is well described and implemented by the Atmospheric Transmission at Microwaves (ATM) code [**?**]. Here we provide a brief summary of the theory underpinning the package but see the original paper for further detail. ATM is commonly used in the Atacama Large Millimeter Array (ALMA) community [**? ?**] and has been tested with atmospheric transmission spectra taken on Mauna Kea [**?**].

We start from the unpolarised radiative transfer equation, which is unidirectional in the absence of scattering,

$$\frac{dI_\nu(s)}{ds} = \epsilon_\nu(s) - \kappa_\nu(s) I_\nu(s), \tag{1.18}$$

where s is the coordinate along the signal path through the atmosphere. The assumption of local thermodynamic equilibrium (LTE) holds as the collisional timescale is much smaller than the time for spontaneous emission up to altitudes ≥ 80 km, after which there is only $\sim 0.001\%$ of mass left [**?**]. Applying **??**, multiplying by $\exp(-\tau_\nu)$ and integrating from the top of the atmosphere ($s = 0$) yields,

$$I_\nu(s) = I_\nu(0)e^{-\tau_\nu(0,s)} + \int_0^s B_\nu(s')e^{-\tau_\nu(s',s)}\kappa_\nu(s')ds', \tag{1.19}$$

where s' is a dummy variable in the same direction as s and $\tau_\nu(0,s) = \int_0^s k_\nu(s')ds'$. $I_\nu(0)$ is normally taken as the radiance from the cosmic background. **??** will allow us to calculate the noise temperature of the atmosphere by converting the output radiance at the ground $I_\nu(s)$ to the equivalent blackbody temperature through inversion the Planck function. To calculate the opacity and complete the above integral, κ_ν needs to be calculated over the frequency range. The time delay δt can be calculated using the Kramers-Kronig relations.

A general equation for the absorption coefficient for a transition between a lower l and upper u states is given in the original paper. Here we merely point out that it should be proportional to the energy of the photon, $h\nu_{l\to u}$, the transition probability or Einstein coefficient, $B_{l\to u}$, the line-shape, $f(\nu, \nu_{l\to u})$ and the number densities N of electronic populations. Line profiles which describe pressure broadening (perturbations to the Hamiltonian due to the presence of nearby molecules) and Doppler broadening are used. The condition of detailed balance further requires that decays from the upper state are included yielding, $g_u B_{u\to l} = g_l B_{l\to u}$, where g is the degeneracy of the electronic state. Putting this together we find,

$$\kappa(\nu)_{l\to u} \propto h\nu B_{l\to u}\left(\frac{N_l}{g_l} - \frac{N_u}{g_u}\right) f(\nu, \nu_{l\to u}), \tag{1.20}$$

where the Einstein coefficients are calculated from the inner product of the initial and final states with the dipole transition operator,

$$B_{l\to u} = \frac{2\pi}{3\hbar^2} | < u|\mu|l > |^2, \tag{1.21}$$

where $|u>$, $|l>$,$|\mu>$ are the wavefunctions of upper and lower states and the dipole transition operator respectively. The number densities of the two states, N_u and N_l in local thermodynamic equilibrium (LTE) are simply related to the local number density and temperature via Boltzmann statistics,

$$\frac{N_n}{N} = g_n \frac{\exp -\frac{E_n}{kT}}{Q} \tag{1.22}$$

where Q is the partition function. $Q = \sum_i g_i \exp -E_n/kT$.

Physically, the lineshape originates from perturbations to the hamiltonian due to proximity to neighbouring molecules, called pressure broadening, and at lower pressures, thermal doppler broadening. A Van Vleck -Weisskopf (VVW) profile is used for pressure broadening. At lower pressures this is convolved with a gaussian which arises from the Maxwellian distribution.

Transition lines at radio wavelengths result from rotational state transitions. To calculate the inner product given in equation **??**, hamiltonians for linearly symmetric rotors (e.g. O_2, CO) and asymetric rotors are used. The asymetric rotations are decomposed into three principal rotation axes with differing rotational constants governing each axis. Rotational constants were measured by the authors as well as drawn from a variety of literature. Partition functions and transition probability are calculated using approximations taken from the literature.

Far wing broadening of H_2O lines > 1.2 THz extends to lower frequencies and is not completely represented by the VVW profile. This is believed to be due to self-self collisions of water molecules. Additionally there are terms from the dry atmosphere related to transient dipoles and Debye absorption which are not represented in the lineshape. To correct for these effects two pseudocontina are used, which take the form of a power law dependence on frequency, temperature and the molecular densities.

