Mémoire

Auteur : Furst, Théo
Promoteur(s) : De Becker, Michael; 17353
Faculté : Faculté des Sciences
Diplôme : Master en sciences spatiales, à finalité approfondie
Année académique : 2021-2022
URI/URL : http://hdl.handle.net/2268.2/14770

FACULTY OF SCIENCES

DEPARTMENT OF ASTROPHYSICS, GEOPHYSICS AND OCEANOGRAPHY

High resolution radio imaging of colliding-wind massive binaries using Very Long Baseline Interferometry

Théo Furst

A thesis presented in order to obtain the Master degree in Space Sciences

Supervisors : Michaël DE BECKER
University of Liege
Benito MARCOTE
Joint Institute for VLBI ERIC, JIVE

Reading Committee : Valentin CHRISTIAENS
University of Liege
Michaël DE BECKER
University of Liege
Marc-Antoine DUPRET
University of Liege
Benito MARCOTE
Joint Institute for VLBI ERIC, JIVE

Academic year 2021-2022

Acknowledgements

First and foremost, I would like to express a huge thank you to my supervisor, Mister Michaël De Becker, without who this thesis would not have been possible. I would like to firstly thank you for proposition of the thesis that I enjoyed from the very first moment to the end. Then, for the exceptional constant support that I received all along this year as well as the very instructive and profitable discussions that we had but also for your availability and the time that you spent reading my work and helping me to understand to science of Particle-Accelerating Colliding-Wind Binaries and much more.

Secondly, I would like to express all my gratitude to my co-supervisor, Mister Benito Marcote, who taught me how to calibrate radio data as well as many other things concerning radio interferometry. I would like to thank you also for the welcome you have given me during my stay at the JIVE Institute at which I learned so much thanks to discussions that I had with you but also with members of your team. I would like to finally thank you for all the time you granted me whether it was for the reading of my work or for answering my questions after my stay at the JIVE Institute.

Then, I would like to thank my family, who have supported me during those five years. You allowed me to find what I want to do in my life and to pursue those studies. I would also like to thank Magali and my friends, notably Jakob, for the support they provided me during this year as well as all the discussions that I had with them to exchange ideas and to learn from each other.

Finally, I would like to thank Mister Valentin Christiaens, Mister Michaël De Becker, Mister Marc-Antoine Dupret and Mister Benito Marcote for being part of the reading committee and the time they granted me for the reading of this thesis.

Contents

1 Introduction

1.1 Massive stars

Massive stars or also called early-type stars, are stars characterized by a mass greater than 8 $M_\odot$. Those stars are playing a major role in the chemical composition of the Milky Way. Indeed, because of their important mass, they have a short lifetime and they are led to experience a lot of nuclear reactions. As a consequence, massive stars are an important source of heavy elements in the interstellar medium (Jamet et al., 2004). In addition, because of their significant masses, early-type stars are associated to an energetic ionizing radiation field and an energy release of the order of a million times the energy release of the Sun (Massey, 2013), which might have numerous impacts onto the surrounding regions. Two other properties of massive stars are of major interest for this work. The first one is the fact that, due to their important luminosity, massive stars are able to drive significant stellar winds thanks to the radiation pressure. Those stellar winds are highly supersonic plasma of particles being associated to velocities of $\sim$ 1000-3000 km.s^{-1}. Because of those winds, early-type stars are submitted to mass losses ranging between 10^{-7} and 10^{-4} $M_\odot$.yr^{-1} (Stevens, 2005). Given the important velocities of those winds, shock physics will be an important tool to be able to describe the interaction between the winds and their surrounding.

The other important property of massive stars is their multiplicity. Indeed, massive stars are more likely to be found in multiple systems than low-mass stars (Zinnecker, Yorke, 2007). It has been estimated that the binary fraction of OB clusters is ranging from 20% to 60% (Sana et al., 2005). It is then interesting to study multiple systems of massive stars, not only to better understand their peculiar behaviour but also because multiple systems of early-type stars are an important part of the stellar population. Combining the two properties listed above (i.e. the stellar winds and the binarity) one can therefore wonder what happens when both stars in the system display stellar winds. In this case, there will be a collision between the winds of the two components producing what is called an interaction region. The shape and position of this region within the system will depend on the strength of the two winds and on the stellar separation. Systems of multiple massive stars for which there is a collision between the winds of the different components of the system are called *Colliding-Wind Binaries*. Once again arises the importance of shock physics in order to describe the interaction and the collision between the two winds.

The stars that are considered in this work are O-type stars. Those O-type stars are forming Colliding-Wind Binary systems in which a possible particle acceleration can be detected (cfr. Sect. 1.2).

1.2 Shock physics

It is first interesting and surely very important to well define what is exactly a shock when talking about shock physics. An hydrodynamic shock is a discontinuity in the properties of a flow such as its pressure, density, temperature, etc. A requirement to be considered as a shock is that the discontinuity in the flow must be significant enough in order to not be considered as a perturbation. The fact that there is a discontinuity in the flow allows to separate the flow in two different regions, the *upstream* (pre-shock) region and the *downstream* (post-shock) region. An important feature of shocks is the fact that shocks are able to convert some of the mechanical energy of the flow into heat. Shocks can therefore be referred as to dissipative systems (De Becker, Van Grootel, 2021).

When considering the stellar winds of massive stars, shocks arise because of the enormous velocity of the winds compared to the value of the sound speed inside the interstellar medium which is ranging between 1 and 10 km.s^{-1} (De Becker, Van Grootel, 2021). Indeed, having a supersonic flow is a requirement for shock formation. One can consider a polytropic gas in the framework of an adiabatic process. The adiabaticity of those shocks is considering that there is no exchange of energy between the flow and the surrounding medium. The trends followed by the properties of the flow along it are a decrease of the velocity while the density, the pressure and the temperature will increase. One has to remember that those trends are not continuous but are a sudden discontinuity at the position of the shock front. An important parameter to define in the framework of shock physics is the so-called *upstream Mach number* which is defined by

$$M_u = \frac{u_u}{a_u} = \left(\frac{\rho_u u_u^2}{\gamma P_u} \right)^{1/2} \tag{1}$$

where the index u is referring to properties of the upstream flow. This Mach number is the ratio between the velocity of the upstream flow, u_u, and the sound speed in the upstream flow, a_u. Considering an ideal gas, it can be written in terms of the upstream density, ρ_u, the upstream velocity, the upstream pressure, P_u, and the adiabatic index, γ. This adiabatic index can be defined using the number of degrees of freedom, f, using

$$\gamma = 1 + \frac{2}{f} \tag{2}$$

where in the case of a mono-atomic gas, f=3 and therefore γ=5/3. The requirement mentioned above, stressing that a supersonic flow is needed for shock formation, can then be written as

$$M_u > 1 \tag{3}$$

Another parameter of interest in the frame of shock physics is the so-called *compression ratio*, χ, defined by

$$\chi = \frac{\rho_d}{\rho_u} = \frac{u_u}{u_d} \tag{4}$$

where the index d is referring to properties of the downstream part of the flow. In the case of a normal (i.e. velocity field perpendicular to the shock front), adiabatic, high Mach-number shock for a mono-atomic gas, the compression ratio takes an interesting value given by

$$\chi = \frac{\gamma + 1}{\gamma - 1} = 4 \tag{5}$$

which tells that the gas is compressed to a value of 4 times the value of the upstream density, independently of the properties of the flow (De Becker, Van Grootel, 2021).

An important mechanism to define in order to discuss the radio emission that will be analyzed in this work is the *Diffusive Shock Acceleration - DSA*. This mechanism is converting some mechanical energy of the flow into kinetic energy of particles, therefore accelerating those particles. DSA is in fact able to accelerate particles up to relativistic speeds, which is a major requirement to produce synchrotron emission that will be discussed later on. Some assumptions are needed in order to only derive the basics of the DSA mechanism. Firstly, one can consider non-relativistic shocks for which the velocity of the flow is way smaller than the speed of light, allowing to neglect some relativistic effects. Then, high Mach number adiabatic shocks can be considered. Those two assumptions allow that, as strong shocks are considered, a significant amount of mechanical energy is available and that, as no energy is transferred to the surrounding medium, this significant amount of energy is available for the conversion into kinetic energy of particles. Finally, collision-less shocks are considered in this case in order to avoid important energy losses during collisions that may prevent a significant acceleration of particles. Once all those assumptions have been considered, the basic principle of DSA can be obtained as follows. If among the upstream flow there is a population of particles that have a sufficient velocity, those particles may be able to cross the shock front. However, because of some magnetic turbulences that are also propagating along the flow, those particles may be scattered in every directions. This is why those magnetic turbulences are also referred to *scattering centers*. It is important to mention here that the interactions with the scattering centers are considered to be exclusively head-on interactions. There is thus a non-zero probability that some particles that have crossed the shock are scattered back towards the upstream region. If so, those particles will cross again the shock front, but in the opposite direction. By crossing the shock front back and forth, the particles will gain energy, thanks to the head-on interactions with the scattering centers, proportionally to the *shock velocity jump*, V. (De Becker, Van Grootel, 2021).

$$\left\langle \frac{\Delta E}{E} \right\rangle = \frac{4}{3}\frac{V}{c} \tag{6}$$

where V is nothing more than the relative velocity between the upstream and the downstream flow (Drury, 1983). The situation is represented in Figure 1.

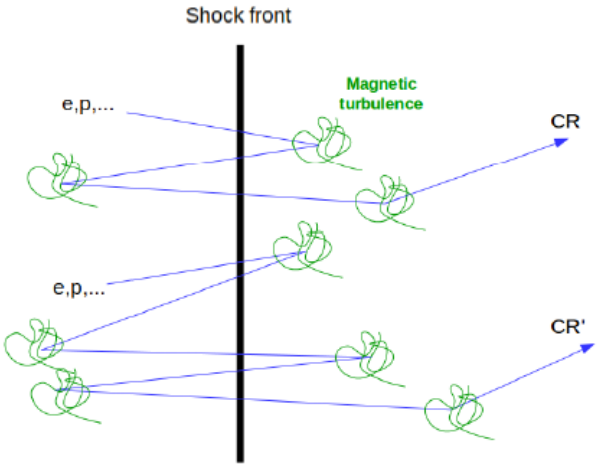


Figure 1: Representation of the DSA mechanism. e and p stand for electron and proton respectively while CR is standing for Cosmic Ray. Image taken from De Becker, Van Grootel (2021).

One very important consequence of the DSA mechanism is the resulting distribution of particles as a function of their energy. As mentioned above, when the particles are scattered by a scattering center, there is a probability that the particles are scattered back towards the upstream region and therefore there is a probability that those particles remain inside the acceleration region. However, there is also a probability that those particles are scattered away, therefore leaving the acceleration region. Only a few particles will remain for a long time inside this region, gaining energy at every cycle (upstream-downstream-upstream) while a lot of particles are likely to leave this region. The important consequence of both the fact that DSA is an iterative process and of this escape probability is that the distribution of particles will be a power law. This can be described as

$$N(E)dE = KE^{-p}dE \tag{7}$$

where p is called the *power law index* of the distribution of relativistic particles. As a matter of fact, the above distribution of particles is not dependent on the temperature. Therefore, radiative processes associated to such distribution of particles are called *non-thermal processes*. Thanks to the negative index of this power law, one retrieves that only a few particles will remain inside the acceleration region for a long time, therefore gaining a lot of energy.

As once again mentioned in De Becker, Van Grootel (2021), it is interesting to note that this power law index is related to the compression ratio defined above by the following relation

$$p = \frac{\chi + 2}{\chi - 1} \tag{8}$$

and by considering a normal, adiabatic, high Mach number shock for a mono-atomic gas, the obtained value for the power law index is

$$p = \frac{4 + 2}{4 - 1} = 2 \tag{9}$$

The DSA mechanism is of major importance in the frame of this work as particle acceleration is very likely to occur inside the interaction region of Colliding-Wind Binaries. It means that, inside a massive binary system, particle acceleration is likely to occur at the position of the collision between the stellar winds of the components of the system. This is why the systems that are studied inside this work are referred to *Particle-Accelerating Colliding-Wind Binaries (PACWBs)*. In addition, the DSA mechanism will be the key to obtain relativistic populations of particles that are needed to obtain synchrotron emission.

1.3 Radio emission of PACWBs

As PACWBs are made of massive stars with stellar winds at temperatures of a few 10^4 K, it is reasonable to anticipate that they will notably be a source of *thermal Bremsstrahlung*. It is caused by the fact that, inside the plasma which is constituting the stellar wind, some electrons are accelerated in the electric field created by protons or other ions. Because those electrons are accelerated, they are emitting some radiation and this radiation is produced in the radio domain (given the temperature of the plasma). As the electrons that are emitting the radiation are not bound to any nucleus, Bremsstrahlung radiation is also called *free-free radiation*. The population of particles involved in free-free radiation is distributed as a Maxwell-Boltzmann distribution

$$f(v) = \frac{4v^2}{\sqrt{\pi}} \left(\frac{m}{2kT}\right)^{3/2} exp\left(-\frac{mv^2}{2kT}\right) \tag{10}$$

where k is the Boltzmann's constant. Radiative processes associated to a Maxwellian distribution of particles are called *thermal processes*. Unlike non-thermal processes, the distribution of particles involved in thermal processes is dependent on the temperature, as it can be seen inside the above equation.

In order to obtain the spectrum produced by Bremsstrahlung, the associated absorption mechanism must be taken into account. This mechanism is called *free-free absorption* and is the fact that free electrons inside the plasma are able to absorb some photons that have just been emitted. In order to characterize the frequency dependence of the free-free absorption, it is useful to consider the radiative transfer equation. This equation is describing the variation of the intensity of a radiation by taking into account both emission and absorption processes. It can be written as

$$\frac{dI_\nu}{ds} = \epsilon_\nu^{ff} - \kappa_\nu^{ff} I_\nu \tag{11}$$

where I_ν is the specific intensity, ϵ_ν^{ff} the free-free emission coefficient and κ_ν^{ff} the free-free absorption coefficient. By assuming thermodynamic equilibrium, emission and absorption processes have equal efficiencies at every frequency. The specific intensity is therefore constant, setting the left hand side of the above equation to zero. Because of the thermodynamic equilibrium, the specific intensity is equal to the Planck's function $B_\nu(T)$ and therefore

$$\kappa_\nu^{ff} = \frac{\epsilon^{ff}}{B_\nu(T)} \tag{12}$$

Considering low frequencies that are relevant in the radio domain, the Rayleigh limit can be used in which the expression of the Planck's function is modified and then

$$\kappa_\nu^{ff} = \frac{\epsilon^{ff} c^2}{2kT\nu^2} \tag{13}$$

where the ν^{-2} frequency dependence of the free-free absorption coefficient has been obtained (De Becker, 2021). In order to obtain a basic idea of the spectrum created by free-free emission and absorption, it is interesting to consider an ideal case. One can consider an homogeneous ionized cloud inside which some Bremsstrahlung radiation is emitted and some free-free absorption is taking place. In this particularly simple case, the obtained spectrum is the following and can be divided into two regions.

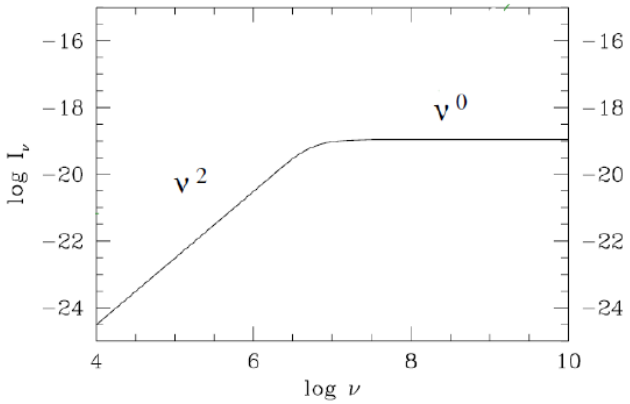


Figure 2: Spectrum produced by Bremsstrahlung for an ionized homogeneous cloud. Image taken from De Becker (2021).

One has to realise that this ideal case is unlikely to be the situation that will be encountered for this work, some deviations are then expected from this spectrum. In order to define the two regions inside the above spectrum, it is useful to introduce the *optical depth*, τ_ν. This optical depth can mathematically be defined as the natural logarithm of the ratio between incident and transmitted radiation through a plasma. In the case of free-free radiation, it is interesting to use the definition of the optical depth as a function of the free-free absorption coefficient, κ_ν^{ff}.

$$\tau_\nu = \int \kappa_\nu^{ff} \, ds \tag{14}$$

where the free-free absorption coefficient is given, in cgs units, by (De Becker, 2021)

$$\kappa_\nu^{ff} = \frac{\pi^2 Z^2 e^6}{4c} \left(\frac{1}{2\pi (m_e k)^3} \right)^{1/2} \frac{1}{T^{3/2}} \nu^{-2} n_e n_i ln\Lambda \tag{15}$$

where n_e and n_i are the electron and ion number density and where $ln\Lambda$ is the *Coulomb logarithm*. This particular factor is the logarithm of the ratio between maximum and minimum values of integration boundaries of the impact parameter.

An important point is the fact that the optical depth is proportional to ν^{-2}. Using this frequency dependence, one can divide the spectrum in two regions. Indeed, by resolving the radiative transfer equation and using the Rayleigh-Jeans limit, the specific intensity created by free-free radiation can be written as

$$I_\nu = \frac{2kT}{c^2} \nu^2 (1 - e^{-\tau_\nu}) \tag{16}$$

Because of the frequency dependence, at low frequencies, the optical depth is a significant quantity and therefore, $e^{-\tau_\nu} \sim 0$. Because the optical depth is large, the plasma is said to be *optically thick*. Hence, the specific intensity has a ν^2 dependence for an optically thick plasma, which defines the first region on the left part of the Figure 2.

However, for high frequencies and if the dependence of the intensity on the Coulomb logarithm is neglected, the optical depth is tending towards zero. The plasma is therefore said to be *optically thin*. By using a Taylor development, one can obtain

$$1 - e^{-\tau_\nu} \sim 1 - 1 + \tau_\nu \sim \tau_\nu \tag{17}$$

and the specific intensity can therefore be written as

$$I_\nu = \frac{2kT}{c^2} \nu^2 \tau_\nu \tag{18}$$

where, if one considers the ν^{-2} dependence of the optical depth, it means that there is no frequency dependence left for the specific intensity. This explains the plateau of the right part of the Figure 2 and defines the second region of this spectrum, referring to an optically thin plasma.

Actually, taking into account the dependence of the Coulomb logarithm on the frequency, the optically thin spectrum is proportional to $\nu^{-0.1}$, which explains why this slight dependence can be neglected. Therefore, thermal Bremsstrahlung is the radiation that is created inside the winds without the effect of wind collision and eventually acceleration of particles. However, considering the non-uniform density of the wind material, dealing with the radiative transfer problem leads to a frequency dependence of the optically thick wind different from that of the homogeneous cloud (see Sect. 1.4).

If one considers now a PACWBs system, another type of radio emission must be considered, the *synchrotron radiation*. Synchrotron radiation is the radiation emitted by relativistic electrons when they are accelerated in the presence of a magnetic field. Two points are important inside the previous sentence. Firstly, a population of relativistic electrons is needed. As explained above, the DSA mechanism is very likely to occur inside the interaction region of a PACWBs system, allowing to accelerate particles up to relativistic speeds. Secondly, because of the presence of a magnetic field, $\vec{B}$, the electrons are accelerated perpendicularly to their velocity. Indeed, the Lorentz force, in cgs unit, is given by

$$\vec{F} = \frac{q}{c}\vec{v} \times \vec{B} \tag{19}$$

where q, c and $\vec{v}$ are the elementary charge, the speed of light and the velocity of the electrons respectively.

As done previously for free-free emission, it is interesting to consider the associated absorption mechanism in order to obtain the spectrum of synchrotron radiation. This absorption mechanism is called *synchrotron self absorption (SSA)* and is the fact that a synchrotron photon can be absorbed by a relativistic electron of the population responsible for the emission of synchrotron radiation. The SSA coefficient can be given by (Longair, 2010)

$$\kappa_\nu^{SSA} = \frac{\sqrt{3}e^3 c}{8\pi^2 \epsilon_0 m_e} K B^{(p+2)/2} \left(\frac{3e}{2\pi m_e^3 c^4}\right)^{p/2} b(p) \nu^{-(p+4)/2} \tag{20}$$

where it is the $\nu^{-(p+4)/2}$ frequency dependence which is of interest here. Inside this expression appears K, the normalization factor of the distribution of relativistic particles (cfr. equation 7). It makes sense as the more relativistic electrons there is, the more efficient will be the absorption. Appears also the magnetic field, B, to a certain power of the index of the power law distribution of relativistic electrons (once again, cfr. equation 7). Finally, $b(p)$ is a function of this power law index, made of a series of gamma functions. It is important to note that, in this frequency dependence, appears p which is the index of the power law distribution of relativistic electrons.

By combining the emission and absorption processes, one is able to retrieve the spectrum created by synchrotron radiation that is displayed in Figure 3. Once again, the spectrum can be divided into two regions.

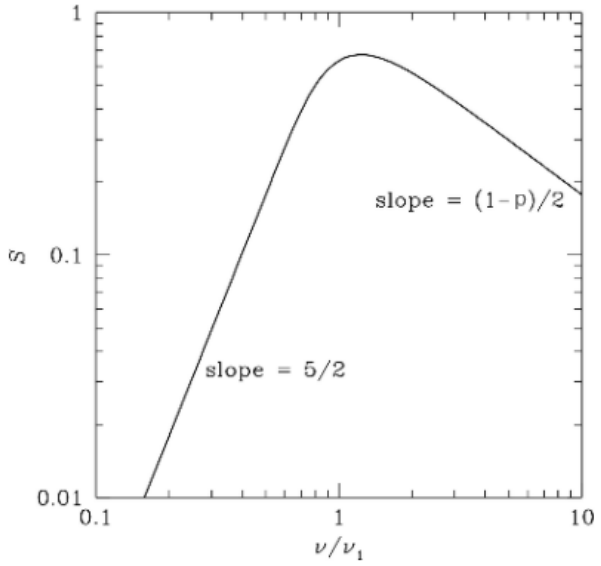


Figure 3: Spectrum created by synchrotron radiation. S is the radio flux density (i.e. the integration of the intensity over the source). The frequency units are given in terms of ν_1 which is the frequency for which τ_ν =1. Image modified from Condon, Ransom (2016).

There are several important features that are displayed in this spectrum. Firstly, the spectrum can be divided into two regions, as done for the free-free radiation. The left part of the spectrum, associated to a positive slope, is accounting for an optically thick synchrotron emission while the right part, associated to a negative slope, is accounting for an optically thin synchrotron emission. Secondly, the value of the slopes in both regions of the spectrum are well defined as displayed in Figure 3. The slope of the optically thick part is a constant value of 5/2 while the slope of the optically thin part is a function of p, the index of the power law distribution of relativistic particles. If one is considering synchrotron radiation produced by a population of relativistic particles, this population must have been accelerated up to relativistic speeds. Therefore, this population of electrons was distributed as a power law (cfr. DSA mechanism). In the frame of synchrotron radiation, if the index of the distribution of electrons as a function of their energy is given by p, the slope of the optically thin part of the spectrum is given by what is called the *spectral index* which is given by

$$\alpha = \frac{p-1}{2} \tag{21}$$

In other words, the spectrum of synchrotron radiation produced by a population of relativistic particles distributed as a power law, will be a power law, as it can be seen in Figure 3. In addition, the indices of the two laws are related by the above relation. Being able to measure the value of the spectral index will therefore allow to obtain information about the population of relativistic particles responsible for the emission of synchrotron radiation.