Turbulent phase fluctuations

Visibility phase instability $\delta\phi(t)$ due to tropospheric turbulence is a fundamental limitation to producing high fidelity, science-quality maps with a mm-VLBI array [**?**]. The coherence time-scale is typically too rapid ($\lesssim 10$ s) for fast switching calibration, so other calibration procedures (e.g. water vapour radiometry, paired antennas, and/or self-calibration) must be performed. Self-calibration is the most commonly used but is limited by the integration time needed to obtain adequate SNR to fringe fit. Phase decoherence often leads to the use of closure quantities to perform model fitting [**? ? ?**].

At centimeter wavelengths and below, water vapour in the troposphere dominates errors in delay rate over hydrogen maser clock errors. Stochastic phase errors lead to a decrease in measured flux due to incoherent vector averaging. This makes it difficult to obtain adequate SNR for calibration. The uncertainty in the phase degrades structural information which makes conventional imaging difficult and hence closure quantities are often used (e.g. Fish et al 2011). Incoherent phases represent a fundamental limit to all types of interferometry.

Following from section **??**, we can model the statistics of $\delta\phi(t)$ with a thin, frozen, Kolomogorov-turbulent phase screen moving with a bulk velocity, v. We set the height h of the screen at the water vapour scale height of 2 km above ground. We will show later that the thickness Δh of the atmospheric turbulent layer can be neglected in our implementation. At 1.3 mm, the Fresnel scale is $r_F \approx 0.45$ m and experiments show annual variations of $r_0 \sim 50 - 500$ m above Mauna Kea [**?**] and $r_0 \sim 90-700$ m above Chajnantor [**?**], where both sites are considered to have excellent atmospheric conditions for millimetre astronomy. As $r_F < r_0$, this is an example of weak scattering.

The required field-of-view (FoV) of a global mm-VLBI array is typically FoV < 1 mas or ~ 10 μm at a height of 2 km, which is roughly 7-8 orders of magnitude smaller than the tropospheric coherence length. The tropospheric corruption can therefore be considered constant across the FoV and, from the perspective of the Measurement Equation, modeled as a diagonal Jones matrix per time and frequency interval. As VLBI baselines are much longer than the coherence length, $|\mathbf{b}| \geq 1000$ km $\gg \mathrm{r}_0$, the phase screen at each site must be simulated independently.

Our aim then is to produce a phase error time sequence $\{\delta\phi(t_i)\}$ for each station which is added to the visibility phase. We invoke the frozen screen assumption and write the structure function as a function of time, $D(t) = D(r)|_{r=vt}$. The temporal structure function $D(t)$ provides an efficient route to sample the variability of the troposphere at the typical integration time of the dataset, $t_{\mathrm{int}} \sim 1$ sec.

The temporal variance of the phase is a function of the temporal structure function, and accounting for time integration yields [see **?** , B3]

$$\sigma_\phi^2(t_{\mathrm{int}}) = (1/t_{\mathrm{int}})^2 \int_0^{t_{\mathrm{int}}} (t_{\mathrm{int}} - t) D_\phi(t) dt. \tag{1.23}$$

Assuming power-law turbulence and integrating yields,

$$\sigma_\phi^2(t_{\mathrm{int}}) = \left[\frac{1}{\sin\theta(\beta^2 + 3\beta + 2)}\right] \left(\frac{t_{\mathrm{int}}}{t_0}\right)^\beta, \tag{1.24}$$

where $t_0 = r_0/v$ is the coherence time when observing at zenith and $1/\sin\theta$ is the approximate airmass which arises as $D_\phi \propto w$. As $r \ll \Delta h$, where Δh is the thickness of the turbulent layer, an thin screen exponent of $\beta = 5/3$ is justified [**?**]. The phase error time-series takes the form of a Gaussian random walk per antenna. At mm-wavelengths, the spectrum of water vapour is non-dispersive up to a few percent [**?**] and so we can assume a simple linear scaling across the bandwidth.