In addition to free-free absorption and SSA, the *Razin-Tsytovitch effect* is likely to affect the radiation. This effect appears when relativistic particles pass through a non-relativistic plasma. In the case of synchrotron radiation, relativistic effects such as relativistic beaming play an important role. The synchrotron emission is beamed in the direction of motion of the relativistic particles, enhancing the emitted power. The beamed radiation is concentrated around a small angle along the direction of motion given by (Marcote, 2016)

$$\theta_b \propto \sqrt{1 - n^2\beta^2} \tag{22}$$

where n is the refractive index of the medium and β is defined as the ratio between the velocity of the particles and the speed of light. However, the passage through a thermal plasma is associated to a refractive index that can become smaller than unity, making that $\theta_b \simeq 1$ and therefore increasing the beaming angle. Hence, the beaming effect is suppressed and the intensity of the synchrotron radiation decreases. This effect is not an absorption mechanism but rather a mechanism that directly affects the efficiency of the emission by inhibiting it.

It is important to notice that both radio emissions, thermal and non-thermal, that can be found in a PACWB system are not produced in the same region of the system. Free-free radiation is created inside the winds of both components while synchrotron radiation is emitted inside the interaction region, where there is a collision between the winds.

Another important point to notice is that, in the frame of a study of particle acceleration by massive star binaries, only the synchrotron radiation is relevant even if both emissions are present. The challenge will therefore be to be able to resolve the region of interest (i.e. the interaction region) to observe and analyze the synchrotron emission. Finally, even if this work refers to massive stellar objects, when studying PACWBs, this is not the emission coming from the stars that is of interest here but rather what is going on in between the stars.

1.4 Observational indicators for synchrotron radiation

On the one hand, it will be important to be able to resolve the region of interest and to distinguish between synchrotron and free-free radiation. On the other hand, it will be as important to know what are the hints that allow to detect synchrotron radiation and to point that particle acceleration might have happened. De Becker, Raucq (2013) is listing some indicators that allow to detect non-thermal radiation (i.e. synchrotron radiation).

The first one is the *spectral index*. Indeed, as it can be seen on Figure 2 and 3, the slopes of the spectra are different for thermal and non-thermal radiation and given the value of the spectral index, one may be able to detect the presence or not of synchrotron radiation. What is important to note is that in reality, the radiation is submitted to absorption inside the binary system that is likely to affect the shape of the obtained spectrum. Considering free-free radiation being submitted to free-free absorption, the thermal Bremsstrahlung can be associated to a power law of the kind $S_\nu \propto \nu^\alpha$ where S_ν is the radio flux density and α is the spectral index. For a spherically symmetric and homogeneous single star wind, this spectral index is close to 0.6 (Panagia, Felli, 1975; Wright, Barlow, 1975). Indeed, the situation is not anymore the one displayed in Figure 2. Firstly, the radiation is emitted inside the stellar winds, that are not homogeneous and uniform at all. Secondly, given the important densities of those winds, the optical depth is significant and the situation can therefore be associated to an optically thick plasma. By looking at Figure 2, one can see that there is a deviation, as expected, from the ν^2 dependence. However, for non-thermal radiation, the spectral index is significantly lower than 0.6 and can even be negative due to the negative slope of the optically thin part of Figure 3. Using measurements of the spectral index could be a possibility to detect particle acceleration in massive stars binaries. Pittard, Dougherty (2006) mentioned that it is possible to find a thermal contribution to the radio emission inside the interaction region. Indeed, inside this region, the winds are shocked. This small contribution is likely to cause the spectral index to be smaller than 0.6 (i.e. the expected value for thermal radiation) even if there is no synchrotron radiation taking place. As a consequence, values of the spectral index in between 0.3 and 0.6 may not be representative of the presence of synchrotron emission. Therefore, only a value of the spectral index below 0.3 could be used as an observational indicator for non-thermal radiation (De Becker, Raucq, 2013).

The second observational indicator is the *radio variability*. Considering the thermal contribution of the radio emission, no significant variation is expected as the thermal contribution is dominated by the winds of the system that should be constant over timescales of interest. However, considering the non-thermal contribution, a significant variation is expected. Indeed, the non-thermal contribution is produced in the interaction region that is periodically changing with the orbital period of the system. As a matter of fact, a varying stellar separation (in an eccentric orbit) will lead to varying physical conditions at the position of the wind-wind collision. In addition, the interaction region is embedded inside the winds of the components, leading to some free-free absorption along the line of sight. But, as a function of the orbital period, the thickness of absorbing plasma crossed by the line of sight is expected to vary, making the importance of the free-free absorption to vary also. Because the free-free absorption is dependent on the frequency, it will alter the measured flux densities differently at various frequencies, and therefore alter the measured spectral index. In addition, the variation of the absorption as a function of the orbital phase will cause the measured spectral index to vary as well. (De Becker et al., 2017).

The observation of a varying radio emission during an orbital period may therefore be a hint that some synchrotron radiation is emitted. However, using radio variability cannot be considered as a sufficient observational indicator as there are some exceptions. A famous example, cited inside De Becker, Raucq (2013), is ηCar which is a known particle accelerator for which the variability of the radio emission is explained by varying properties of the stellar wind of a component of the system rather than because of a varying non-thermal radiation.

The third observational index that can be used is the *brightness temperature*. For low frequencies relevant in the radio domain, the Rayleigh-Jeans limit is used in which the Planck's function can be written as

$$B_\nu^{RJ} = \frac{2\nu^2}{c^2} kT \tag{23}$$

In this limit case, the brightness temperature is defined as

$$T_B = \frac{c^2}{2k\nu^2} B_\nu^{RJ} \tag{24}$$

and interpreted as the temperature of a black body that would emit the same radiation as the radio source of interest. It is interesting to use this physical parameter because the brightness temperature for non-thermal and thermal radio emitters is believed to be significantly different. Purely thermal emitters can be characterized by a brightness temperature that can be as high as $T_B \sim 10^4 K$ while it is up to $T_B \sim 10^6 - 10^7 K$ for non-thermal radio emitters (De Becker, Raucq, 2013). However, because of the expected radio variability mentioned above, the flux density may reach low levels. As a consequence, the brightness temperature may also reach low levels for certain points of the orbit and using only T_B to detect non-thermal radiation is not sufficient for targets where the orbit is poorly sampled. Indeed, one has to be sure that the value of the obtained brightness temperature is not associated to a particular point of the orbit for which the free-free absorption is such that the flux density is temporarily very low.

The fourth possible observational evidence may be obtained through *high angular resolution imaging*. Radio interferometry is a technique that combines several telescopes in order to improve the angular resolution as compared to single dish observations. With Very Long Baseline Interferometry (VLBI), it is possible to resolve the winds of the two components as well as the interaction region, provided that the system is not too far away. This allows to obtain more precise information about the possible non-thermal radiation that is produced inside the wind-wind collision region. This peculiar technique of observation is further detailed inside Sec. 2.3.2.

The best way to investigate the presence or not of non-thermal radiation in massive binary systems and therefore the presence or not of particle acceleration is to combine several observational indicators among the four ones mentioned above. Depending on the observations (sampling of the orbit, resolution of the observation,...) some indicators might be favoured. By combining those observational indicators over the last years, a catalogue of known PACWBs has been established containing confirmed particle accelerators or suspected ones. The online version of this catalogue is available via the following url link : Catalogue of Particle-Accelerating Colliding-Wind Binaries.

1.5 Objectives

The objective of this work is to detect and characterize, thanks to high angular resolution radio maps, the non-thermal radio emission in two Colliding-Wind Binary systems. The data used in this work have been obtained in the framework of an European VLBI Network (EVN) pilot project containing 5 different systems. The objectives of this pilot project were to, firstly, demonstrate the capability of the EVN to map such non-thermal emission in Colliding-Wind Binary systems and to, secondly, use the EVN to extend the number of known PACWBs for which the synchrotron radio emission is mapped at high angular resolution. Given the pilot nature of the project, the detection of non-thermal emission in itself in this work would already be significant.

Among the 5 systems that have been observed during this pilot project, two of them are considered in this work. The first one is HD 168112, for which no such high angular resolution radio image has been obtained for the moment. The objective concerning this target is therefore to obtain the first image of the non-thermal emission using VLBI data. Concerning the second system, HD 167971, previous images have been obtained. The objective concerning this target is therefore to complete other previous observations of the system and to compare the various results obtained concerning this system. The former and new images are intended to be discussed in the framework of the evolution of the measured synchrotron radio emission as a function of the orbital phase.

2 Observational techniques for radio astronomy

Observations in the radio domain are quite different from observations inside other wavelength domains. Radio astronomy is using antennas to carry out observations which differ significantly from the telescopes used for observations in other parts of the electromagnetic spectrum. In this section, the philosophy of radio astronomy and how it deviates from astronomical observations in other wavelength domains is explained inside section 2.1. A brief discussion of radio antennas and their properties can be found in section 2.2. The technique of radio interferometry is detailed inside section 2.3. This section also contains a discussion about the *uvw space* (2.3.1), which is useful for mathematical considerations, the peculiar case of Very Long Baseline Interferometry (2.3.2) and the specific VLBI network that has been used for this work (2.3.3).

2.1 Philosophy of radio astronomy

It is important to understand the properties of radio astronomy before going into the details of radio interferometry. The main difference with observations inside other wavelength domains is that, in the frame of radio astronomy, it is possible to record both the amplitude and the phase of the incoming photons. This will allow to combine several signals, using some mathematical considerations, in order to significantly increase the sensitivity and the resolution of the processed images. The main idea behind the mathematical considerations is that the detected electric field is linked to the intensity of the source by a Fourier transform. Hence, this Fourier transform will be inverted in order to retrieve the intensity of the target. Those mathematical developments are detailed in the Annexe 1.

2.2 Radio antennas

An important feature of radio antennas is the *power pattern*. In order to define this power pattern, the *reciprocity theorem* is very useful. According to De Becker (2021), this theorem states that :

> *An antenna can be treated either as a receiving device, gathering the incoming radiation field and conducting electrical signal to the output terminals, or as a transmitting system, launching electromagnetic waves outward. These two cases are equivalent because of time reversibility : the solutions of Maxwell's equations are valid when time is reversed.*

In other words, the power pattern of a single dish radio antenna is the same in receiving and transmitting mode. This theorem is quite useful as it is sometimes easier to think about the antenna emitting a signal rather than receiving one. The power pattern of a single dish antenna is represented in Figure 4.

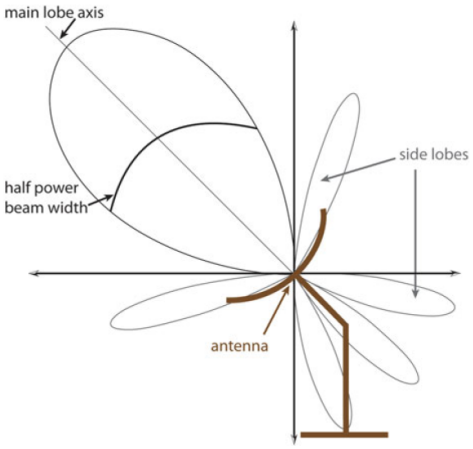


Figure 4: Representation of the power pattern of a single dish radio antenna. Image taken from Pannuti (2020).

Several features are noticeable in Figure 4. Firstly, there is a main lobe inside this pattern directed along the axis of the dish. This lobe is describing the maximum sensitivity of the antenna in the direction of the dish and referred to as the *primary beam*. The size of this beam is in fact related to the angular resolution of the antenna and is therefore antenna dependent. Secondly, there are several side lobes inside which there is also some sensitivity of the antenna. Those lobes need to be taken into account because due to their orientation, it is possible that a signal which is not coming from the observed source will be detected by the antenna, and thus contribute to the integrated signal. As an example, observations carried out at low elevation with respect to the horizon may lead to the detection of some signal coming from the Earth by the side lobes. As a conclusion, the sensitivity of radio antennas is not isotropic at all and this sensitivity pattern will have to be taken into account during calibration of the data.

Finally, the last important parameter of radio antennas to define is what is called the *system temperature,* $T_{\rm sys}$. During the observation, each dish will record a given power. For historical reasons, in radio astronomy, this power is measured as a temperature by using the Planck equation for a black body. However, this recorded temperature is accounting for several contributions. Firstly, there is the internal noise of the antenna as well as other internal noises that might arise from the receiver or other electronic devices. Secondly, there is the astronomical emission coming from the sky. It will be therefore of major importance to disentangle the astronomical signal from the total signal that the antenna receives. Then, the associated temperature will be converted to standard physical units, i.e. a flux. In this way, each antenna is characterized by an intrinsic antenna temperature, $T_{\rm a}$, describing the internal noises and a system temperature, $T_{\rm sys}$, which is the temperature that the antenna detects at a given moment during the observation of a source (Thompson et al., 2017). A good knowledge of the system temperature is therefore needed in order to recover the actual flux coming from the target itself (from the power or temperature that the antenna records).

2.3 Radio interferometry

An important parameter of any astronomical observation is the angular resolution. For an observation carried out by a single reflecting radio antenna, or a single telescope, the resolution can be estimated from

$$R \propto \frac{\lambda}{D} \tag{25}$$

where λ is the wavelength and D the diameter of the telescope. However, for radio observations, the wavelength is of the order of meters or centimeters while it can be only a few hundreds of nanometers for observations in the visible for example. Therefore, the resolution is significantly worse in the radio domain if only a single antenna is used. This is where the concept of radio interferometry arises. Radio interferometry is a technique where several antennas can be used to simultaneously observe a source and all the signals of the different antennas will be combine afterwards by what is called a *correlator*. The interest of such a technique is that, with interferometry, the angular resolution is now given by

$$R \propto \frac{\lambda}{B} \tag{26}$$

where B is what is called the *baseline* and corresponds to the maximum separation between two antennas. In radio interferometry using connected interferometers, such as the Karl G. Jansky Very Large Array (VLA), the baseline is typically in the order of 10 km therefore allowing a significantly better resolution that will allow to resolve smaller sources and to observe smaller details. What is measured by an interferometer is the interference pattern created by the addition of the different signals, similar to a 2-slit experiment. However, in a two-slit experiment, the two signals are simply added while with a correlator, the signals are digitalized. Indeed, each analog signal received at each antenna is firstly converted into a digital signal. Then, once the signals are combined, they are averaged in both time and frequency (Perley et al., 1989).

As mentioned above, the different signals coming from different antennas must be correlated by a correlator. This task is not as easy as it seems and will require some data calibration before being able to use the data and draw relevant conclusions from their analysis. If one considers interferometry performed with only two different antennas, the main goal of the correlator will be to add the two signals together but in a coherent way. Indeed, there will be a time delay between the two signals as it is represented in Figure 5. This time delay is simply due to the fact that the travel length from the source to the antenna will not be the same for the two signals as they are received at two different spatial positions on Earth. This is therefore a geometrical effect that must be taken into account when calibrating the data.

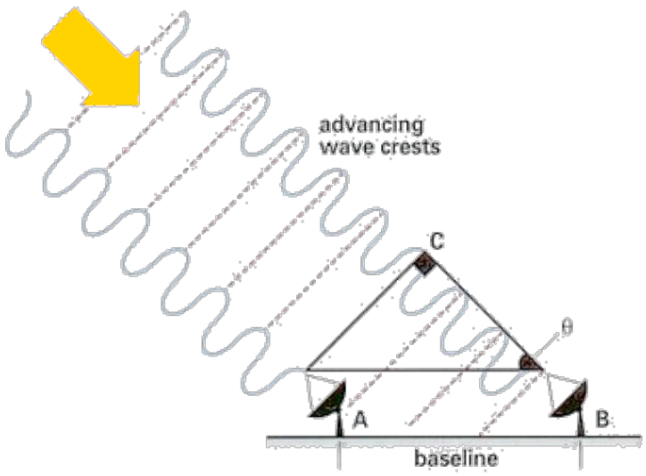


Figure 5: Representation of radio interferometry with two antennas. The time delay is visible at the reception point of the two signals. Image taken from Hessels (2013).

In the case of radio interferometry using connected interferometers, the fact that the time delay must be corrected requires a very good model of the Earth's surface in order to accurately know the position of the telescopes. In addition, due to the reasonable distance between the telescopes, they are connected by optical fibers. Hence, another time interval must be considered during which the signals are travelling from the antennas to the correlator. Because the distance between the antennas is rather small, one can assume that the same atmospheric conditions will be present above all the antennas and the same atmospheric corrections will be applied for all antennas. Indeed, because the wavelengths that are used in radio observations are close to the characteristic frequency of the plasma inside the atmosphere, the passage through atmospheric layers can introduce some different optical lengths for the signals. It is important to realise that atmospheric conditions are not affecting radio observations at the same level than e.g. optical observations. When observing at frequencies ≤ 5 GHz, the ionosphere is the only relevant influence in the signal given the characteristic frequency of the plasma. At higher frequencies (not used in this work), the only relevant influence comes from the troposphere. Nevertheless, ionospheric perturbations cannot be neglected in the case of radio interferometry using connected interferometers. However, the correction will be the same for all antennas, which will not be the case for Very Long Baseline Interferometry. Once the correlation of the signals is done, the *visibility* can be obtained.

2.3.1 Visibility and *uvw space*

As mentioned above, this is the electric field of the radiation that is detected by the antenna and all the signals are then combined in order to obtain what is called the *visibility*. This visibility is a representation of the Fourier transform of the radio brightness of the source in the sky (Atri, 2020). In Annexe 1, the visibility is also called the *spatial coherence function* and the mathematical developments are detailed in order to obtain the visibility as a function of the actual brightness of the source, I_ν, through a Fourier transform.

$$V_\nu(\vec{r}_1, \vec{r}_2) \simeq \int I_\nu(\vec{s}) e^{-2\pi i \nu \vec{s}.(\vec{r}_1 - \vec{r}_2)/c} \, d\Omega \tag{27}$$

where $\vec{r_1}$ and $\vec{r_2}$ are the position vectors of the two antennas of the interferometer, $\vec{s}$ the unit vector in the direction of the position of the source that can also be written as $\vec{s} = \vec{s_o} + \vec{\sigma}$ in which $\vec{s_o}$ is pointing towards the center of the source and is called the *phase center*. $\vec{\sigma}$ is perpendicular to the phase center. Finally, the integral is over all the solid angles Ω.

What is often done is to define a coordinate systems, (uvw), attached to the Earth and using its barycenter as center of reference to define the position of the antennas. This system is chosen such that u is in the direction of the east, v in the direction of the North Pole and w points in the direction of the phase center defined above. The magnitude of those three coordinates is given in wavelength corresponding to the observing frequency. What is important to define as well are the direction cosines of the u, v and w axes, (l, m, n), in order to define the (l, m) plane that allows to project the celestial sphere introduced in the Annexe 1. In most cases, having all the information in the projected plane, (l, m), orthogonal to the direction to the source, is enough to reconstruct the source information. The plane of the sky is sometimes also referred as the $xy-plane$. The situation is represented in Figure 6.

All those definitions allow to write

$$\begin{cases} \vec{r_1} - \vec{r_2} = (u, v, 0) \\ \vec{s} = (l, m, \sqrt{1 - l^2 - m^2}) \end{cases}$$

which allows to re-write equation 27 where the positions are defined with the new coordinates

$$V_\nu(u, v, w \equiv 0) = \int I_\nu(l, m) \frac{e^{-2\pi i(ul + vm)}}{\sqrt{1 - l^2 - m^2}} \, d\Omega \tag{28}$$

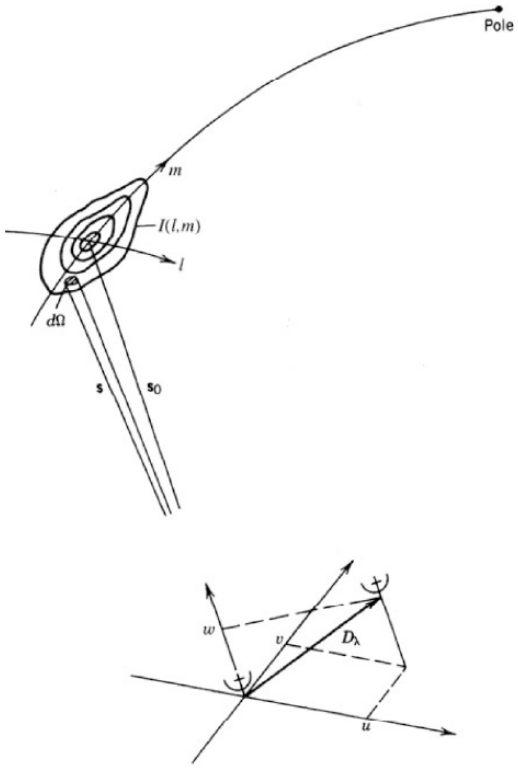


Figure 6: Representation of the (u, v, w) system coordinate as well as the (l, m) plane. $I(l, m)$ is the intensity of the source in the (l, m) plane and D_λ is the baseline between the two antennas. Image taken from Thompson et al. (2017)

where the measurements are assumed to be taken inside a plane (i.e. $w = 0$). The left-hand side of the above equation is also referred as the *complex visibility*. Indeed, this value is a complex number that can be decomposed into an amplitude and a phase thanks to the following equation

$$V_\nu(u, v) \equiv |V_\nu| e^{i\phi} \tag{29}$$

The detection of the signal is therefore consisting in the detection of both an amplitude and a phase.

Thanks to those assumptions, the Fourier transform defined by the equation 28 can be inverted and it is therefore possible to retrieve the intensity of the target. The method defined above is valid no matter what the value of the baseline is and will be therefore also used in the case of VLBI.

2.3.2 Very Long Baseline Interferometry - VLBI

The content of this section is highly based on everything I learned during my stay at the JIVE Institute (Joint Institute for VLBI in Europe) with the help of Dr. Benito MARCOTE.

Very Long Baseline Interferometry (VLBI) is a peculiar case of radio interferometry in which the baseline can reach values in the order of 10 000km. In fact, the limit is set by the diameter of the Earth (12 742km) if one considers the maximum distance between two telescopes that are on the Earth's surface, observing the same target. The fact that the maximum separation between two telescopes of the network is so large have several implications. To begin with, it requires an even better model of the Earth as for radio interferometry using connected interferometers in order to correct for the time delay between the signals. Indeed, the position of the antennas on the Earth's surface must be known with extreme accuracy. The uncertainty on the position of the antennas must be smaller than the observed wavelength length. Then, it is not possible anymore to connect all the antennas with optical fibers in order to send the data to the correlator. Traditionally, the data have been stored in some local hard disks which are then sent physically to the institute which hosts the correlator. Nowadays, several antennas just send the data electronically via internet to the same institute. It is then important to record accurately the time at which the data have been taken to be able to correlate the signals coherently, and this is ensured by local atomic clocks associated to all antennas in the network. Finally, as highlighted previously, atmospheric conditions cannot be assumed to be the same above every antenna in the case of VLBI. Therefore, a knowledge about those conditions is needed in order to derive and compensate the effect of the atmosphere on the path length for all signals.

In the case of VLBI, as the baseline can be up to $\sim$ 10 000km, the resolution that is achieved is in the order of milli-arcseconds (mas) which allows to be able to observe smaller features of a particular target. But, a given baseline is only sensitive to a certain angular scale given by equation 26 (i.e. a certain spatial frequency). Given that VLBI lacks of very short baselines, images from these arrays are not sensitive to extended radio emissions. For example, in the case of an Active Galactic Nucleus (AGN) associated to a very bright core and some radio jets, very long baselines will allow to resolve the core of the AGN while shorter baselines will allow to observe the extended emission coming from the jets. Therefore, a longer baseline is somewhat limiting the field of view that can be observed.

As defined above, the antennas are placed inside the $uv-plane$, provided that the coordinate w can be set to 0. One can then create a value of the baseline by connecting every antenna of the network to all the other antennas. This will lead to a certain number of baselines available for the observations but it is not filling all the $uv-plane$. Indeed, one can represent this plane by using the u and v distances from a point of reference as the two axes of a diagram. Because this is a complex number that is measured during radio observations (an amplitude and a phase), each value of the baseline will correspond to two points inside the $uv-plane$ (the real baseline and its complex conjugate) as represented in the Figure 7, which is commonly called the *uv-coverage* of a given observation.

In addition, the Earth is spinning during the observations. As a consequence, the points defined by the baselines inside the $uv-plane$ will rotate inside that plane and the coverage will now be represented by curves rather than a series of points. Those curves are not expected to be continuous because the observations will not observe the target source continuously (one alternates between the target and calibrator sources, as it will be explained later on). The fact that the uv-coverage is not fully filled with data points implies that not all the spatial frequencies are recovered. Indeed, as mentioned before, a given baseline allows to recover a given spatial frequency (Marcote, 2016). In other words, a better uv-coverage will lead to a more accurate imaging of the source.

In the case of VLBI, the uv-coverage is generally poorer than in the case of radio interferometry using connected interferometers as the range of possible wavelength is wider, and the number of antennas is generally more limited. This will result in a less straightforward data processing as explained later in this work. To each visibility (point in the uv plane from a given baseline) will be associated a sinusoidal fringe pattern in the $xy-plane$. As the two planes are related by a Fourier transform, a shorter baseline will lead to a larger period of the sinusoidal pattern while a short-period sinusoidal pattern will be obtained thanks to a long baseline.

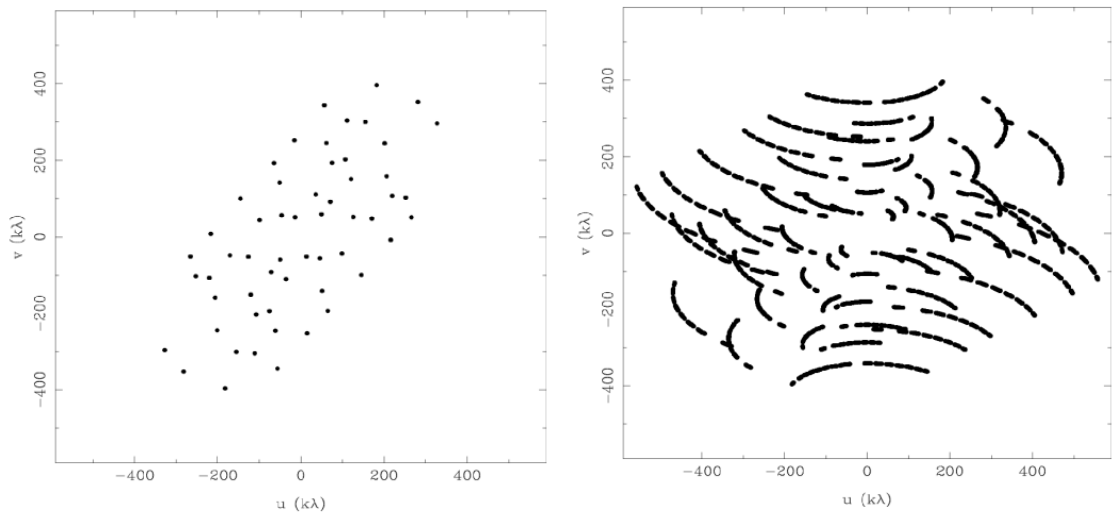


Figure 7: *Left*: Representation of the $uv-plane$ for 8 antennas without any rotation of the Earth. The u and v distances are given in kilo-wavelengths. *Right*: Same as left but where the rotation of the Earth is considered. Images taken from Isella (2011).

The fact that several baselines are available results in the superposition of several sinusoidal fringe patterns inside the $xy - plane$. Those fringe patterns will create constructive and destructive interferences and will eventually lead to a central peak created by constructive interferences between the patterns at the expected location of the source inside the $xy - plane$. Around this location, the interferences will lead to the appearance of some ripples that are showing the alternation of constructive and destructive interferences. As mentioned previously, because of the rotation of the Earth, the baselines in the $uv - plane$ are rotating. But as the $uv - plane$ and the $xy - plane$ are related by a Fourier transform, it implies that the fringe patterns are also rotating inside the $xy - plane$. Figure 8 represents the addition of several fringe patterns that have different orientations due to the rotation of the Earth.

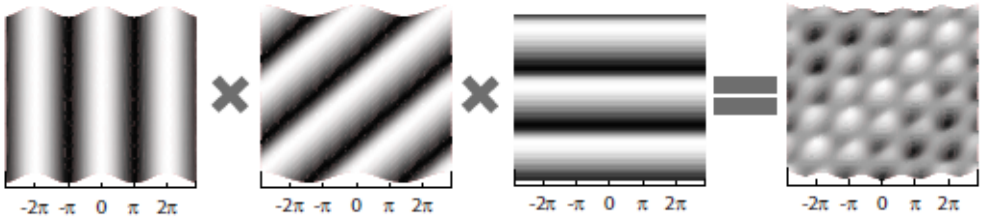


Figure 8: Representation of the addition of several fringe patterns coming from a pair of antennas from different times (i.e. including the rotation of the Earth). Image taken from Burns (2016).

In Figure 8, one can observe that at the center of the three patterns that are combined lies a white peak. Therefore, this white peak is retrieved in the final pattern because of constructive interferences. In the case of VLBI, the combination of the patterns is somewhat similar to what is represented in Figure 8 but in addition, patterns obtained with different baselines (i.e. different spatial frequencies) are combined. One has to remember that, given the equation 26, a longer baseline will result in a smaller fringe spacing in the $xy - plane$. What is done during VLBI is that several baselines are combined in order to obtain one fringe pattern that contains the contribution of all the baselines. Then, in order to produce the resulting image, several of such patterns obtained at different times are combined. This resulting image is what is called the *dirty image* which is the direct Fourier transform of the visibility data into the image plane and which is represented in Figure 9. In fact, the dirty image is a convolution between the sensitivity of the antenna (introduced in section 2.2) and the uv-coverage.

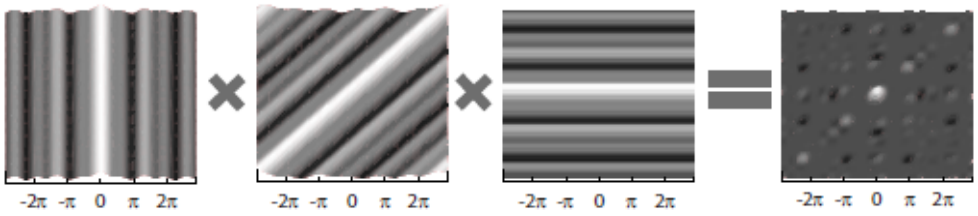


Figure 9: Representation of the dirty image obtained thanks to observation with several baselines and at different times. The three figures on the left are fringe patterns obtained thanks to the combination of signals obtained with three different baselines. Image taken from Burns (2016).

What can be observed in Figure 9 is that, using small and long baselines helps to reduce the side lobes that can be observed inside the final pattern of Figure 8. The required processes to obtain the *cleaned image*, which are not straightforward, are explained in section 3.

2.3.3 European VLBI Network - EVN

The data that have been used in this project have been taken with radio antennas that are part of the *European VLBI Network (EVN)* which is a network dedicated to radio interferometry. This network represents the most sensitive VLBI array in the world and is composed of around 22 antennas spread across Europe, Asia and South Africa [1]. The position of those antennas across the Earth is represented in Figure 10 which is taken from the *EVN and Global VLBI results and images* website [2] .

Once the data have been collected and locally recorded at each antenna, they are sent to an institute that hosts the correlator (Keimpema et al., 2015) and takes care of the data pre-processing. In the frame of this work, this institute is the JIVE Institute (Joint Institute for VLBI ERIC) at which I stayed during a week in order to learn how to calibrate and draw conclusions from radio data with the help of Dr. Benito MARCOTE and his team.

[1] https://www.evlbi.org/telescopes
[2] http://old.evlbi.org/gallery/images.html

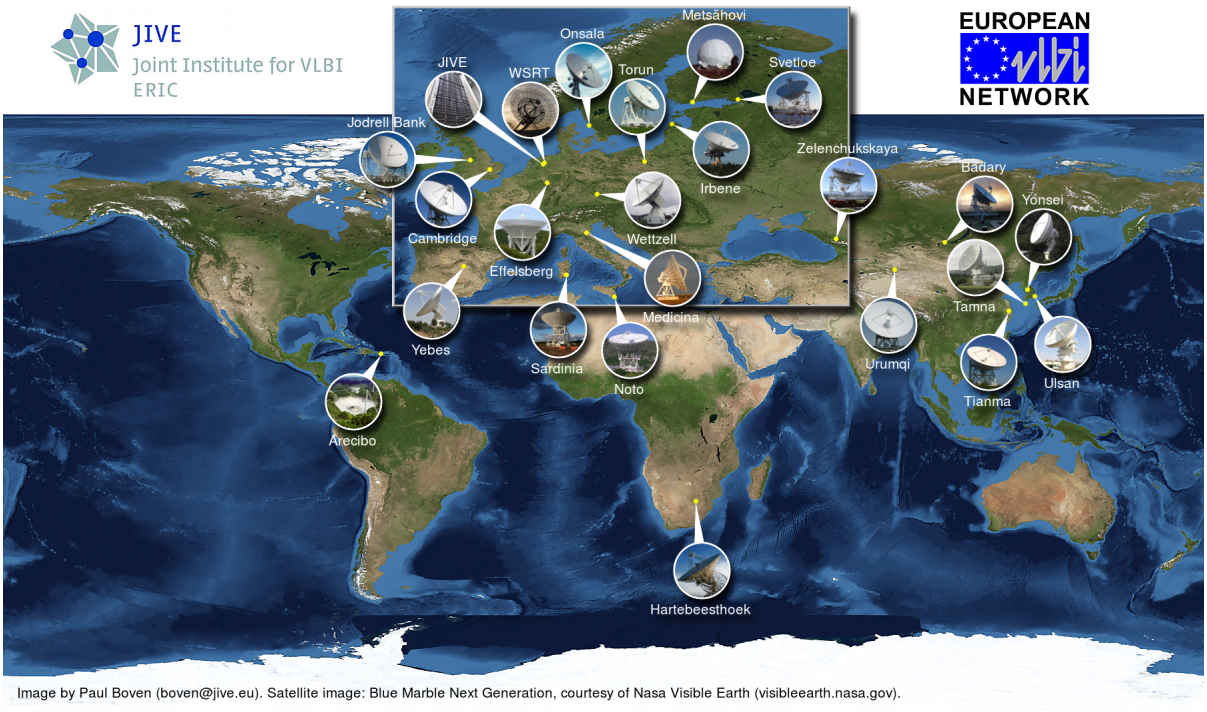


Figure 10: Location of the antennas that are part of the European VLBI Network (EVN). Credit : *JIVE*.

3 Data processing

As mentioned in the previous sections, the fact that this work is making use of radio interferometric data requires an exhaustive processing. There are two main objectives concerning this processing. The first one lies on getting the correct amplitudes for the different sources while the second one lies on the correction of the phases from the different effects that altered the signals.

The data processing for this work has been performed with the CASA software (Common Astronomy Software Applications)[3] that is tending to replace the AIPS software (Astronomical Image Processing System)[4] in radio astronomy data processing. The DIFMAP software [5] (Shepherd et al., 1994) has also been used for a more interactive visualisation of the data as well as manual flagging and cleaning of the final images. The content of this third section is the result of what I learned during my stay at the JIVE Institute thanks to long and profitable discussions that I had the chance to have with Dr. Benito MARCOTE as well as with members of his team.

3.1 A priori calibration

Before starting to calibrate the data, an *a priori calibration* of the antennas is required. Two physical quantities are of interest for this step. The first one is called the *antenna gain*. The antenna gain is a description of the sensitivity of the antenna as already mentioned in the section 2.2 (cfr. Figure 4). The a priori calibration of the antenna gain consists in accurately calibrate the conversion between the power measured at each antenna in order to translate it into astronomical flux units (i.e. Jansky units). It is important to note at this stage that the gain may evolve as a function of time due to different reasons. The main reason consists in the elevation of the sources while external contamination of the signal like radio frequency interferences (RFI) may also affect the gain. In the case of the observation of a target with a significant elevation with respect to the horizon, the antenna reaches its nominal sensitivity and it is unlikely that the signal will be greatly affected by radio signals coming from the Earth for example. However, in the case of a low elevation, the sensitivity drops and it is more likely that the signal will be perturbed by non-astronomical radio signals. The a priori calibration consists therefore in multiplying the signal coming from the target by a certain factor in the case of low sensitivity. In order to characterize the gain of each antenna, a physical quantity, called the *gain curve*, is used. The gain curve is a polynomial which is describing how the nominal sensitivity of an antenna is dropping with e.g. elevation.

[3] https://casa.nrao.edu/index.shtml
[4] http://www.aips.nrao.edu/index.shtml
[5] https://science.nrao.edu/facilities/vlba/docs/manuals/oss2013a/post-processing-software/difmap

In addition to the gain curve, the second physical quantity of interest for this step is the *system temperature,* T_{sys}. This quantity helps to characterize the internal noise of an antenna. In order to compute the value of this system temperature, an artificial signal is injected inside the antenna during every observation. Firstly, the antenna is receiving a signal coming from the sky. The amplitude that will be detected by the antenna will, at this stage, contain both the amplitude of the sky signal and the amplitude of the internal noise. Suddenly, an artificial signal, that has a known amplitude, is injected in the antenna while it is turned off to avoid observing the sky. The amplitude detected by the antenna will this time contain the known amplitude of the artificial signal and the amplitude of the noise. From this manipulation, it is therefore possible to infer the amplitude of the internal noise of each antenna. The noise amplitude can then be converted into a temperature thanks to the following relation (Petrachenko, 2013).

$$T_{sys} = \frac{\eta_A \pi D^2}{8k} SEFD \tag{30}$$

where η_A is the antenna efficiency to transmit the power that has been fed to the antenna, D the diameter of the antenna, k the Boltzmann's constant and $SEFD$ is the System Equivalent Flux Density. This physical quantity is the flux density that produces an antenna temperature equal to T_{sys}. Therefore, the above relation is assuming that the antenna temperature, T_{a}, is equal to the system temperature, T_{sys}. Knowing the $SEFD$, which is a constant for each antenna, therefore allows to convert temperature into flux density. Once the system temperature of the antenna is obtained, the corresponding noise amplitude will be subtracted from the data in order to subtract the internal noise. It is even more important to consider this T_{sys} as it is not constant with time as it can be seen in Figure 11. Therefore, the correction that will be applied to the obtained signal will not be the same for all the data and will also be time-dependent.

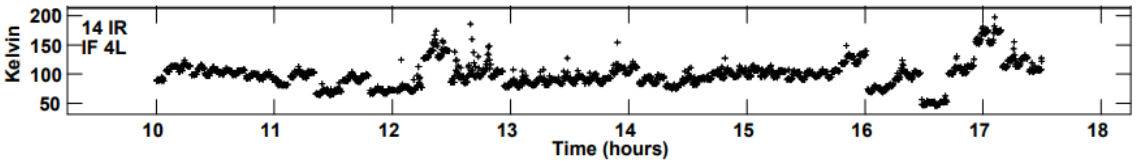


Figure 11: System temperature of the Irbene antenna (Latvia) as a function of time for the 4^{th} sub-band and the left polarisation.

3.2 Flagging

The following step before really entering into data calibration is what is called *flagging*. Among all the data obtained by the correlation of the signals of all the antennas, some of them are highly corrupted or simply useless. This can be the result of some errors coming from the antennas. For example, a wrong pointing of the source by one antenna of the network or one antenna that was not observing due to instrumental issues. An other frequent error that appears is known as *Radio Frequency Interference (RFI)*. Those RFI result from other radio signals that are not coming from an astrophysical source and that can create interferences with the signal of the target (satellites, phones, global communications,...). Keeping those data will lead to a high value of the noise inside the image and will lower the quality of the final image. In this situation, the corrupted data are manually removed from the data set. In addition to manual flagging, some systematic flagging is also performed. Firstly, the auto-correlations coming from the correlation of the signal from one antenna with itself are flagged in order to only keep the correlations between signals coming from different antennas. Secondly, the first 5 seconds of every scans are also flagged. By doing so, it ensures that no data remain in the data set during which an antenna was eventually still moving towards the source of interest.

It is important to mention that a lot of flagging has been performed for this work. It is highly due to the fact that, when considering the nature of the targets, the Signal to Noise Ratio (SNR) is pretty low. As a consequence, the data calibration is performed on a very noisy signal, leaving traces of this noise inside the final data. In addition, some minor instrumental issues appeared during the observations, forcing some entire scans to be flagged for given baselines. Nevertheless, the remaining data are sufficient to ensure meaningful final images and scientific conclusions, thanks to the very large number of initial baselines and scans available.

3.3 Fringe fitting

The remaining (or unflagged) data need to be calibrated. As mentioned previously, what is measured is both an amplitude and a phase along both time and frequency. It is important to keep in mind that the data are resulting from the correlation of two signals coming from two different antennas. Because the signals are expected to be correlated in a coherent manner, the phase difference between the two signals is expected to be zero, both as a function of time and frequency. As for the amplitude, one must remember that the $uv-plane$ and the $xy-plane$ are related by a Fourier transform. Inside the $xy-plane$, the amplitude is expected to be described by a delta function if the source can be approximated by a point-like source. Therefore, the expected amplitude in the $uv-plane$ is a constant function, which is the Fourier transform of a delta function. However, phase and amplitude often display wrong values that are deviating from the one expected and explained above. In order to correct those deviations, some gain factors must be calculated and applied to the data thanks to some various calibration steps.

The first step in radio data calibration consists in what is called a *fringe finder*. In the frame of interferometry, two signals will be correlated together. However, the signals must be combined coherently in order to ensure that the interference that is created is a constructive interference. This step is done by the correlator. What is in fact done is to try different time offsets between the signals until a sharp increase of the amplitude of the combined signals is observed. If such sharp increase is observed for a given time delay, it means that this particular value of the time delay is the time difference between the signals. In other words, this is the time offset that must be applied to a signal to combine it coherently with the second signal. This feature is what is called a *fringe*. It is important to note that this pattern is a two-dimensional pattern. The shift between the two signals that are being correlated can be made either in time or either in frequency. Therefore, it is mandatory that the two antennas are observing with the exact same frequency (i.e. the one at which the source is emitting) and that a time delay is introduced in order to obtain a fringe. It is important to mention, especially for a discussion that will arise later in the work, that the correlator only accounts for first-order corrections in time and frequency during the fringe finder step. Those corrections include the geometric model of the Earth, where all the antennas are located, and the pointing direction to the sources.

However, in the case of Colliding-Wind Binaries, the SNR is pretty small as the targets of observation are pretty faint. Therefore, the values of the amplitude and the phase will be consistent with noise inside the raw data because the low SNR does not allow to observe some clear trends. In order to obtain a fringe, a bright source will be used as it will be associated to a larger SNR. Because this bright source is observed in order to find the fringe, it is called the *fringe finder*. As the same source will be also used to calibrate the amplitude response for different frequencies (called *bandpass calibration*), the fringe finder is also referred as the *bandpass calibrator*. This is just a matter of vocabulary depending on the calibration step that is considered.

This strong source must be observed by all the antenna of the network simultaneously in order to increase the SNR for all the antennas. However, it is possible that during the observation of the fringe finder, some antennas were still moving towards the strong source or were simply turned off due to technical issues. This is why the fringe finder is observed multiple times. It ensures that at least one observation was performed by all the antennas as represented in Figure 12.

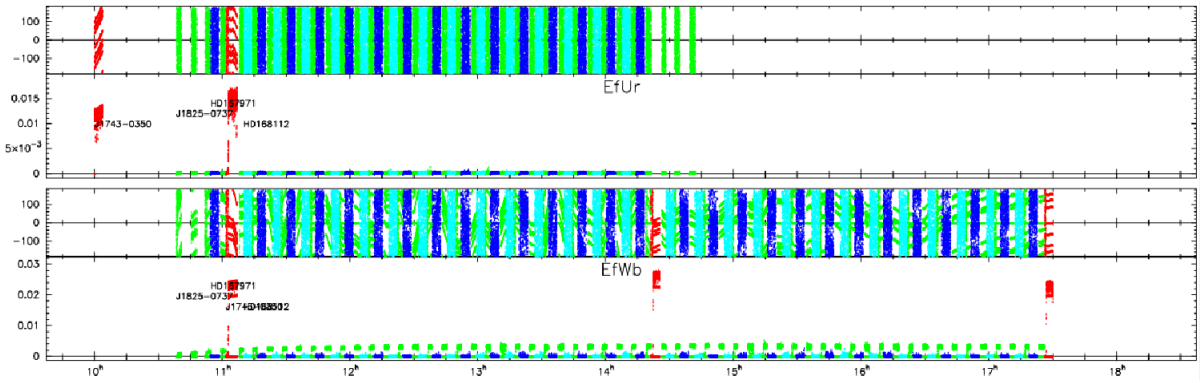


Figure 12: Phase and amplitude detection during the observation of the calibrators and the targets.

A lot of information are displayed in Figure 12. This figure is made of two boxes corresponding to data obtained thanks to two different baselines. The upper box is corresponding to data obtained with the baseline between the Effelsberg antenna in Germany (Ef) and the Urumqi antenna in China (Ur) while the other box is corresponding to data obtained with the baseline between the Ef antenna and the Westerbork antenna in the Netherlands (Wb). Each box is divided in two parts, the upper part representing the phase as a function of time and the other part the amplitude as a function of time once again. The different colors are accounting for the different sources that have been observed. In red is represented the data from the observations of the fringe finder, in green the data from observations of the phase calibrator and in light and dark blue data from observations of the two targets.

What is of interest here is the fact that the fringe finder was not observed by all the antenna during the first red scan (as represented in Figure 12). Therefore, it is very useful to use this kind of diagram in order to select the scan that will be used as the fringe finder during the calibration processes. An other important observation is that, as it can be seen in Figure 12, the fringe finder is by far the brightest observed source (higher amplitudes), which also implies much better trace of the phases (i.e. their temporal evolution is better constrained).