Phase fluctuations $\delta\phi(t)$ can also be simulated by taking the inverse Fourier transform of the spatial phase power spectrum. However this approach is much more computationally expensive, e.g. for an observation length t_{obs} involving $N_{\mathrm{ant}} = 8$ independent antennae with dish radii $r_{\mathrm{dish}} = 15$ m, wind speed $v = 10$ m s^{-1} and pixel size equal to r_{F}, the number of pixels $N_{\mathrm{pix}} \approx$

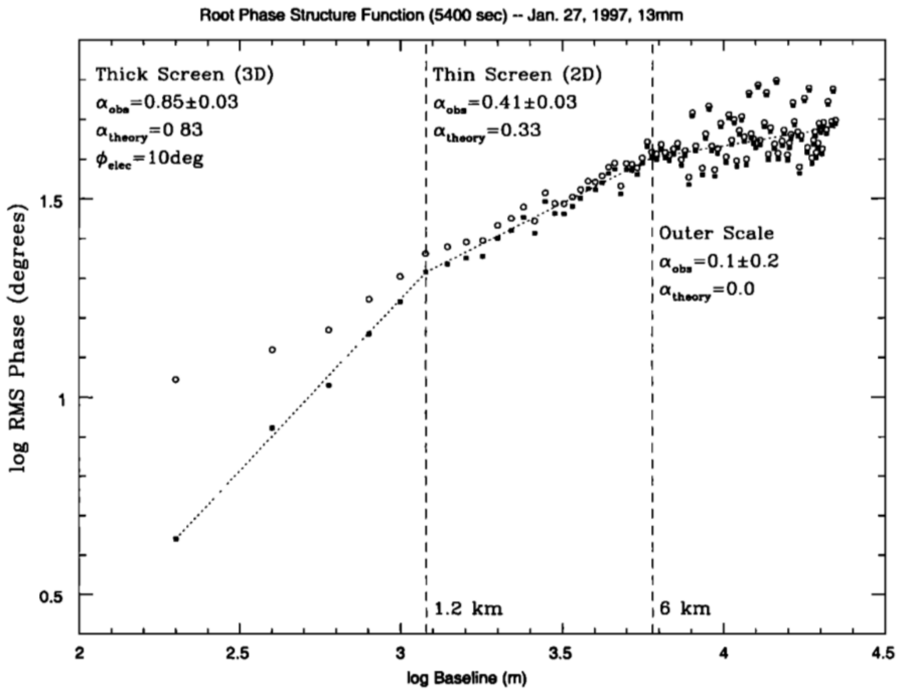

Figure 1.5: RMS visibility phase variations versus baseline length for an observation of 1-Jy celestial calibrator 0748 + 240 with VLA in BnA configuration at 22 GHz for 90 min.The open circles show the rms phase as measured whereas the solid squares show these same values with a constant noise term of 10 subtracted inquadrature.showing β changes with distance on the phase screen. Note the three distinct regime but continous variation in betweenThe three regimes of the root phase structure function as predicted by Kolomogorovturbulence theory are indicated. Reproduced from **?**]

$N_{\rm ant} t_{\rm obs} r_{\rm dish}^2/(v r_{\rm F}^3) \sim 10^8$. Additionally, due to fractal nature of ideal Kolmogorov turbulence, the power spectrum becomes unbounded as the wavenumber approaches zero which makes it difficult to determine the sampling interval of the spatial power spectrum [**?**].

1.2.4 Instrumental

Thermal Noise

The level of thermal noise in the data will define the sensitivity of the interferometer to detect a source and also to distinguish fine source characteristics. Closure quantities are especially prone to high levels of thermal as multiple visibilities are combined. A derivation of the thermal noise of an interferometer can be made through derivation of the thermal noise of an antenna and then correlating the result [**?**]. The RMS thermal noise of an interferometer $\{i, j\}$ over a bandwidth $\Delta\nu$ and an integration time is given by

$$\Delta S_{\rm ij} = \frac{1}{\eta_{\rm s}} \sqrt{\frac{SEFD_{\rm i}\ SEFD_{\rm j}}{2\Delta\nu t_{\rm int}}}, \tag{1.25}$$

where $\eta_{\rm s}$ is the system efficiency and $2\Delta\nu t_{\rm int}$ is the number of independent samples. The $SEFD$ is a measure of the sensitivity of an antenna, accounting for the effiency, collecting area and thermal noise and is defined as the flux density of a source with the same power,

$$SEFD = 2k_{\rm B} T_{\rm sys}/(\eta_{\rm a} A), \tag{1.26}$$

where A is the antenna area, $\eta_{\rm a}$ is the antenna efficiency, $T_{\rm sys}$ is the system temperature and the factor $\frac{1}{2}$ accounts for only sampling 1 polarisation.