3.3.1 Instrumental delay

The major use of the fringe finder is in fact to "align the phases". During a radio observation, a given frequency range is chosen such that only the signals having a frequency inside this range will be detected. However, this range is divided in several sub-bands, also called *intermediate frequencies (IF)*. Because some phase offsets are still present after the correlation of the signals, some patterns will be observed in the phases and it will therefore deviate from a constant zero value consistent with a coherent correlation of the signals coming from a source at the phase center.

Indeed, some frequency-dependent phase offsets may be created by the hardware associated to the antenna. When the signal is received by an antenna, it is converted into smaller frequencies by what is called the *local oscillator* in order to transport the signal via internal electronics. However, not the full recorded bandwidth goes through the same electronics. Therefore, this conversion can introduce a frequency dependent phase offset. The instrumental delay is the fact that some jumps in phase will be observed between different IF and for the same signal as represented in Figure 13.

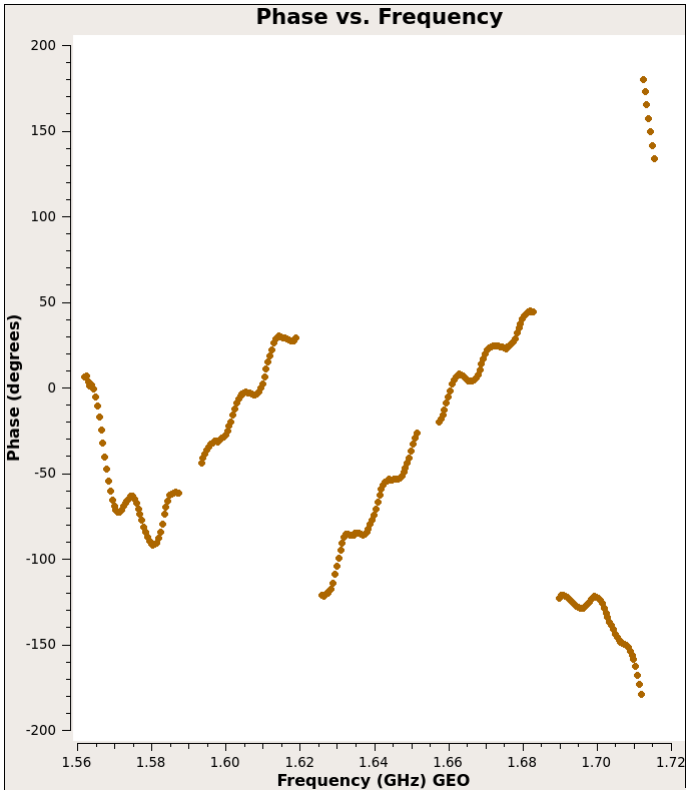


Figure 13: Phase as a function of the frequency for the baseline between the Ef antenna and the Hartebeesthoek antenna in South Africa (Hh).

Several information are displayed in the Figure 13. Firstly, one can see that the value of the phase as a function of the frequency is measured in several sub-bands. Indeed, it seems like several curves are displayed in the plot. Then, one can clearly identified that there are some jumps in between the different curves (i.e. in between the different sub-bands). This will be corrected by the instrumental delay correction. Finally, one can see that the value of the phase inside each sub-band is not consistent with zero, as expected for a coherent correlation. Those trends will be corrected by the multi-band delay correction (see Sec. 3.3.2).

Thanks to observations of the fringe finder, the corrections that need to be applied to the phases in order to align all the signals of the different IF will be computed. It is important at this stage to note that the raw data will always remained untouched. Indeed, at each calibration step, a table containing the corrections that need to be applied (called *calibration table*) will be computed. At the end of all the calibration steps, a data set containing the raw data to which the corrections have been applied will be computed.

Thanks to the first step, the instrumental delay has been corrected and the phases detected in each sub-bands will now be aligned with each other or at least, the overall signal will be nearly continuous over all the sub-bands. It is important to note that a constant value of the phase is not expected at this point. The goal is rather to obtain a continuous trend over all the sub-bands. However, inside a given IF, some trend can still be observed. The correction of this trend will be the goal of the multi-band delay correction.

3.3.2 Multi-band delay

Considering an interferometer only made of two antennas, the phase offset between the two signals is described by the differences inside the optical paths of both signals. After the correlation of the signals, some second order offsets can remain. This overall phase offset can be divided in several contributions as follows:

$$\tau = \tau_{\text{geom}} + \tau_{\text{atm}} + \tau_{\text{instr}} + \tau_{\text{source}} + \epsilon \tag{31}$$

where each term of this equation requires a small discussion. The left-hand side of the equation is representing the overall phase offset between the two signals. The first term, τ_{geom}, is the geometrical phase offset between the two signals due to the separation between the two antennas. Because of this distance, there is a time delay between the two signals and this time delay is therefore introducing a phase offset. This geometrical phase offset is illustrated in Figure 5 and is corrected during the correlation of the signals (cfr. fringe finder).

The second term, τ_{atm}, is the phase offset created by the atmospheric conditions above each antennas. In the case of VLBI, the antennas can be far from each other and the assumption under which the atmospheric conditions are the same for every antenna that is considered for radio interferometry using connected interferometers is no longer valid. This phase offset could be divided into two different contributions

$$\tau_{\text{atm}} = \tau_{\text{tropo}} + \tau_{\text{iono}} \tag{32}$$

where the two terms of the right-hand side are the phase offsets created by the troposphere and the ionosphere respectively. Among those two contributions, only τ_{iono} will be of interest in this work as, at a frequency below 5 GHz, the frequency of observation is close to the characteristic frequency of the plasma constituting the ionosphere and the interactions between the radio waves and this atmospheric layer can become significant.

The third term of equation 31, τ_{instr}, is an intrinsic phase offset to the antennas. Indeed, each antenna is associated to a specific hardware and it is possible that a phase offset arises because of the different hardwares of each antenna. This phase offset has been corrected during the correction of the instrumental delay.

The fourth term, τ_{source}, is the phase delay that is created if the source is not perfectly a point-like source at the center of the image (i.e. at the correlated position). Indeed, if inside the $xy - plane$ the source is described by a delta function (i.e. a point-like source), by taking the Fourier transform of this delta function, one obtains a constant function. However, if the source is not point-like (or not perfectly centered), it can be described by a Gaussian inside the $xy-plane$ and another Gaussian will be obtained in the $uv - plane$, leading to a non-constant value of the phase. Once again, this offset will be corrected during this calibration step. Finally, the last term of equation 31, ϵ, is a phase offset due to some possible noise in the data.

As mentioned previously, the expected phase offset between the two signals is zero, provided that they have been combined coherently. However, as the correlator only corrects for first order shifts in frequency and time, some offsets remain both as a function of time but also as a function of the frequency. Indeed, the phase offset created by atmospheric conditions, τ_{atm}, is frequency dependent as the propagation of a wave inside a plasma is frequency dependent. The multi-band delay characterizes the trends that can still be observed inside the phases that have already been corrected for the instrumental delay. In other words, the goal of the multi-band delay correction is to "flat" the phase over all the sub-bands in order to obtain a signal consistent with zero. However, in the case of Colliding-Wind Binaries, the SNR is very small and it is nearly impossible to observe such patterns. Instead, if one observes the phase as a function of the frequency or the time, only random points will be observed that could be assimilated to noise. Therefore, a peculiar calibration technique has to be used in order to observe some trends in the phase of the target.

In order to correct for this effect, the *phase referencing technique* is used. In this technique, a radio source which is near the target, called *phase calibrator*, is observed. It is important that the phase calibrator is placed not too far from the target to ensure that the effects responsible for the phase offsets will be the same for both the phase calibrator and the target. In addition, the phase calibrator must be sufficiently bright to ensure a larger SNR than for the target and must ideally be a point-like source.

The main idea of the phase referencing technique is then to perform an alternation of observations of the target and the phase calibrator. Indeed, when observing the phase calibrator, the SNR will be such that some trends will be visible. However, it will be impossible to observe some trends when observing the target because of the low SNR. Hence, observing the target in between two observations of the phase calibrator allows to extrapolate the trend inside the observation points coming from the target. This is now an appropriate time to define the two following quantities which are part of the VLBI vocabulary.

$$\text{delay} = \frac{\partial \phi}{\partial \nu} \tag{33}$$

$$\text{rate} = \frac{\partial \phi}{\partial t} \tag{34}$$

Indeed, for VLBI observations, the phases evolve very rapidly both in time and frequency. Therefore, the fit must be performed on the derivatives of the phases rather than directly on the phases. This particular feature is not encountered in radio interferometry using connected interferometers. The idea of the phase referencing technique is therefore to fit both the delay and the rate, using a least-square fit, by extrapolating the trends observed in the calibrator's data to the target's data. Indeed, the first-order shifts in time and frequency are already corrected by the correlator. Therefore, second-order effects still have to be calibrated and this is the goal of the fringe-fitting calibration step. The situation is represented in the Figure 14.

The phase referencing technique will therefore allow to create a continuous trend inside the data points despite the low SNR associated to the target. In the framework of this work, observations of the target were carried out during $\sim$ 3-4 minutes while observations of the phase calibrator were carried out during $\sim$ 1 minute. Several cycles were performed in order to obtain a situation similar to the one represented in Figure 14.

The final point of the phase referencing technique is that, thanks to the fit of both the delay and the rate, one is able to create a calibration table containing the corrections that one needs to apply to the data in order to obtain phases consistent with zero (Beasley, Conway, 1995).

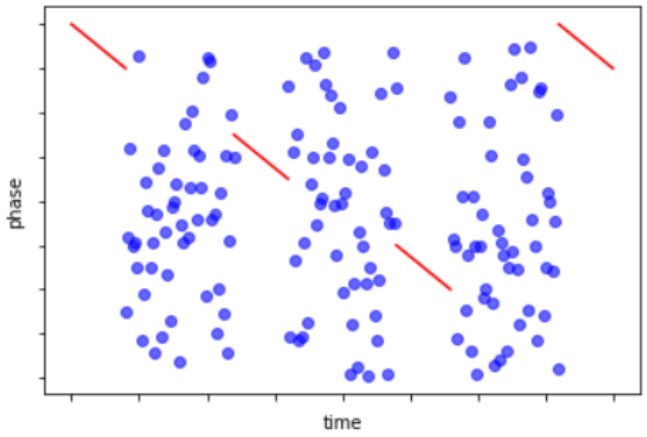


Figure 14: Representation of the phase referencing technique. The blue points represent the target's data while the red lines represent observations of the phase calibrator.

At this stage of the discussion arises a very important point concerning data calibration in VLBI. As in the case of phase referencing technique, the computation of the calibration is fully dependent on observations of the phase calibrator, any error in the measurement of the bright source's phase will propagate inside the data. If for example the position of the phase calibrator is not perfectly known or if it cannot be approximated as a point-like source, the phase that will be derived will be slightly wrong. As a result, the phase referencing technique will consist in fitting wrong values of the delay and the rate and the computed gains will also be wrong. It is therefore of major importance that the source used as a phase calibrator has a well known position and is compact enough to be approximated by a point-like source.

In conclusion, the phase referencing technique has for objective to fit the derivatives of the phase, by observing a strong source, in order to see the trends inside the low SNR target's data. Afterwards, the corrections are applied to the raw data to correct for those trends and obtain a phase offset that is consistent with zero both as a function of time and frequency.

3.3.3 Bandpass calibration

Once the phase have been calibrated, it is time for amplitude calibration. As mentioned at the beginning of the section 3.3, the expected amplitude both as a function of time and frequency is a constant value. This is because the amplitude is expected to be described by a delta function in the $xy - plane$. In order to obtain the amplitude inside the $uv - plane$, a Fourier transform has to be calculated and the Fourier transform of a delta function is a constant function. However, without data calibration, this is not what is obtained.

One can focus first on the evolution of the amplitude as a function of the frequency. When a radio observation is performed, a given frequency range ($\Delta\nu$) is chosen such that only signals having a frequency inside $\Delta\nu$ will be detected. To do so, a filter is applied in order to be sensitive to the frequency range of interest.

However, as already mentioned in the section 3.3.1, the $\Delta\nu$ used for the observation is divided into several sub-bands, the so-called intermediate frequencies (IF). This division is a consequence of the fact that the hardware of each antenna is able to process only a given frequency range, smaller than the frequency range used for the observation. Finally, each IF is divided in different channels (64 in the case of this work) and each channel is receiving signals at a given frequency which can be considered monochromatic, without loss of generality. This is because the $\Delta\nu$ of observation is divided into several IF that the obtained amplitude will not be exactly as expected. Indeed, each IF can be described as a filter which has nearly a rectangular shape (cfr. Figure 15) but not perfectly. Therefore, the edges of each IF will correspond to frequencies where the sensitivity is dropping. As a consequence, instead of observing a constant value of the amplitude over the sub-bands, several drops in the amplitude will be observed, consistent with the edges of the IF. The situation is represented in Figure 15.

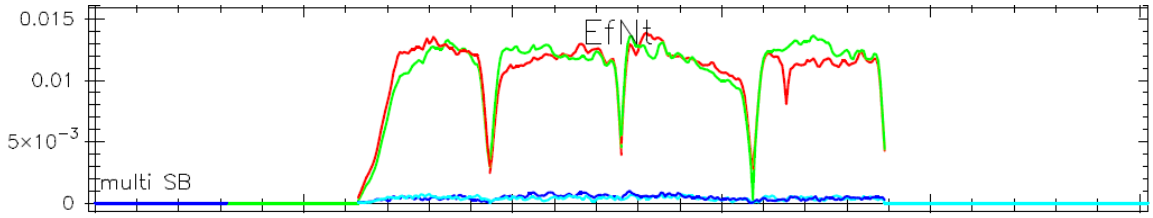


Figure 15: Representation of the evolution of the amplitude as a function of the frequency for the baseline between the Ef antenna and the Noto antenna in Italy (Nt).

A lot of information are displayed in Figure 15. This figure is showing the amplitude that is detected, prior to any calibration, as a function of the frequency for the baseline between the Ef antenna and the Nt antenna. Four color-coded signals are in fact present inside the figure. This is because two different circular polarisation states are observed, the right circular and the left circular polarisation. However, the four signals are actually correlated between each pair of antenna. As a consequence, four different combinations of those polarisation states are possible, right-right (rr), left-left-(ll), right-left (rl) or left-right (lr). The red and green signals of Figure 15 correspond to the rr and ll combinations, respectively. The clear blue and the dark blue signals account for the lr and rl combinations, respectively. One does normally expect an amplitude consistent with zero for the rl and lr combinations as the right and left polarisation are orthogonal. It would lead to incoherent signals and therefore, no amplitude would be detected for the correlated signal. However, the correlation and the detection of the signals exhibit some errors and is never perfect. It can lead to some amplitude of the rr and ll combinations to be transposed inside the rl and lr combinations.

In the practical sense, many antennas have a linear receiver which is made of two metal sticks separated by a 90° angle. The first possible explanation for a non-zero amplitude in the rl and lr combinations is that the metal sticks are not always perfectly orthogonal. The second possible explanation is that, during the conversion of the signal into circular polarisation, there can be some signal leakage from the ll and rr combinations into the rl and lr combinations. In the case of circular receivers, the discussion is roughly the same as the signal leakage during the conversion is also possible. Nevertheless, this is not the angle between the receivers that matters this time. In the case of circular receivers, two spiral-like receivers with different directions are superposed. One will always be slightly on top of the other, therefore receiving the signal a bit before the second receiver. This will introduce a phase offset that can lead to a signal leakage.

It is therefore always useful to display the values represented inside Figure 15 in order to evaluate the amount of signal that has been transposed inside the combinations of the orthogonal polarisation states. This amount is generally not significant compared to the amplitude detected inside the red and green signals but it can be important or even dominant for some baselines as in the case represented in Figure 16.

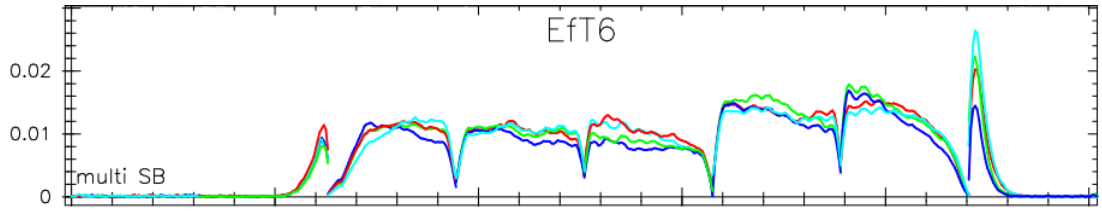


Figure 16: Representation of the evolution of the amplitude as a function of the frequency for the baseline between the Ef antenna and the Tianma antenna in China (T6).

In addition to this effect caused by the division of the frequency range into sub-bands, it is possible that the value of the amplitude which is obtained in the center of the sub-bands is not totally constant as displayed in Figure 16. Indeed, due to some contamination of the signal by non-astronomical radio signals, this level can present variations or small peaks.

Firstly, because the edges of the IF are dropping, the sensitivity must be corrected. To do so, the data points (visibilities) taken at frequencies corresponding to the edges of the sub-bands will be systematically flagged. It will ensure that all the available data are corresponding to a maximum value of the sensitivity. Secondly, the possible peaks and variation that might be created by contamination will be corrected as the amplitude will be averaged for all the IF. What will remain after those two corrections is some nearly constant values of the amplitude as a function of the frequency in which some points are missing due to systematic flagging.

Finally, it is possible that even after those two corrections, the amplitude will not be represented by a constant value over all the sub-bands. Indeed, it is possible that some jumps will be present between the amplitudes of different IF. To correct for those jumps and obtain a constant value of the amplitude, a *bandpass calibrator* is used. This calibrator has already been introduced when talking about the fringe finder. It must be a bright radio source that can be approximated by a point-like source. The bandpass calibration consists in observing this calibrator in order to derive a gain table that will later be applied to the data in order to align the amplitudes. However, in the case of VLBI, the sources are starting to be resolved. In addition, the resolution that is reached is in the order of milli-arcseconds. Therefore, the only sources that are compact and bright enough to be considered as amplitude calibrator are highly variable (Marcote, 2016). It means that the flux varies as a function of the baseline and also as a function of the position of the baseline. As a consequence, the technique involving the amplitude calibrator cannot be used in this case. Instead, the characteristics of the antennas will be used. Using the system temperature that has been computed in the a priori calibration (Sec. 3.1), one is able to calibrate the amplitude and to derive the amplitude of a signal coming from the sky.

In the case of the amplitude as a function of time, what might be observed is a decreasing amplitude as a function of time. This feature will be observed in the case of a target that is approaching the horizon during the time of the observation. Indeed, as previously mentioned, the sensitivity of the antenna is decreasing if the elevation angle with respect to the horizon is decreasing. This variation of the amplitude is therefore significant on timescales of hours and depends on the target's position. As mentioned in the section 3.1, this feature will be corrected by applying a gain to the signal depending on the elevation angle of the target at the moment of the observation.

As a conclusion, after both the bandpass calibration and the gain applied to the data, the amplitude should be represented by a constant value both as a function of the time and the frequency. Before being able to actually image the target, two last steps called *cleaning* and *self-calibration* can be useful in order to obtain even more accurately calibrated data set.

3.4 Cleaning

At each calibration step described above, a gain table will be computed. All those tables can therefore be applied to the data in order to calibrate them. After having applied those tables to the data, it is possible to image both the calibrators that have been used as well as the targets. However, the target images that are obtained at this stage are what is called the *dirty images*. Indeed, the uv-coverage is far from being perfect and the image that is created at this stage is the convolution between this uv-coverage and the constant amplitude of the source in the $xy-plane$. The uv-coverage used for this work is represented in Figure 17.

Because the coverage is not perfect, all the spatial frequencies cannot be retrieved. As a consequence, the intensity that is retrieved thanks to the inversion of the equation 28 is in fact called the *dirty intensity*. This is given by Perley et al. (1989) as

$$I^D(l,m) \equiv \iint S(u,v)V(u,v)e^{2\pi i(ul+vm)}\,du\,dv \tag{35}$$

where the function $S(u,v)$ is called the *sampling function*. This function describes the uv-coverage and characterizes the fact that the intensity is in fact only sampled due to the poor coverage often associated to VLBI. This sampling function is a function of the u and v coordinates and can be described by using Dirac delta functions. It would mean that this function deviates from zero where there is some coverage by the antenna and is equal to zero otherwise.

In fact, this dirty intensity, or also dirty image, is the convolution between the real intensity of the source and what is referred as the *dirty beam* (Garrett, 2001)

$$I_\nu^D = I_\nu * B \tag{36}$$

where

$$B(l,m) = \iint S(u,v)e^{2\pi i(ul+vm)}\,du\,dv \tag{37}$$

is the dirty beam while I_ν is the real intensity of the source. The dirty beam is the equivalent of the Point Spread Function (PSF) often used in the frame of astronomical observations at other wavelength ranges. Indeed, this dirty beam is basically an Airy function where the main peak is equivalent to the PSF but also display bright secondary peaks which will need to be removed.

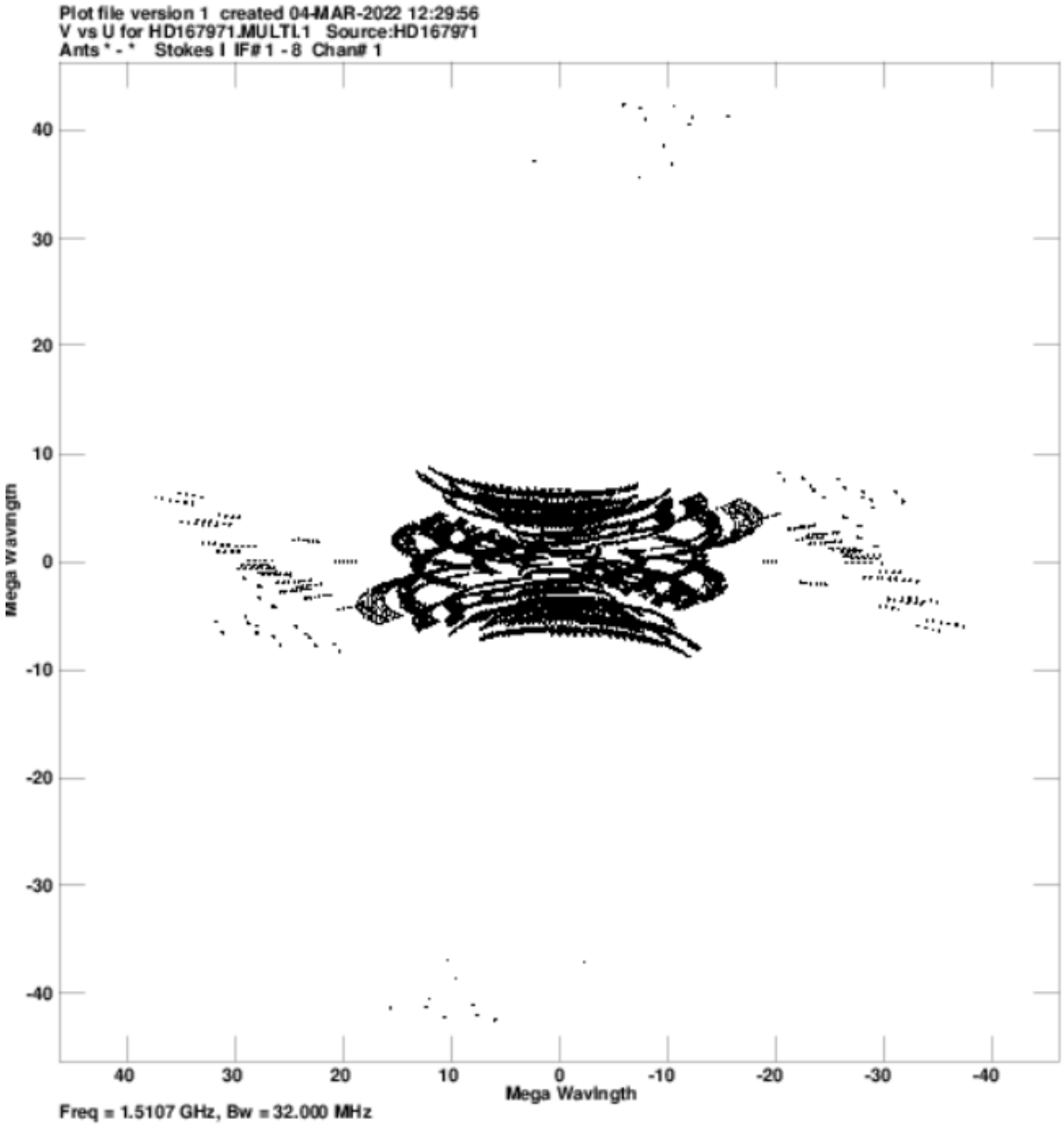


Figure 17: Representation of the uv-coverage used during observations of the targets HD 167971 and HD 168112. The u and v distances are given in Mega wavelengths.

Therefore, the image that is obtained after the above calibration steps is not perfect and some features can still be observed. Indeed, in radio interferometry, there is no unique conversion from the Fourier plane into the image plane that can provide the sky image from the uv data, given the poor uv-coverage. In order to clean the dirty beam effects from the resulting image, the *CLEAN algorithm*[6] has been used. This is an iterative algorithm which assumes that any source emission can be characterized as a set of different point-like sources. Each iteration is removing some flux from the dirty image. The first step is to define the position of the source. In the case of the cleaning of the calibrator image, it is an easy step as the calibrators are characterized by a high SNR and are clearly visible inside the image. However, in the case of the targets that are associated to a low value of the SNR, it is much more complicated and one has to be careful not to point to an other location than the one of the source (this will be discussed in more details further in this section).

To avoid noise contamination from parts of the field of view where one knows that there is no real emission, one can limit a region with the cursor. This step is also referred as *adding clean components*. Indeed, one is telling the algorithm that the source is located where the components have been placed. This region is supposed to contain the full emission of the source at the end of the cleaning. The idea of the CLEAN algorithm is then to evaluate the flux inside this region and to remove a fraction (typically between 1 and 10% (Marcote, 2016)) of this flux everywhere inside the image. The goal is to repeat this step several times until the peak that was assumed to be the source at the first step has a flux density value consistent with the value of the surrounding noise (Högbom, 1974). The cleaning is done in a interactive manner in such a way that it guarantees that the side lobes, or effects of the synthesized beam, will not introduce any artificial flux at a random position in the map because one can verify that this is the peak of emission which is considered in between each iteration. What remains after the iterations is what is called the *residual map* where the residual noise in the field of view is shown. In addition to this residual noise, there are also the components of the CLEAN algorithm that remain, which define our model of the source. In order to obtain the *cleaned image*, the components are added inside the residual map. Those components are added by convolving them with the *synthesized beam*, which is the equivalent of the point spread function.

[6] https://www.cv.nrao.edu/ abridle/deconvol/node7.html

The first step of this iterative process is therefore of major importance and the region where the source is assumed to be located must be chosen carefully. Indeed, by defining the region around the source, an assumption is made. Indeed, the flux contained in the defined region is assumed to be real (or at least a part of it). It means that, by defining this region, one is telling the algorithm that this flux at this precise location is coming from a real source and not from noise that remains inside the image. If the region is not defined at the right location, the cleaning will be performed by telling the algorithm that the flux of interest is a source while it is in fact only noise.

At the end of the cleaning, once the cleaning components have been superimposed onto the residual noise, a cleaned image will be produced and there will be no way of telling that this is not the real position and flux of the target as this flux will have been cleaned during the several iterations. In other words, it is pretty easy to obtain a biased cleaned image where the remaining flux is coming from noise and where the flux of the target has been suppressed by the algorithm.

An alternative method to place clean (point-like) components during the cleaning process in the image is to fit a two-dimensional Gaussian profile in the uv data and not in the image plane. The benefit of this alternative is that this is less affected by strong residuals in the map, allowing to recover a more accurate source size as it does not require to be convolved with the dirty beam. This is a more robust method that is often used in the case of fainter and close to point-like/Gaussian-like sources. In the framework of this thesis, the Gaussian profile fit that has been used is a χ^2 fit. At every iteration, the Gaussian profile is fitted on the location that one defined to be the location of the target. As a result, the surrounding side lobes will become less and less important at every iteration as they will be subtracted by the cleaning process.

The discussion about the importance of the placement of the location of the source is still valid. Indeed, if the Gaussian profile is placed on a bright peak that is in fact noise, the peak associated to the real source will be subtracted along the iterations and it will be impossible to say, at the end of the fit, if the obtained position for the source corresponds to the real one, or if it has just been shifted due to the assumption on the bright peak. However, the celestial coordinates (declination and right ascension) of the targets are known, this will help to locate the target in the image and to place the Gaussian profile.

An additional step that can be done during the cleaning process is to artificially put more weight on some data. As mentioned previously, the uv-coverage is not perfect at all and depending on the baseline that is considered, a given spatial frequency might be recovered. In other words, longer baselines will contribute to a better angular resolution while the shorter baselines will contribute to a better sensitivity. This concept of weighting can be made following two different strategies. The first one is called the *natural weighting*. With natural weighting, more weight will be attributed to the data where the uv-coverage is denser. In most cases with the available VLBI networks, the denser part of the coverage is accounting for short baselines. By doing this, the sensitivity is enhanced as the coverage in the region of short baselines is denser than the overall coverage of the uv-plane. This could be seen as an equivalent of a richer uv-coverage than the one really available. However, by using the natural weighting, some long baselines are, in a sense, almost neglected. As a result, the angular resolution will drop given its definition by the equation 26. During the cleaning of the targets, this is the natural weighting that will be favoured in order to increase the sensitivity and therefore decrease the level of noise around the targets.

The second weighting strategy is called the *uniform weighting*. This case is the opposite of the natural weighting. Therefore, more weight will be attributed to the longer baselines, where the uv-coverage is less dense (cfr. Figure 17). As a result, the angular resolution will increase. It is equivalent to an artificial longer baseline in the network. However, as the weight is placed where the coverage is less dense, the sensitivity will drop and the noise will become more important. This is not negligible in the case of Colliding-Wind Binaries for which the flux of the target is only of a few milli-Jansky (mJy). Indeed, as Colliding-Wind Binaries are slightly resolved when compared to the synthesized beam (or resolution of the image), their flux are spread over some given surface. Thus, the peak brightness at any of these positions drops, given that the spreading of the flux. Therefore, the sources can become less and less visible inside the map because the surface brightness is consistent with the noise, even where there is a significant flux. Therefore, it is important to note that a mix of those two strategies can be used during the cleaning of the data. By doing so, one is using a *Briggs robustness parameter* which is an intermediate situation between the natural and uniform weighting (Briggs, 1995). It is useful to firstly observe the difference between the two processes but secondly to vary the data reduction strategy given the uv-coverage that has been used for the observation. It is possible that the coverage will not be such that the weighting technique that must be used is well defined. In this case, a mix of the two is likely to be the best way to proceed. During the cleaning of the calibrator, this is the uniform weighting that will be favoured. Indeed, as this calibrator is expected to be a bright source, there is no need to decrease the level of noise around the source. Therefore, a better angular resolution will be favoured in order to better characterize the calibrator source emission.

Finally, it is important to stress once again that the cleaning process is done manually. It is possible to do it automatically where the algorithm will perform the different iterations and provide the cleaned image at the end of the process. However, as mentioned above, the targets observed in this work are expected to be quite faint and VLBI arrays are sparse arrays of antennas, implying that the noise in the maps is not Gaussian any more. It is therefore wiser to do it manually. By doing so, it is possible to confirm the position of the region in which the components will be placed at each iteration. Hence, this region could be enlarged if needed in between two steps of the algorithm. More importantly, it ensures that the region is not displaced to an other peak inside the image that would be consistent with noise. As a conclusion, doing the cleaning process manually ensures that the components are not placed on a constructive interference among the noise of the image, that could be brighter than the faint peak created by the target.

3.5 Self-calibration

The self-calibration is an additional (and the last) step that can be performed in order to increase the quality of the data calibration. What is done during the cleaning process is that, normally, components have been placed at the location of every bright sources inside the field of view. Hence, those components are building a new model of the sky. With this model, one can calibrate the full data set, including the target sources (which cannot be used for calibration purposes as they are quite faint).

The model that has been created by the cleaning process and the calibration steps is expected to yield a phase consistent with zero as well as a constant amplitude over both time and frequency. This is not always what is obtained. Indeed, the calibration processes may display some errors, this is why an additional self-calibration step makes sense. At this stage, it is important to stress that, as the self-calibration depends on the model created during the previous steps, the model must be as accurate as possible. If some artifacts are present inside the model at the end of its building, they will propagate during the self-calibration and therefore inside the final image. It is thus important to exclude the artifacts that are very likely created by calibration errors before using the model of the sky for the self-calibration (Cornwell, Fomalont, 1999).

This self-calibration is firstly performed on the phase calibrator before being performed on the targets. In the case of target self-calibration, several cycles can be performed until a stage where the model has stopped to improve. At the end of the self-calibration step, it is expected that the level of noise will be smaller and that, inside the contour plot, less features will be present around the main source of the image. In other words, this would mean that the flux density coming from the targets would have been correctly isolated inside the image. At this stage, all the calibration steps have been performed, the cleaning algorithm has been used and a self-calibration has been applied on both the targets and the phase calibrator. Therefore, one is now able to image the calibrator and the targets meaning that the heavy technical part, which consists in data processing, is finished.

4 Targets

In this section, some information found in the literature about the multiple systems that have been observed are given. In addition, a short description of the source that have been used as calibrators for the calibration processes is also presented. Because both systems which are considered in this work are located inside the same cluster (NGC6604), the same distance will be considered for both sources, consistent with the distance to the cluster (Reipurth, 2008).

4.1 HD 168112

HD 168112, which is a known non-thermal radio emitter (Bieging et al., 1989), is a system for which the binarity has been suggested by De Becker et al. (2004) due to a variability of both its radio and X-ray fluxes. In this case, the non-thermal emission could be due to a stellar wind collision. Indeed, it is explained in this paper that both variabilities can be used as hints of a binary system. Concerning the X-ray flux, the variability is compatible with a binary system since the intensity of the X-ray flux is varying as $1/D$ where D is the distance between the two stars of the hypothetical binary system (Stevens et al., 1992). Therefore, an orbit associated to an eccentricity of $e \geq 0.5$ could explain the long term X-ray variability observed in De Becker et al. (2004).

Concerning the non-thermal radio flux, the variability could be explained by a variable efficiency of the free-free absorption. Indeed, as the stellar separation is varying, the thickness of the winds inside which the non-thermal emission is located is also varying. As a consequence, when the two stars are very close to each other, the non-thermal emission is embedded inside very dense stellar winds and the free-free absorption is therefore significantly efficient. A decrease of the non-thermal radio flux would therefore be expected around the periastron passage. In other words, the maximum radio flux can be expected when the stellar separation is larger whereas the maximum X-ray flux can be expected when the stellar separation is minimum. This radio variability has been confirmed later by Blomme et al. (2005) thanks to a larger amount of measurements.

In addition to several propositions of a binarity inside the HD 168112 system, Sana et al. (2014) made use of high angular resolution interferometry in order to provide a strong hint for a companion inside this system. In this case, no orbital solution has been proposed so far and several characteristics of the system are still to determine. Indeed, Blomme et al. (2005) sampled the non-thermal emission inside the system using data from the Very Large Array (VLA) and the Australian Telescope Compact Array (ATCA). Thanks to those data, a radio lightcurve has been obtained. Despite the fact that the sampling is quite poor, Blomme et al. (2005) concluded that the short-term radio variability could indicate a period shorter than 2 years. However, only a more strict constraint on the period could allow to derive a reliable value. Some physical parameters about the system are listed inside Table 1.

Parameter	Value	Source
RA (J2000)	18:18:40.868	[1]
DEC (J2000)	-12:06:23.39	[1]
Spectral type	O5 IV + O6 IV	[2]
V_∞(A) (km.s^{-1})	2400	[3]
V_∞(B) (km.s^{-1})	3165	[3]
$\dot{M}$(A) (M$_\odot$.yr^{-1})	2.34 10^{-6}	[3]
$\dot{M}$(B) (M$_\odot$.yr^{-1})	7.94 10^{-7}	[3]

Table 1: Physical parameters of the HD 168112 system. *Note* : the index [1] mentions that the values are taken from Maíz Apellániz et al. (2016), index [2] is for the values taken from Maíz Apellániz et al. (2019) and the index [3] for the values taken from Muijres et al. (2012)

The values of the Table 1 noted $V_\infty(A)$ and $V_\infty(B)$ represent what is called the *terminal velocity* of the stellar winds of components A and B of the system, respectively. This velocity can be defined as the velocity reached by the wind at the point where the radiation is not sufficiently strong anymore to continue to accelerate the particles. Indeed, the radiation intensity coming from the stars is decreasing as a function of the distance. Therefore, the winds will be submitted to geometric dilution and one can consider that they will reach a maximum velocity, noted V_∞.

4.2 HD 167971

The second system of interest, HD 167971, is in fact a triple system (Maíz Apellániz et al., 2019) consisting of two stars on a short-period orbit and a third star on a larger orbit. Re-using the notation of De Becker (2015), the members of the short-period binary system are denoted as components Aa and Ab while the third object, which is on a wider orbit, is denoted as component B. It is interesting to note that the first observational proof of a gravitational bound between the component B and the components Aa and Ab has been obtained by interferometric measurements few years ago (De Becker et al., 2012). However, the two central stars are so close to each other that no wind-wind collision can be observed in the radio domain. Indeed, all the synchrotron emission that is emitted by the eventual collision between the winds of the two central objects is embedded inside the dense stellar winds and submitted to significant free-free absorption. Therefore, the collision of interest for this work is the wind-wind collision between the wind of the third external object and the common wind created by the two central stars. Hence, only the wider orbit, which has a period around 20 years (De Becker, 2018), is of interest for this work. Some relevant physical parameters of the system are presented in Table 2.

Parameter	Value	Source
RA (J2000)	18:18:05.895	[1]
DEC (J2000)	-12:14:33.29	[1]
Spectral type	(07.5III + 09.5III) + 09.5I	[2]
V_∞(A) (km.s^{-1})	4400 - 2600	[3]
V_∞(B) (km.s^{-1})	2921	[3]
$\dot{M}$(A) (M$_\odot$.yr^{-1})	1.58 10^{-7} - 2.51 10^{-7}	[3]
$\dot{M}$(B) (M$_\odot$.yr^{-1})	3.16 10^{-7}	[3]

Table 2: Physical parameters for the triple system HD 167971. The index [1] mentions that the values are taken from Maíz Apellániz et al. (2019), the index [2] is for the values taken from De Becker (2018) and the index [3] for the values taken from Muijres et al. (2012).

This system, which is the most luminous synchrotron radio emitter among PACWBs exclusively made of O-type stars, has already been observed in the frame of previous studies. Indeed, other equivalently luminous systems are including at least one Wolf-Rayet star, which increases the interest to observe this system. More importantly, it has already been observed with radio observations and several important features were observed (Blomme et al., 2007). In the frame of this paper, the system was observed using different radio frequencies, allowing to obtain different radio lightcurves. The first important result is the fact that a long term radio variability has been observed. At this stage it is useful to remember that radio variability has been listed as an observational indicator for synchrotron radiation in section 1.4. Indeed, because of the eccentricity of the orbit, the separation between the two central stars and the outer one is expected to vary as a function of time, leading to a variation of the physical conditions. In addition, as a function of the orbital phase, the efficiency of free-free absorption is also expected to vary since the thickness of the winds inside which the non-thermal emission is embedded is also varying if the stellar separation changes.

However, as again mentioned in Section 1.4, a variable radio flux is not always sufficient to infer the emission of synchrotron emission. Some exceptions are indeed possible. Thankfully, in the paper of Blomme et al. (2007), a negative spectral index as well as a very high flux compared to the expected thermal radio flux have been observed. Therefore, by combining those three observational indices, it is sufficient to infer the production of synchrotron emission inside the HD 167971 system and therefore the presence of particle acceleration.

The second important result of the paper is concerning the two central stars. Those two stars are believed to be on an orbit characterized by a 3.3-day period. However, no modulation of the radio flux on this timescale has been observed inside the lightcurve. It allows therefore to confirm what has been presented above concerning the fact that the free-free absorption is such that no synchrotron emission can be detected from the wind-wind collision region in the short period binary system.

An other very useful result that has been obtained is concerning the astrometry of the system. Indeed, an orbital solution has been obtained by Le Bouquin et al. (2017) using the astrometry as well as a spectroscopic constraint on several O+O binary systems and HD 167971 was one of them. The orbital parameters that have been derived for HD 167971 during this study are displayed inside Table 3. The various parameters presented inside Table 3 are respectively the period (P), the semi-major axis (a), the eccentricity (e), the inclination angle (i), the longitude of the ascending node (Ω) and the longitude of the pericenter (ω).

Parameter	Unit	Value	Uncertainty
P	days	7806	540
a	mas	18.15	0.9
e		0.443	0.02
i	deg	145.2	3.2
Ω	deg	197.7	8.8
ω	deg	108.5	4.7

Table 3: Orbital parameters of the HD 167971 system obtained by Le Bouquin et al. (2017).

This astrometry has been of major importance for previous studies of the system. Indeed, Sanchez-Bermudez et al. (2019) already observed HD 167971 inside the X- (3.5 cm ; 8.4 GHz) and C-Band (6cm ; 5 GHz) using the Very Long Baseline Array (VLBA), in 2016. This 2016 observation has been completed with archival VLBA data from 2006 taken at the same frequencies. Because two frequencies have been used for the observations, Sanchez-Bermudez et al. (2019) were able to obtain a constrain on the spectral index as the flux density is expected to follow a power law as a function of the frequency, given the non-thermal nature of the emission. Using the orbital solution previously obtained by Le Bouquin et al. (2017), a representation of the orbit inside the plane of the sky has been determined and is displayed in Figure 18.

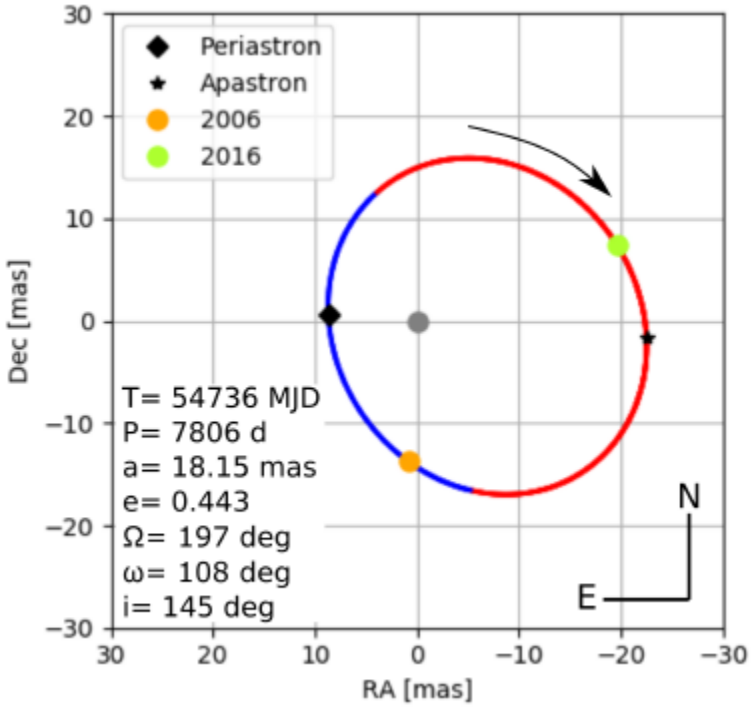


Figure 18: Orbit of the outer star with respect to the two central stars of the HD 167971 system projected in the plane of the sky. The black arrow represents the projected motion of the outer star along the orbit. Image taken from Sanchez-Bermudez et al. (2019).

4.3 Calibrators

The first source that has been used as a calibrator is J1743-0350. Also known as QSO B1741-038, it is a quasar that is often used as a calibrator in the framework of radio astronomy. Indeed, as it is a quasar, it ensures that the SNR will be sufficient and significantly larger than the SNR of the two targets presented above. In addition, by using a quasar as a calibrator source, it ensures that the calibrator will be compact enough, filling all the requirement needed to be selected as a calibrator. For the purpose of this work, this quasar has been used as a fringe finder or also called bandpass calibrator.

The second source, that has therefore been used as a phase calibrator, is J1825-0737 which is a quasar that fills all the requirements in order to be selected as a calibrator like J1743-0350. It is important to note that in this work, the two Colliding-Wind Binary systems are located inside the same cluster (NGC6604). As a consequence, only one phase calibrator was enough as the ionospheric corrections can be assumed to be the same for both targets, given their proximity.

Those two radio sources are commonly used as calibrators for radio observations and are listed inside the *VLBA Calibrators Search Form*[7] which gathers all the suitable radio sources that can be used for data reduction. Inside the following Table are gathered the positional parameters of J1743-0350 and J1825-0737. All the values of Table 4 are taken from the *SIMBAD Astronomical Database* [8]

Parameter	J1743-0350	J1825-0737
RA (J2000)	17:43:58.8561	18:25:37.60955
DEC (J2000)	-03:50:4.616	-07:37:30.0128

Table 4: Positional parameters of the calibrators that have been used for this work.

[7] http://www.vlba.nrao.edu/astro/calib/
[8] http://simbad.u-strasbg.fr

5 Results

In this section are presented the results obtained for both systems studied in this work. Firstly, the results of all the calibration steps as well as a short discussion on their effect and quality is presented in section 5.1. Then, the final images and the properties that can be obtained from it are presented in section 5.2.

5.1 Calibration data files

The first part of the data reduction that can be discussed concerns the calibration tables that have been obtained thanks to the calibrators, J1743-0350 and J1825-0737. As explained in the section 3, the first steps of radio data calibration consist in what is called *a priori calibration* and *flagging*. On the one hand, a priori calibration includes the corrections linked to the system temperature, T_{sys}, as well as the corrections linked to the variation of the antenna gain as a function of the elevation angle. On the other hand, flagging consists in (manually at this stage) deleting some data that display obvious contamination or corruption.

For each of this a priori calibration step, a calibration table is computed and will be applied to the data later on. After the a priori calibration, the data look pretty noisy and one can see that still a lot of trends are included as it can be seen in the following Figures illustrating the effect of the a priori calibration.