Antenna Pointing

All antennas suffer pointing errors to some degree due to a variety of factors including dish flexure due to gravity, wind and thermal loading, as well as drive mechanics. This corresponds to an offset primary beam, which should only translate to minor amplitude errors if the pointing error $\theta_{\rm PE}$ is significantly smaller than the primary beam (i.e. $\theta_{\rm PE} \ll \theta_{\rm PB}$). In the Measurement Equation formalism, this offset can be represented by a modified (shifted) primary beam pattern in the E-Jones term

$$\mathbf{E}_p(l, m) = \mathbf{E}(l_0 + \delta l_p, m_0 + \delta m_p), \tag{1.27}$$

where $\delta l_p, \delta m_p$ correspond to the directional cosine offsets. This could be a problem for millimetre observations as the primary beam is significant, e.g. for a 30 m dish at 1.3 mm, $\theta_{\rm PB} \sim 10$ arcsec, compared to the pointing error which is on the order of arcseconds. A standard beam model which we will make use of later is the analytic WSRT beam model [**?**]

$$E(l, m) = \cos^3(C\nu\rho), \qquad \rho = \sqrt{\delta l_p^2 + \delta m_p^2} \tag{1.28}$$

where C is a constant, with value $C \approx 130$ GHz^{-1}. Note that the power beam EE^H becomes $\cos^6$.

An antenna tracking a source will suffer a slow, continuous time-variable pointing error associated with the tracking error σ_{track}. Physically this could be attributed to changes in wind, thermal and gravitational loading which all change with telescope pointing direction and over the course of a typical few hour observation. Using the MeqTrees software package, such behaviour has been demonstrated to occur with the Westerbork Synthesis Radio Telescope (WSRT, [**?**])[1].

Whilst a stationary phase centre is tracked, the pointing error should evolve slowly and smoothly, however, in mm-VLBI observations the phase centre is often shifted to another source/calibrator. This would cause the pointing error to change abruptly, with an absolute pointing error $\sim \sigma_{\text{abs}}$. Source/calibrator change is scheduled every 5-10 minutes in a typical millimetre observation. The point is that even though EHT will be able to determine the pointing offset when observing a calibrator with well known structure, when the antennas slew back to a source (e.g. Sgr A*) with less certain or variable source structure, the pointing error could change significantly. This is exacerbated by the scarcity of mm-wavelength calibrators, which are often widely separated from the source.

Polarisation leakage

[1]See also https://indico.skatelescope.org/event/171/session/9/contribution/20

Chapter 2

Software implementation

2.1 Data simulation

2.1.1 Design objectives

Our primary aim is to test and research mm-VLBI calibration, imaging and parameter estimation algorithms/strategies through the construction of a synthetic data simulation framework. To address the many questions within the wide scope of this objective, one must be able to setup and run a diversity of experiments within the simulation framework. This places definite constraints on the software architechure. In particular, the framework should

- enable the implementation of all relevant classes of signal corruption within a formalism which ensures consistency with the causal signal transmission chain,
- be compatible with time-variable GRMHD source models which are to be used as inputs,
- enable the construction and execution of arbitrary observations,
- be organised in modularised structure so that it is flexible, extendable and could be incorporated by other interferometric algorithms e.g. a calibration or a parameter estimation algorithm.

2.1.2 Architechure and Workflow

In this section, we will review how the architechural design and workflow of the simulator architechure has been designed to meet the above objectives. To fulfill the first objective, we try to cast signal corruptions in the RIME formalism (see section **??**), and where this is not possible, to fit those particular signal corruptions into the casually correct position in the signal transmission chain, with proper consider given to non-communitivity of elements in the signal transmission path. The implementation of each signal corruption is described in the following subsections. The remaining objectives fall into the realm of software design and will be discussed in this subsection.