There are several things to discuss about the two plots of Figure 20. Firstly, the data displayed in those plots are corresponding to data obtained by observing a calibrator, J1825-0737. Indeed, only a significant SNR would allow to observe such trends inside the data. Hence, one does not expect to see such significant trends inside the targets data as they would be embedded inside noise. Secondly, the trends that are visible inside both the amplitude and the phase plot are the expected trends discussed inside section 3. In the left plot, the effect of the sub-bands can be observed as the amplitude is dropping nearly down to zero at each sub-band edge. It is therefore interesting to observe here that only 5 sub-bands among the initial 8 remain after the initial flagging step. This will be valid for all the calibration steps of this work. Then, the amplitude is clearly not constant both over the sub-bands and inside a given sub-band. It therefore makes sense to make use of some further calibration steps to obtain the expected constant amplitude. In the right plot, once again, the expected trends are visible. Because no phase offset has been corrected yet, there is a phase shift that is creating this kind of linear increase over the sub-bands.

Finally, it is interesting to note that the a priori calibration is slightly modifying the data (as it can be seen by comparing the plots of Figure 19 and 20) but the goal here is absolutely not to correct the trends that can be seen inside the data. It will be the objective of further calibration steps.

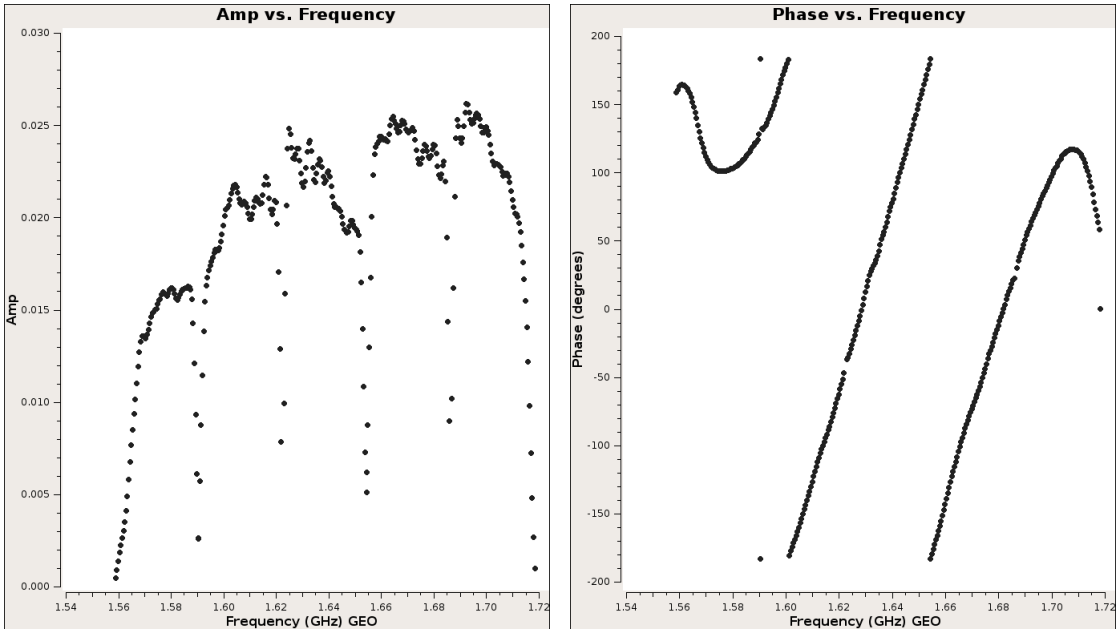


Figure 19: *Left*: Amplitude of J1825-0737 as a function of the frequency without any calibration. *Right*: Phase of J1825-0737 as a function of the frequency without any calibration. Both plots are corresponding to the baseline between the Ef antenna and the Sardinia antenna in Italy (Sr).

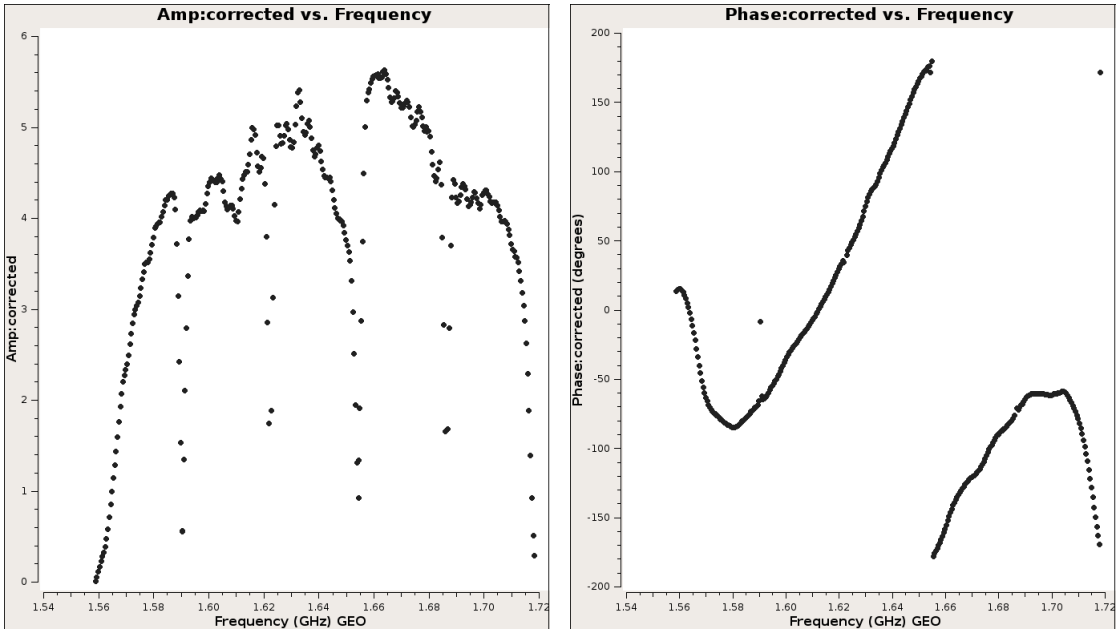


Figure 20: *Left*: Amplitude of J1825-0737 as a function of the frequency after a priori calibration. *Right*: Phase of J1825-0737 as a function of the frequency after a priori calibration. Both plots are corresponding to the Ef-Sr baseline.

After the a priori calibration, some systematic flagging is performed. As mentioned previously, the 5 first seconds of each scan will be flagged in order to delete the data that might have been taken while an antenna was still moving towards its definite pointing position. In addition, the edges of the IF (i.e. the sub-bands) are also flagged in order to avoid data associated to a reduced sensitivity as discussed in section 3.3.3. In the framework of this study, the fraction of the sub-bands that have been flagged accounts for 10%. The effect of those systematic flagging are visible in Figure 21 as the data points corresponding to the sub-bands edges have been removed. Once again, no correction of the trends has been performed at this stage.

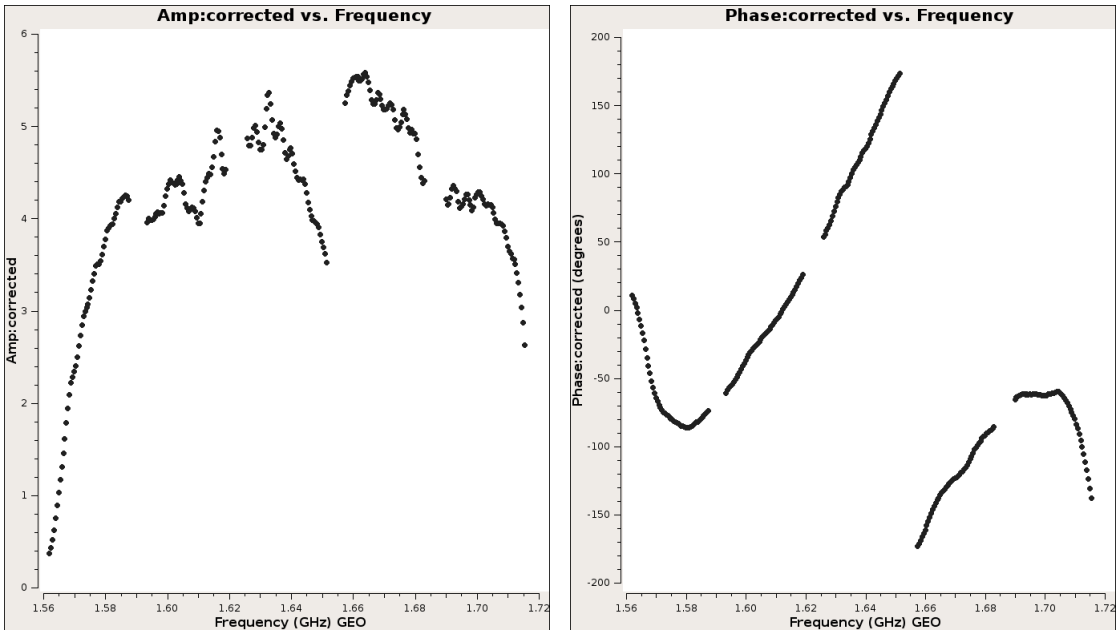


Figure 21: *Left*: Amplitude of J1825-0737 as a function of the frequency after systematic flagging. *Right*: Phase of J1825-0737 as a function of the frequency after systematic flagging. Both plots are corresponding to the Ef-Sr baseline.

After the a priori calibration comes the ionospheric correction. As mentioned in section 3.3.2, this calibration step is of major importance in VLBI measurement given both the fact that the antennas can be separated from several thousands kilometers and that measurements are taken at frequencies below 5GHz. Because of this very important baseline values, the ionospheric conditions cannot be assumed to be similar above each antenna of the network. One important tool to understand the corrections that need to be applied is what is called the *Total Electron Content (TEC)* which measures, as its name indicates, the content of electron inside the ionosphere at a given location. Because the ionosphere is a plasma made of free electrons and because the radio waves that are detected for this work are in the range the electrons are sensitive to, it is interesting to know this TEC value in order to estimate how the electrons will resonate and interact with the radio waves.

By doing so, one can estimate if the travel path of the wave has been significantly affected by the ionosphere or not and therefore if it requires a correction. An example of the representation of the TEC measurement is illustrated in the Figure 22.

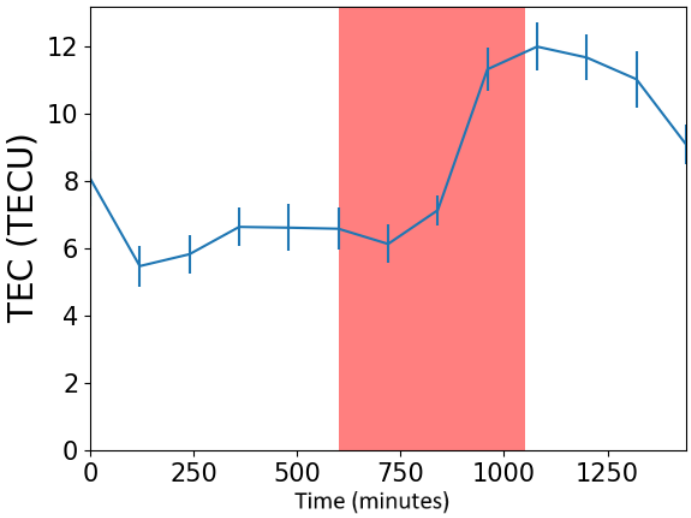


Figure 22: Total Electron Content corresponding to the scan of the phase calibrator chose to perform the calibration steps.

After the a priori calibration come the heavier calibration steps. Firstly, the instrumental delay correction is performed on the data of the calibrator in order to obtain a correction table that will be applied onto the targets' data afterwards. In terms of correction of the trends inside the data, the goal of the instrumental delay calibration step is to create a continuous phase over frequency. As explained inside section 3, the value of the phase will display a shift going from one sub-band to an other. The objective here is therefore to connect the value of the phase in each sub-band. It is important to note at this stage that the value of the phase is still not expected to be close to zero, it will be the goal of the following calibration step. The result of this calibration step on the data of the calibrator can be seen in the Figure 23. It is also interesting to note that, inside a given sub-band, the value of the phase is not perfectly constant and some trends are still visible. Those trends will be corrected thanks to following calibration steps.

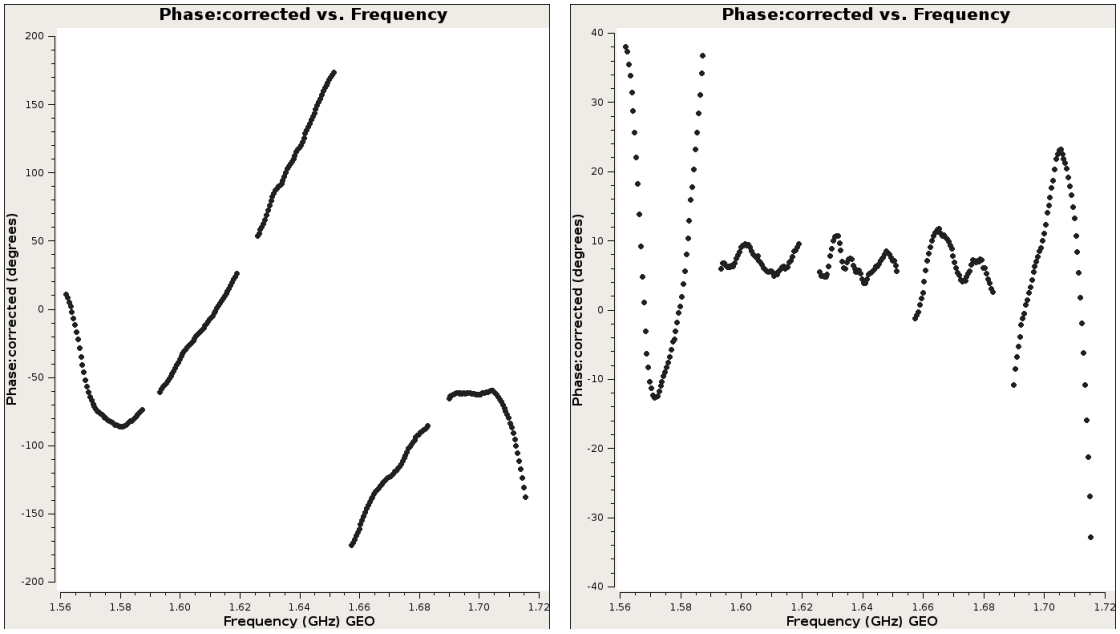


Figure 23: *Left*: Phase of J1825-0737 as a function of the frequency before instrumental delay correction. *Right*: Phase of J1825-0737 as a function of the frequency after instrumental delay correction. Both plots are corresponding to the Ef-Sr baseline.

Secondly comes the multi-band delay correction. The goal of this calibration step is to align the value of the phases as a function of time over all the sub-bands that remain after the manual flagging. Once again, a calibration table will be computed from the data of the calibrator and will be applied later on to the data of the targets. In Figure 24 is represented the data, for a given baseline, before and after that the multi-band delay correction has been performed.

Several information can be deduced from Figure 24. On the one hand, one can easily observe the trend inside the phase and one can see that this trend remains after the multi-band correction. On the other hand, when looking at the range of values that the phase is taking inside the top plot and the bottom plot of Figure 24, it is clear that the range of values is thinner. Indeed, it is a first step towards a phase consistent with zero. By reducing the thickness of this value range, the multi-band correction has somewhat started to align the phase along zero. Yet, it is still not completely around zero, as it can be seen on the bottom plot of Figure 24. This proves that the multi-band delay correction calibration step is not perfect or that the source is not a point-like source. This is why a self-calibration step is making sense. Indeed, thanks to the model that will be created by all the calibration steps, remaining trends will be corrected and one is expecting the phase of Figure 24 to be closer to zero after the self-calibration correction. In other words, the multi-band delay correction helped to define a thinner trend inside the phase that will be better removed thanks to self-calibration.

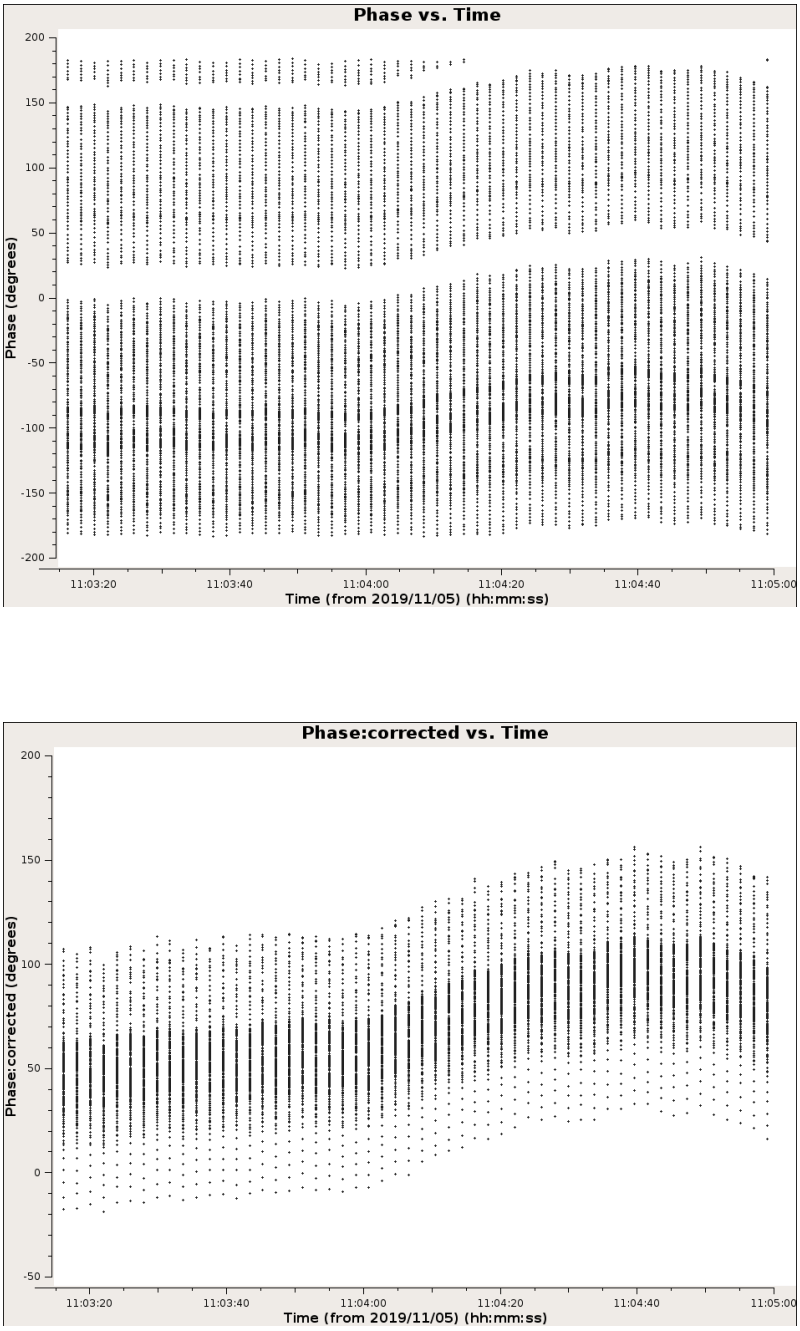


Figure 24: *Top*: Phase of J1825-0737 as a function of time before the multi-band delay correction. *Bottom*: Phase of J1825-0737 as a function of time after the multi-band delay correction. Both plots are corresponding to the Ef-Sr baseline.

It is interesting to note also that the multi-band correction is also affecting the phase as a function of the frequency. After the multi-band correction step, the phase is closer to zero and one can observe on Figure 25 that inside a given sub-band, the phase is more constant than before the multi-band delay correction.

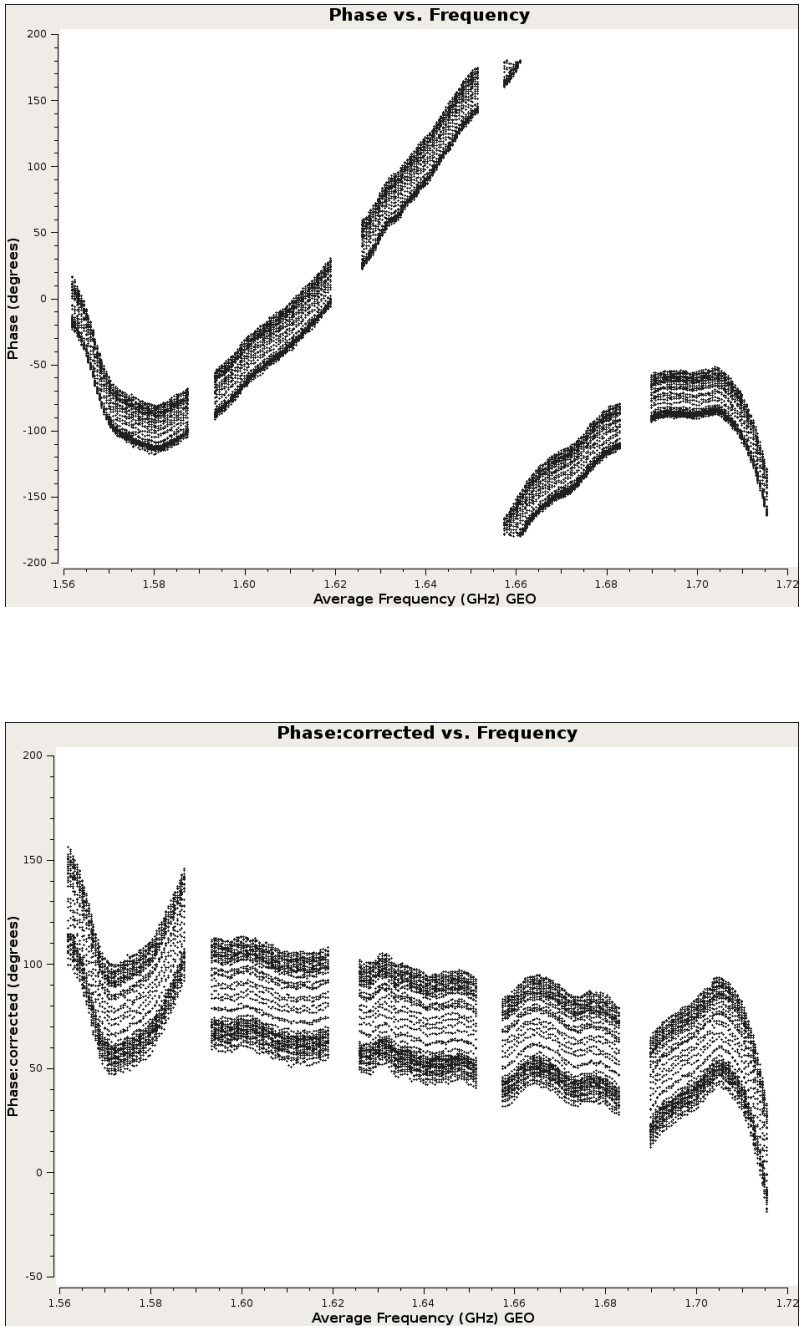


Figure 25: *Top*: Phase of J1825-0737 as a function of frequency before multi-band delay correction. *Bottom*: Phase of J1825-0737 as a function of frequency after multi-band delay correction. Both plots are corresponding to the Ef-Sr baseline.

Several comments are required considering the phase at this time. Firstly, the top plot of Figure 25 is accounting for the systematic flagging but not for the instrumental delay correction. Indeed, by looking at the Figures 23 and 25, one can easily see that the trend that has been corrected inside Figure 23 re-appears in Figure 25. Secondly, by comparing the right plot of Figure 23 and the bottom plot of Figure 25, one can see that the value of the phase inside a given sub-band is more constant, tending towards a constant value of the phase as well as a continuous signal over all the sub-bands.

Finally comes the bandpass calibration step. This step is more focusing on the correction of the amplitude rather than correction of the phase as for the previous calibration steps. Indeed, the expected amplitude is a constant value over both time and frequency. After the removal of some hardware effects (near the edges of the sub-bands as explained before), the bandpass calibration allows to obtain such value of the amplitude. This can be seen in Figure 26. This bandpass correction concludes the part of the calibration called *fringe fitting* which has been described in details inside section 3.3. Once again, several features are visible inside Figure 26. Firstly, if one is focusing onto a single sub-band, it is clear that the effect of the bandpass calibration has been to correct the signal in order to obtain a constant amplitude over the sub-band. This was clearly not the case after the systematic flagging (top plot of Figure 26). Secondly, if one now considers all the sub-bands, it is clearly visible that some jumps are still present between the different sub-bands. However, it was expected from the bandpass calibration step to align those signals in order to obtain a constant value of the amplitude. This is because the bandpass calibration step is sensitive to Radio Frequency Interference (RFI) already mentioned in this work. Those interferences created by non-astronomical radio signals are likely to create some peaks inside the amplitude of a given sub-band as it can be observed in the top plot of Figure 26. During the bandpass calibration step, the amplitude is averaged and normalized, which leads to the fact that RFIs will lead to different averaged amplitudes for the different sub-bands. A RFI dependence is also present during the correlation of the signals. Those features are expected to be corrected during self-calibration, once a confident model of the calibrator will be established.

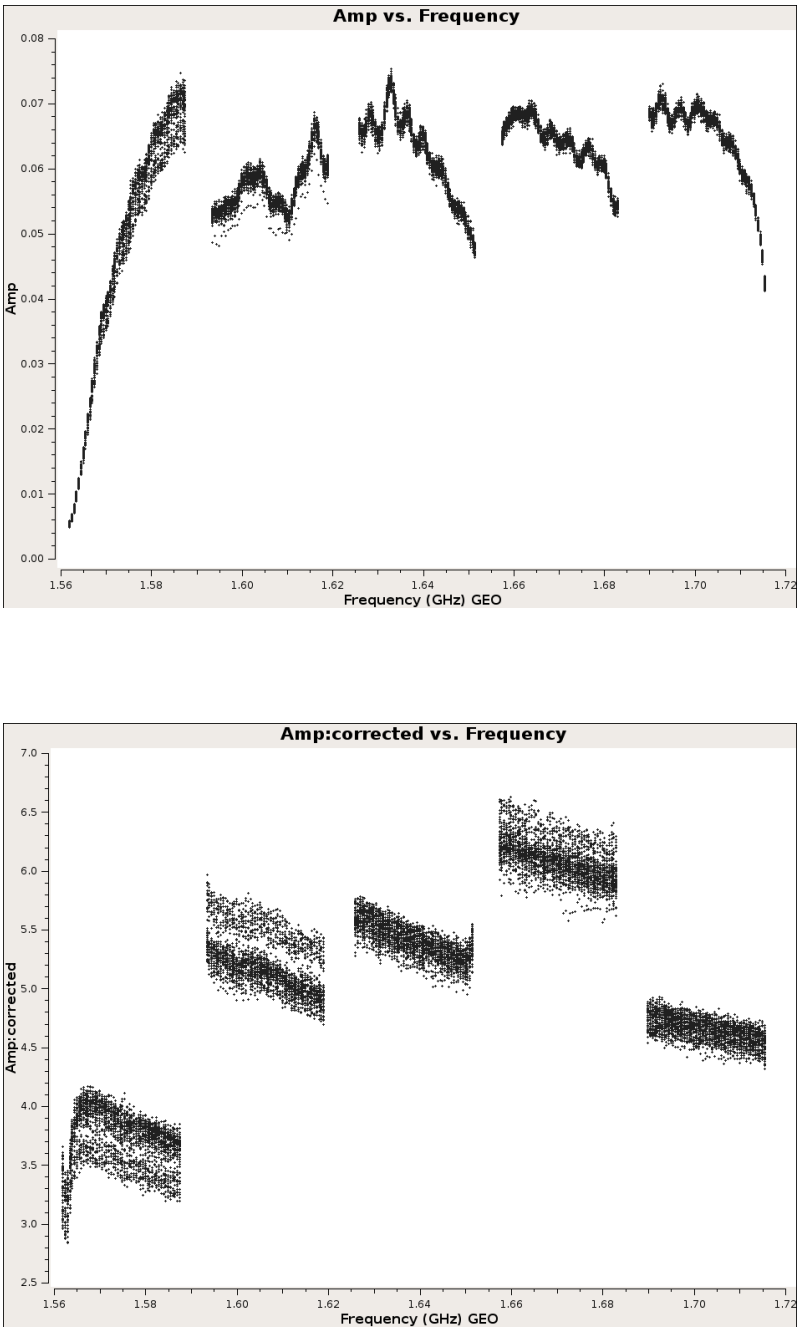


Figure 26: *Top*: Amplitude of J1825-0737 as a function of the frequency before the bandpass calibration. *Bottom*: Amplitude of J1825-0737 as a function of the frequency after bandpass calibration. Both plots are corresponding to the Ef-Sr baseline.

It is important to note that, at this point, no calibration table has been applied to the targets data. The goal here is to use the bright calibrators described in section 4.3 to obtain the calibration tables. Inside the targets data, as the signal is weaker, it is dominated by noise. However, inside that noise is hiding the real signal coming from the sky containing some valuable information. Therefore, it makes sense to apply the calibration tables to the targets data even if the information coming from the sky is not visible over the noise. In addition, thanks to appropriate choices of both phase and bandpass calibrators, the corrections that need to be applied to the targets are the same than the one applied to the calibrators. In other words, the entire calibration code (still under development at the JIVE institute [9]) is using the trends visible inside the calibrator signal. This is only at the end of the process, when all the calibration tables have been computed, that the targets data are submitted to the calibration processes.

Once all those calibration steps as well as the CLEAN algorithm have been performed, a model of the calibrator has been obtained. Indeed, each step has determined a calibration table that describes the trends inside the calibrator data. It is then interesting to apply this model to the data to perform what is called a *self-calibration*. Thanks to this step, one expects to observe a refinement of the calibration, especially for the multi-band and bandpass correction, as discussed above. In the Figure 27, the model that has been created can be seen.

It is not straightforward to interpret the Figure 27. Firstly, if one focuses on the top plot, one can observe a decrease of the amplitude as the uv distance increases, which is expected. Indeed, as the uv distance (i.e. the baseline) increases, the source starts to be resolved if one follows the equation 26. As soon as the source is resolved, the size of the synthesized beam will be smaller than the size of the actual source, meaning that not all the signal will be detected and therefore the amplitude is expected to decrease. In other words, as soon as the green and therefore the red signal is starting to decrease inside the top plot of Figure 27, it means that the source is starting to be resolved. As a consequence, for baselines longer than $\sim 14.10^6\lambda$, the source is resolved and therefore no trend is expected inside the amplitude plot or the phase plot as the flux is spread over some given surface.

If one now focuses on the bottom plot of Figure 27, the first feature that can be observed is the fact that, indeed, more trends are visible for baselines shorter than $\sim 14.10^6\lambda$ as less signal is available for longer baselines. The second observation that can be made on this plot is the fact that the model is fitting correctly the linear trends that are visible inside the plot. A zoom inside that plot is represented in Figure 28 where some trends are more visible. It can be seen that the red model is indeed well fitting the signal which ensures that the model applied to the targets is adapted and therefore that the targets will be rightfully self-calibrated.

[9] https://www.evlbi.org/evn-data-reduction-guide

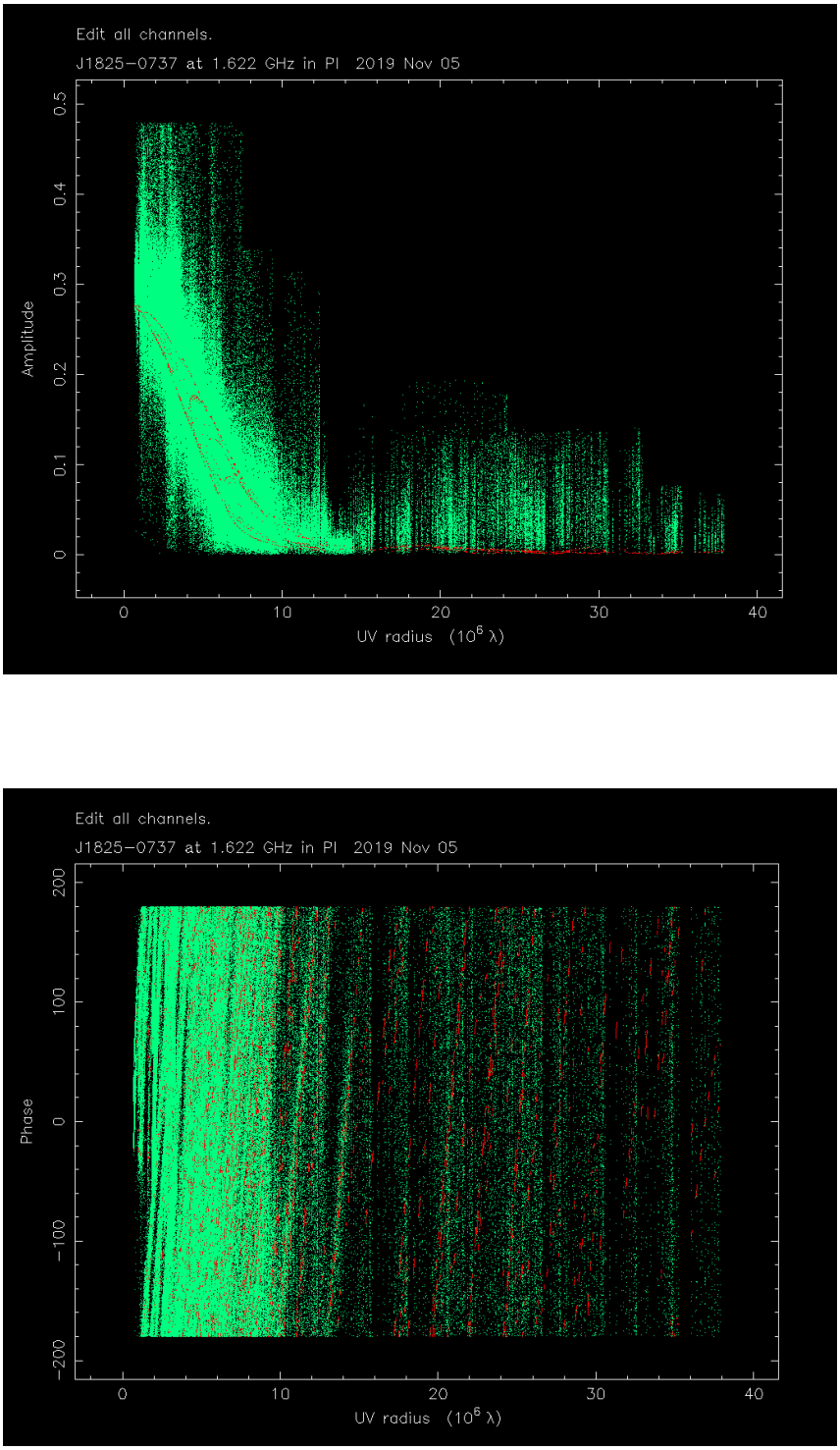


Figure 27: *Top*: In green, amplitude of J1825-0737 as a function of the uv distance. In red, the model associated to the signal *Bottom*: In green, phase of J1825-0737 as a function of the uv distance. In red, the model associated to the signal

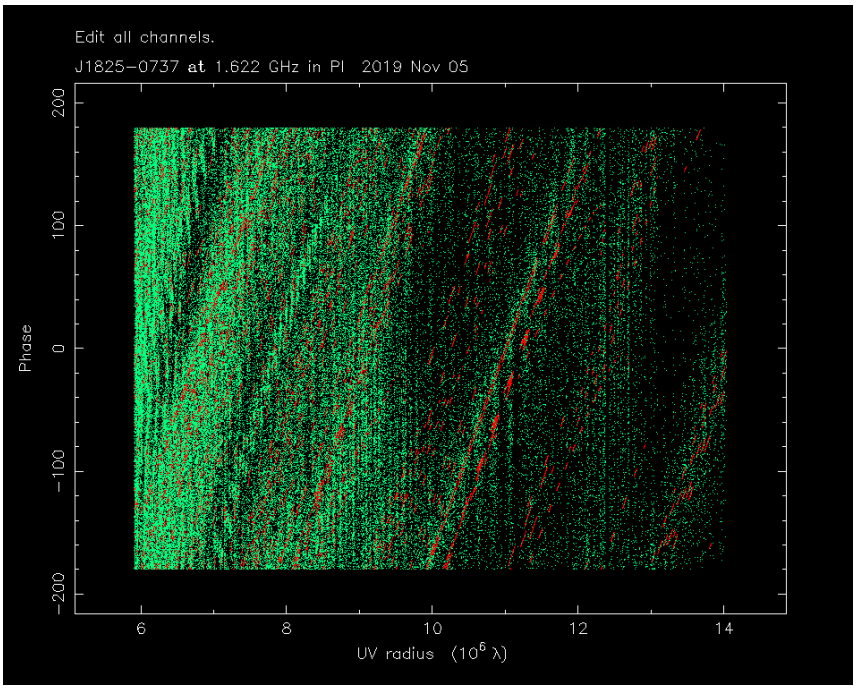
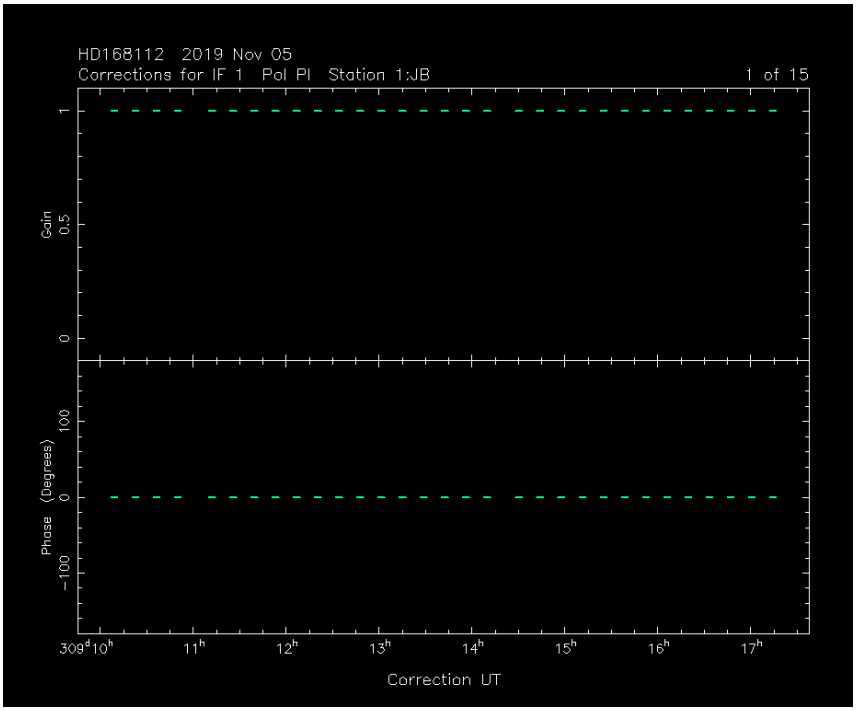


Figure 28: Zoom in the Figure 27 in between 6 and $14.10^6\lambda$.

However, even if the model is fitting the data as it can be seen inside previous Figures, one must be careful when performing the self-calibration step. Indeed, one must be sure that the self-calibration that is applied to the target's data is consistent with reality and that the self-calibration does not just correct randomly the data. If a random correction is applied to the data, the risk is that the self-calibration is not obtaining realistic results anymore and it is possible that some flux is forced inside the peak created by the target inside the image. Therefore, there is a risk that, if the self-calibration is not done properly, the obtained flux is over-estimated. Thankfully, there is a way to check if the self-calibration step is providing some realistic results or not thanks to what is called a *corplot*. This plot is representing the corrections that have been performed on the data during this step. Ideally, those corrections are constant over all the sub-bands of an antenna. It can vary a bit with time but the key point is the fact that, if the corrections are varying in the same way for all the sub-bands it means that the self-calibration is providing realistic results. However, if the corrections are more randomly distributed over the sub-bands, it can be considered as a strong hint that the self-calibration is providing some random corrections on the target's data and therefore that it is not reliable. A representation of a corplot associated to a reliable self-calibration as well as a corplot associated to an un-realistic self-calibration are shown in Figures 29 and 30, respectively.

Once the cleaning and the self-calibration steps have been performed, the images of the sources can be obtained.



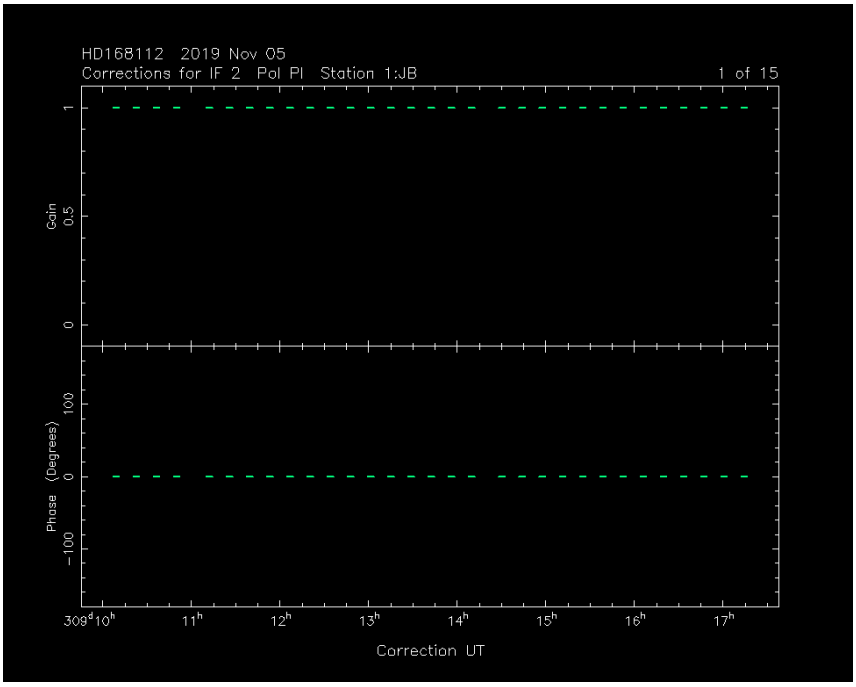
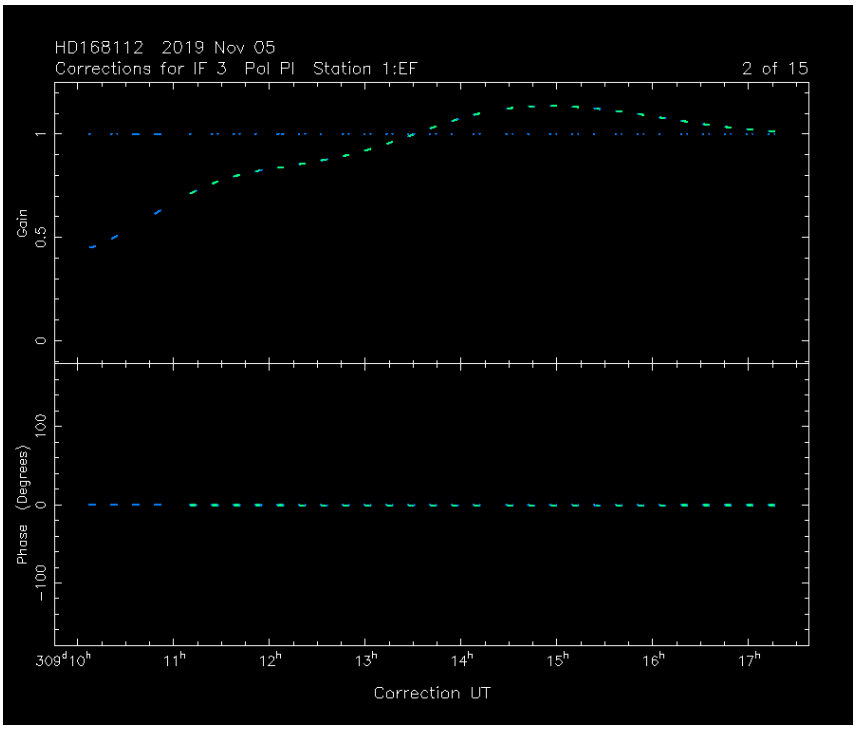


Figure 29: *Top*: Corplot for the first sub-band of the Jodrell Bank antenna in the UK (Jb) during a reliable self-calibration of HD 168112. *Bottom*: Corplot for the second sub-band of the Jb antenna during a reliable self-calibration of HD 168112.



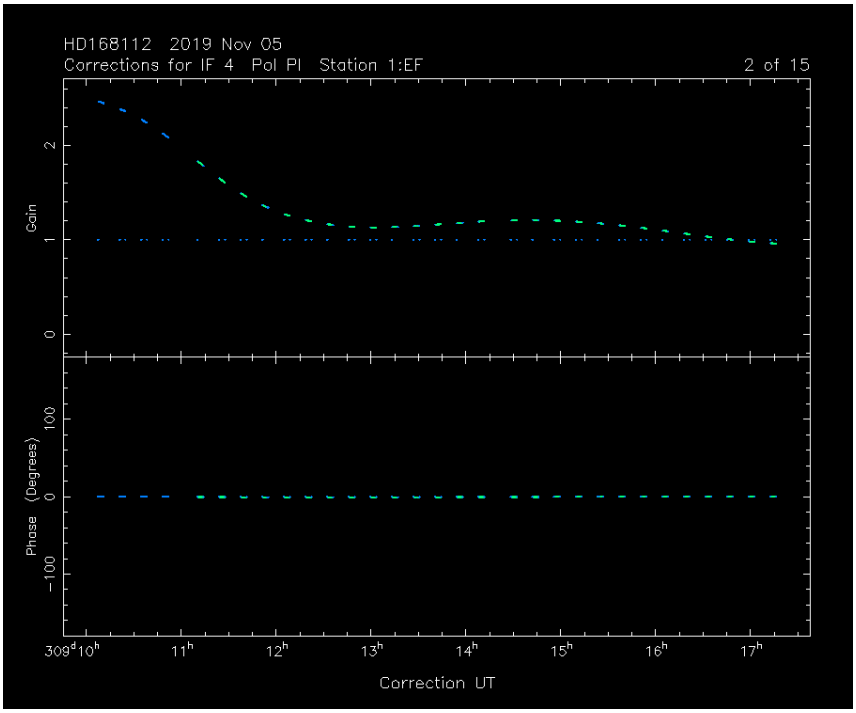


Figure 30: *Top*: Corplot for the third sub-band of the Ef antenna during a non-reliable self-calibration of HD 168112. *Bottom*: Corplot for the 4^{th} sub-band of the Ef antenna during a non-reliable self-calibration of HD 168112.