We have chosen to write the top level simulation code using the PYTHON language. PYTHON is a general purpose language, is geared towards readability,

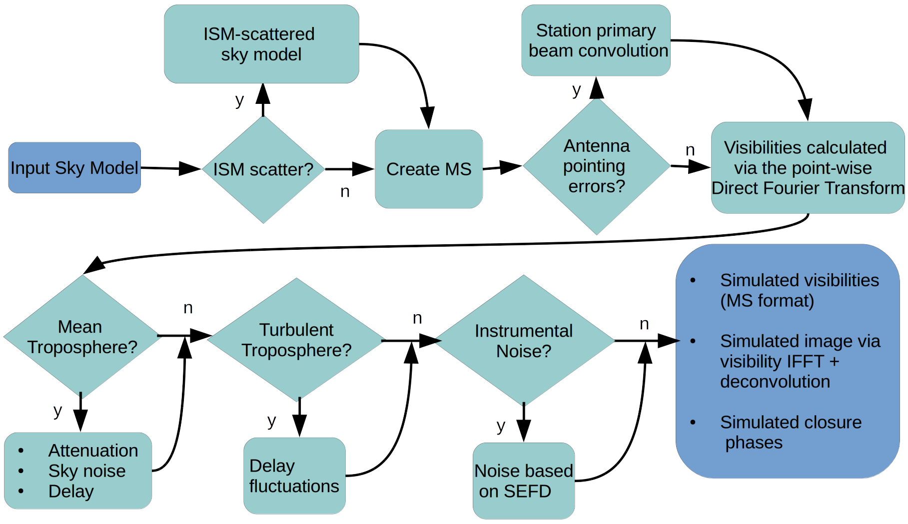

Figure 2.1: Flow diagram showing basic sequence of the MeqSilhouette simulation pipeline. The sky model could include (a) a time-ordered list of FITS images or (b) parametric source model consisting of Gaussians or point sources. The details of the station information, observation strategy, tropospheric and ISM conditions are specified in a user-defined input configuration file. The pipeline is flexible, allowing any additional, arbitrary Jones matrices to be incorporated. Further details in text.

and is well supported by a comprehensive library and wide user base (including astronomers). Specifically Python interfaces well with a modern interferometric toolbox, MeqTrees, as well as our data formats of choice: FITS for image cubes and the measurement set[1] MS for visibilities. Although the higher level functionality is written in Python, the bulk of the computational load (MS and visibility generation) is called through the faster C++ language. We use MS as our data format as it is directly accessible via the pyrap library and is the data format used by MeqTrees which performs the visibility generation and pointing error simulation. Although in the mm-VLBI subfield other data formats are currently still more popular than the MS, i.e. UVFITS or IOFITS, with the completion of ALMA, the MS format should become the next modern data format and already is used at the Joint Institute for VLBI in Europe (JIVE).

A conceptual flow diagram of a MeqSilhouette simulation pipeline is shown in Fig. 1.1. Note that the framework is not restricted to this sequence of operations and allows for the exact pipeline to be quite general. This flexibility is made possible through the use of *Object-Orientation*, which will be elaborated on later.

At highest level of the framework is a driver script and parameter dictionary. The driver script details the sequence of steps to take place in the pipeline whereas the parameter dictionary specifies all parameters needed by the pipeline to determine the particular observation configuration (array, fre-

[1]https://casa.nrao.edu/Memos/229.html

quency, bandwidth, start time, etc), which signal corruption implementation should be employed and where the sky model is located. The sky model is typically a time-ordered list of FITS images, where each image represents the source total intensity[2] over a time interval $\Delta t_{\mathrm{src}} = t_{\mathrm{obs}}/N_{\mathrm{src}}$, where t_{obs} is the observation length and N_{src} is the number of source images. A typical parameter dictionary is shown in table **??**.

The first major step in the pipeline is to create a fairly comprehensive MS. This is performed using the SIMMS[3] tool. SIMMS provides an easy to use interface to construct general MS, given an appropriate antenna table, which can be generated using PYRAP.

In order to make the framework as clean and modular as possible we have made extensive use of object orientation. The first major class, *SimpleMS*, was intended to abstract and modularise the MS and MS-only derived attributes (e.g. visibility data and station positions) and methods (e.g. functions to calculate station elevations and closure phases) as well as expose these attributes and methods more efficiently than following PYRAP procedures which become verbose when used frequently.

The second MS-related class, *TropMS*, handles the calculations relevant to tropospheric and thermal noise corruptions. This class is a child of *SimpleMS* and is initialised with weather and station information. Note that a child contains all the methods and attributes of its parent. This allows the tropospheric corruption implementation to use, whilst being separated from, the core MS functionality. The details of the tropospheric corruption is provided in a later subsection.

The third MS-related class is the *SimCoordinator*, and it is a child of the *TropMS* class. *SimCoordinator* is designed to make arbitrary simulations easy and efficient to construct and execute on a high level. It is the only MS class directly initialised in the driver script and hence the low level functionality and attributes of its parents are abstracted from the user. In addition to inherited functionality, *SimCoordinator* can call the ISM-scattering task (the implementation of which is in the next subsection), and the generation of visibilities and simulation of antenna pointing errors using the MEQTREES : TURBO-SIM task, where the visibilities are calculated through evaluation of the Fourier Transform at each UVW coordinate in the dataset, the time and frequency resolution of which is specified by the user.

The primary outputs of the pipeline are an interferometric dataset in MS format along with the closure phases and uncertainties and a dirty and/or deconvolved image (or spectral cube if desired. The modular structure of the pipeline allows for imaging and deconvolution algorithms to be employed. There exists a CASA task for conversion to UVFITS. As the pipeline is easily flexible other data products can be easily produced as needed e.g. polarisation ratios or time-frequency averaged data.

ISM scattering

As described in section **??**, ISM scattering towards Sgr A* falls into the *average regime*, wherein diffractive scintillation is averaged out but refractive scintillation is still present. As mm-VLBI observations can resolve the scatter-broadened

[2]Later versions of MEQSILHOUETTE will enable the full Stokes cubes as input.

[3]https://github.com/radio-astro/simms

image of Sgr A*, an implementation of scattering is needed which approximates the subtle changes in its extended source structure. Such an approximation has been implemented in the PYTHON-based SCATTERBRANE[4] package, and is based on **(author?)** [1]. The algorithm generates a phase screen based on the two dimensional spatial power spectrum [see 1, Appendix C] which incorporates inner and outer turbulent lengths scales and then implements **??** using an interpolation function modified by the phase screen.

SCATTERBRANE allows variation in all parameters associated with the scattering screen which is essential as aspects of the scattering towards the galactic centre is still unconstrained.

We include the SCATTERBRANE software, which has already yielded important context for mm-VLBI observations towards Sgr A* [e.g. 2], within the MEQSILHOUETTE framework. Our ISM module interfaces the SCATTERBRANE code within an interferometric simulation pipeline. This module enables simultaneous use of time-variable ISM scattering and time-variable intrinsic source structure within a single framework. The user is able to select a range of options relating to the time-resolution and epoch interpolation/averaging of both. By default, if the time resolution chosen to sample the source variability $\Delta t_{\rm src}$ and screen variability $\Delta t_{\rm ism}$ are unequal, we set

- $\Delta t_{\rm ism} = \Delta t_{\rm src} \quad \text{if} \quad \Delta t_{\rm src} < \Delta t_{\rm ism}$
- $\Delta t_{\rm ism} = R(\frac{\Delta t_{\rm src}}{\Delta t_{\rm ism}})\Delta t_{\rm src} \quad \text{if} \quad \Delta t_{\rm src} > \Delta t_{\rm ism},$

where R rounds the fraction to the nearest integer. This modification to the ISM sampling resolution avoids interpolation between different snapshots of the intrinsic source structure.

Atmospheric corruption simulator

Pointing error simulator

2.1.3 RODRIGUES interface

For community use, we host the online, RODRIGUES, interface, found at http://rodrigues.meqtrees.net/. Each of the components of the simulator run in Docker containers. **Looks like the infrustructure is going to change, re: discussions with Gijs and Sphe, so going to wait before writing this.

2.2 Parameter estimation

[4] http://krosenfeld.github.io/scatterbrane

Bibliography

[1] Michael D. Johnson and Carl R. Gwinn. THEORY AND SIMULATIONS OF REFRACTIVE SUBSTRUCTURE IN RESOLVED SCATTER-BROADENED IMAGES. *ApJ*, 805(2):180, jun 2015.

[2] G. N. Ortiz-León, M. D. Johnson, S. S. Doeleman, L. Blackburn, V. L. Fish, L. Loinard, M. J. Reid, E. Castillo, A. A. Chael, A. Hernández-Gómez, D. Hughes, J. León-Tavares, R.-S. Lu, A. Montaña, G. Narayanan, K. Rosenfeld, D. Sánchez, F. P. Schloerb, Z.-q. Shen, H. Shiokawa, J. SooHoo, and L. Vertatschitsch. The Intrinsic Shape of Sagittarius A* at 3.5-mm Wavelength. *ArXiv e-prints*, jan 2016.