5.2 Imaging

In addition to all the calibration steps mentioned in the previous result section, the cleaning of the data is necessary in order to be able to image the sources (cfr. Sec. 3.4). The first image that can be obtain after both the cleaning and the self-calibration processes is the image of the phase calibrator, J1825-0737. It is therefore interesting to observe the differences between the dirty and cleaned images of the phase calibrator in order to check the efficiency of the two processes mentioned above. This comparison can be made thanks to Figure 31.

A first conclusion that can be made is that, in this case, the cleaning is very efficient. Indeed, nearly all the ripples created by the interferences of the signals that can be seen in the top plot of Figure 31 have been removed. This efficiency can be further confirmed if one looks at the color code. Indeed, in this case, the flux density of the source can be distinguished among the flux density of the noise even before the cleaning. This explains the fact that the removal of the features created by the correlation of the signals is efficient. Inside the cleaned image, one can see that there is approximately one order of magnitude between the flux density of the source and the remaining flux density of the noise (i.e. the flux density of the background). A more precise evaluation of the phase calibrator's flux density can be found in Table 5.

If one looks to the bottom plot of Figure 31, it is possible to still observe some features that are forming a vertical line above and below the source which are represented by the positive and negative contours. Those features are the footprints of the poor coverage that is often associated to VLBI observations. A zoomed figure of the phase calibrator can be found in Figure 32, in which the source is more visible. In this Figure as well as in Figure 31, several information are listed in green in the bottom of the image. Firstly, the coordinates (right ascension and declination) of the center of the map are given. The values mentioned in the axes are therefore the deviation from this particular position. Secondly, what is called the map peak is given in Jy.beam^{-1}. One must be careful to not get the map peak and the flux density mixed up. The map peak is the amount of flux that is contained inside a Dirac delta shape located at the position of the source while the flux density is the amount of flux integrated over the full extension of the source. This is why the map peak, also called the *peak brightness*, is given in Jy.beam^{-1} and the flux density in Jy. Finally, some information about the contours present inside the plot as well as the Full Width at Half Maximum (FWHM) of the *synthesized beam* are given. The synthesized beam is the projection of the beam with which the source is sampled. A representation of this beam is generally given in the bottom left corner of the images as it can be seen in the bottom plot of Figure 31 for example. It allows to observe that, with VLBI observations, it is possible to resolve the sources as the size of the beam is smaller or comparable to the size of the sources.

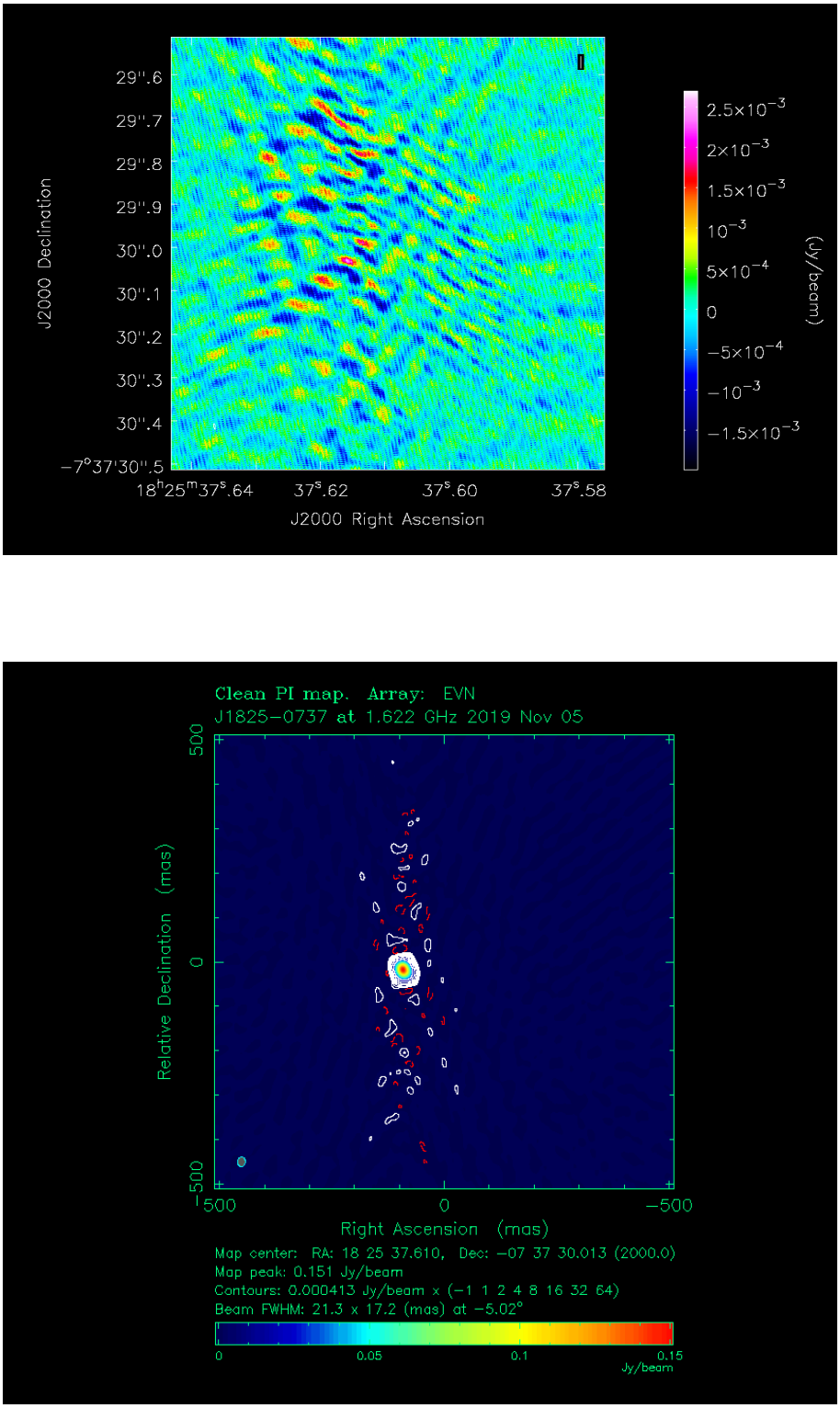


Figure 31: *Top*: Dirty image of the phase calibrator, J1825-0737. *Bottom*: cleaned image of the phase calibrator, J1825-0737, obtained after performing self-calibration and the CLEAN algorithm. The positive contours are represented in white while the negative contours are represented in red.

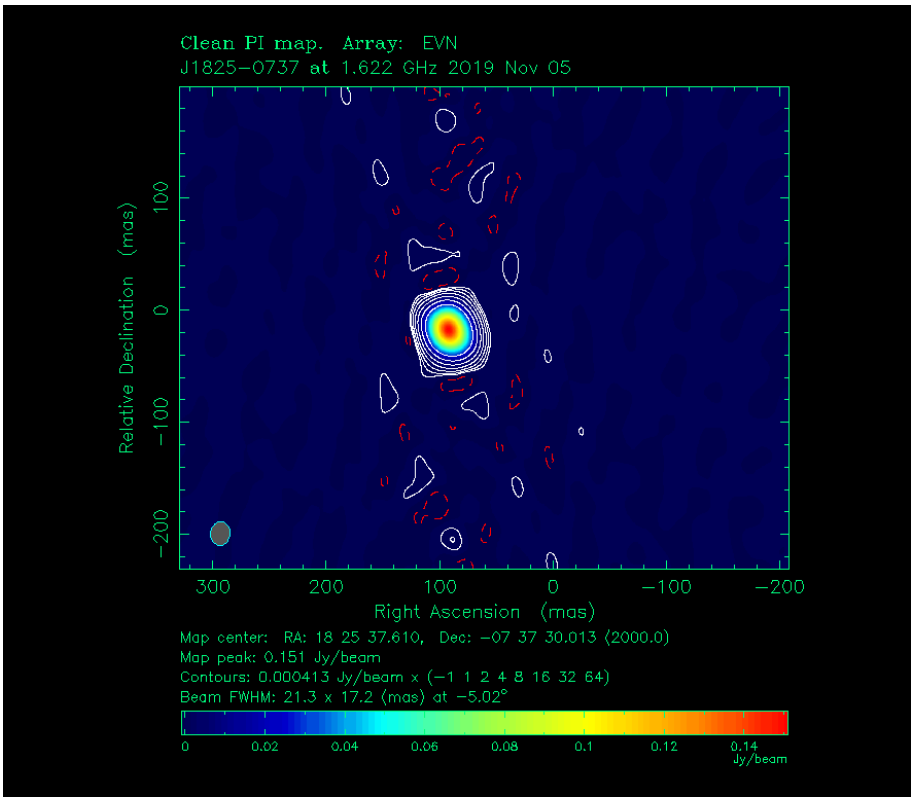


Figure 32: Zoom in the cleaned image of the phase calibrator, J1825-0737.

At this stage, it is possible to confirm that the assumption of J1825-0737 being a point-like source is reasonable as no specific shape can be observed inside Figure 32. If one closely observes Figures 31 and 32, it might seem strange that the value of the flux density (or the peak brightness) at the location of the source increased during the cleaning process. However, this is something which is expected when performing the CLEAN algorithm. During all the calibration processes, some errors propagated or some trends remained inside the data. Because of those effects, some flux has been removed from the source location along all the data processing. What the CLEAN algorithm is doing is to, firstly, remove some flux around the source location and, secondly, re-use a part of this flux to reconstruct the initial flux coming from the source. It is therefore expected to observe an increased flux density after the cleaning process. Thanks to the cleaned image of the phase calibrator, it is possible to obtain the observed flux density of J1825-0737 as well as the error on this flux density which is nothing more than the flux density of the background. Those values are listed inside Table 5.

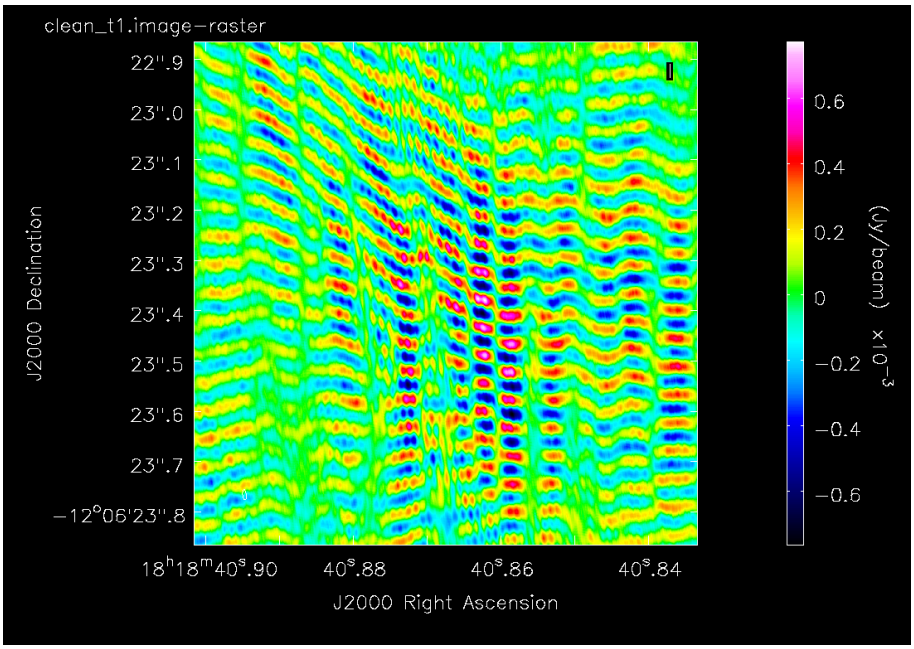
Parameter	
RA (J2000)	18:25:37.616
DEC (J2000)	-07:37:30.030
Frequency (GHz)	1.62249
S_ν (Jy)	0.275
Δ S_ν (Jy)	$6.29.10^{-4}$

Table 5: Parameters of J1825-0737 obtained thanks to the cleaned image. S_ν is the flux density and Δ S_ν is the error on this flux density. The given frequency is the frequency at which the observations have been performed.

Once the image, the various parameters and more importantly the model of the phase calibrator have been obtained thanks to self-calibration and some cleaning, it is time to obtain the images and parameters of both targets. The comparison between the dirty and the cleaned image of the first target, HD 168112, can be done using the Figure 33. The first comment that can be made on the cleaned image is, once again, concerning the efficiency of the cleaning. By combining the cleaning process with some self-calibration, a significant part of the noise has been removed. However, this is not as efficient as in the case of the phase calibrator. Indeed, the main difference is that, in this case, the SNR is lower which leads to a more difficult detection of the source inside the dirty image as already mentioned previously. Therefore, the cleaning is less efficient, which explains the presence of some remaining side lobes inside the background of the cleaned image. Nevertheless, it is important to stress here that even if the cleaning is not perfect, this step is of major importance, given the good quality of the final image and the fact that the source is barely visible inside the dirty image.

It is also important to stress at this point that one of the main difficulty of the cleaning is to find the source in the field of view. The prior knowledge on the position of the source is of great help for this task as well as the fact that the source is expected to be in the vicinity of the center of the dirty image, given the fact that the observational parameters have been correctly implemented in the observation plan.

In order to observe more details on the emission region, a zoom in the cleaned image is given in Figure 34. The same information than for the cleaned image of the phase calibrator are given in green in the bottom of Figures 33 and 34. Once again, thanks to an analysis of the cleaned image, is it possible to retrieve some valuable parameters that are displayed in Table 6.



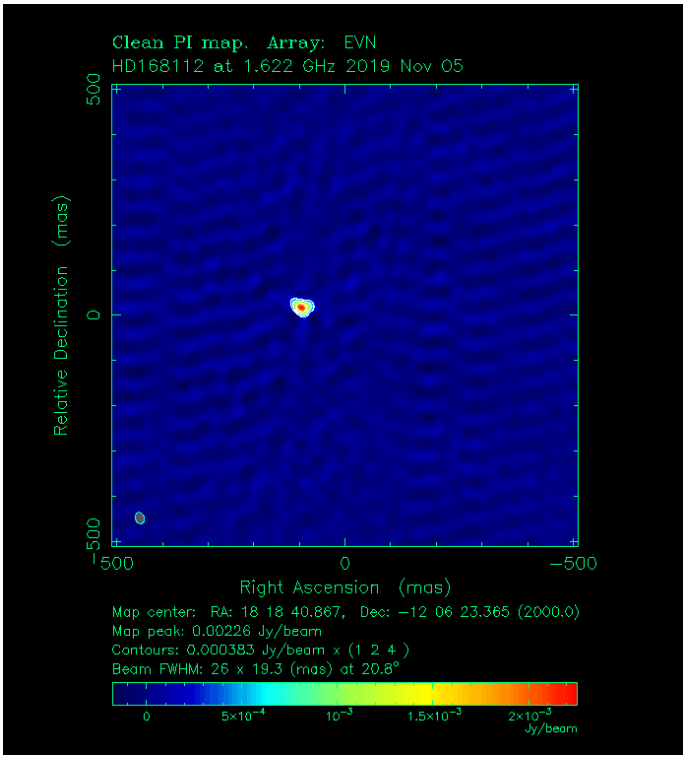


Figure 33: *Top*: Dirty image of HD 168112. *Bottom*: cleaned image of HD 168112, obtained after performing self-calibration and the CLEAN algorithm.

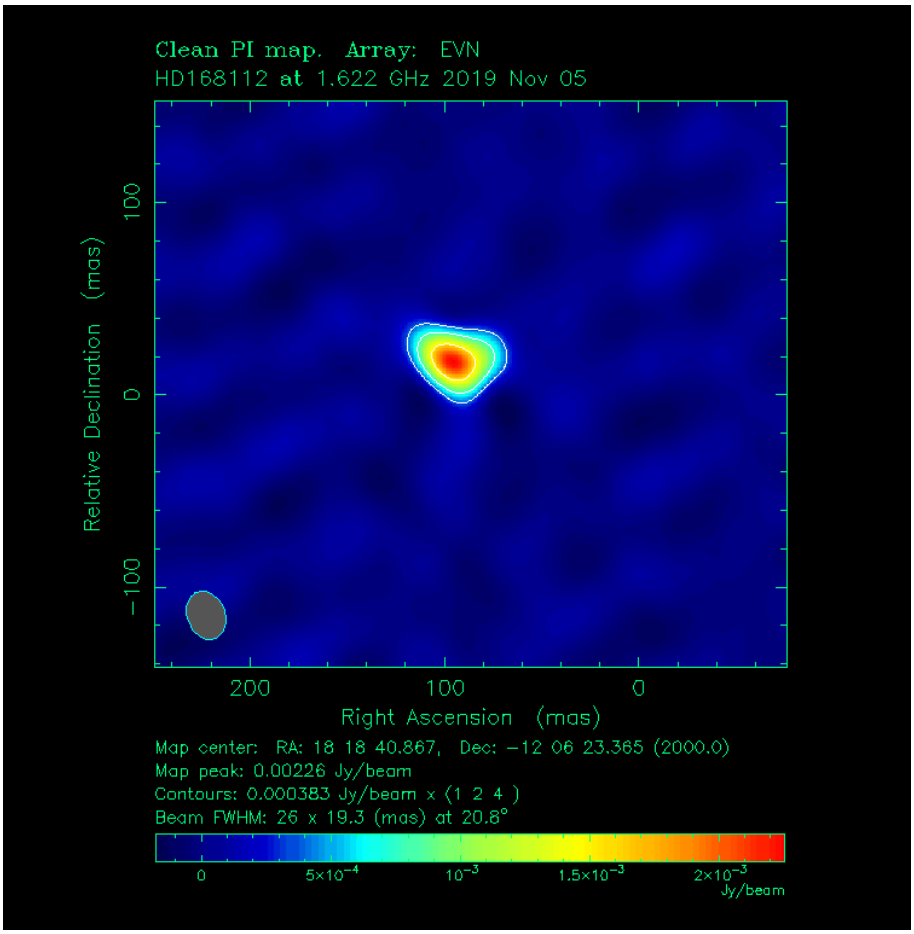


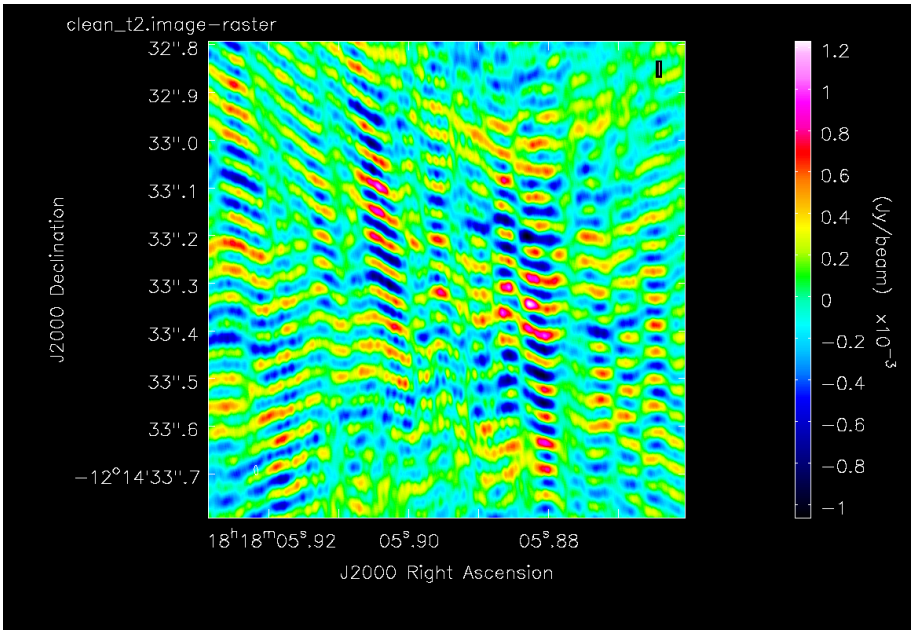
Figure 34: Zoom in the cleaned image of the first target, HD 168112.

Parameter	
RA (J2000)	18:18:40.874
DEC (J2000)	-12:06:23.348
Frequency (GHz)	1.62249
S_ν (mJy)	3.028
Δ S_ν (mJy)	0.59

Table 6: Parameters of HD 168112 obtained thanks to the cleaned image.

It is important to stress here that the error on the flux density, ΔS_ν, is made of two contributions. The first one being the flux density of the background while the second contribution accounts for a 10% error which is associated to VLBI measurements. Indeed, because of the un-perfect uv-coverage and the propagation of those errors during the calibration processes, it is common to add 10% of the final flux density to the value of the error for VLBI data.

Then, it is possible to obtain the same results for the second target, HD 167971, for which a comparison between the dirty image and the cleaned image can be done using Figure 35. The same discussion concerning the efficiency of the cleaning as for HD 168112 is valid because of both the low SNR associated to the dirty image and the good quality of the cleaned one. Once again, a zoom in the cleaned image can be found in Figure 36 in order to be able to observe more details in the emission region. Finally, some valuable parameters can also be obtained by using the cleaned image. Those parameters are gathered inside Table 7.



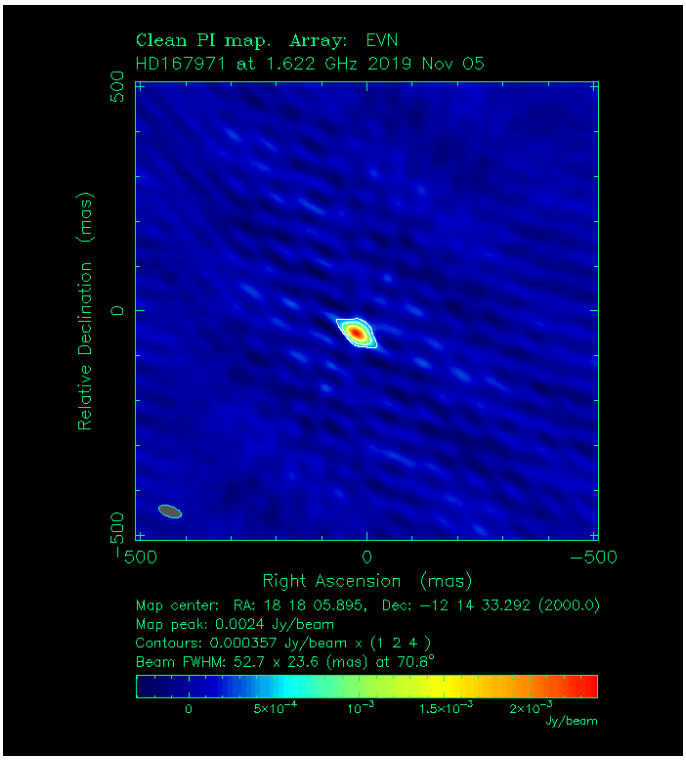


Figure 35: *Top*: Dirty image of HD 167971. *Bottom*: cleaned image of HD 167971, obtained after performing self-calibration and the CLEAN algorithm.

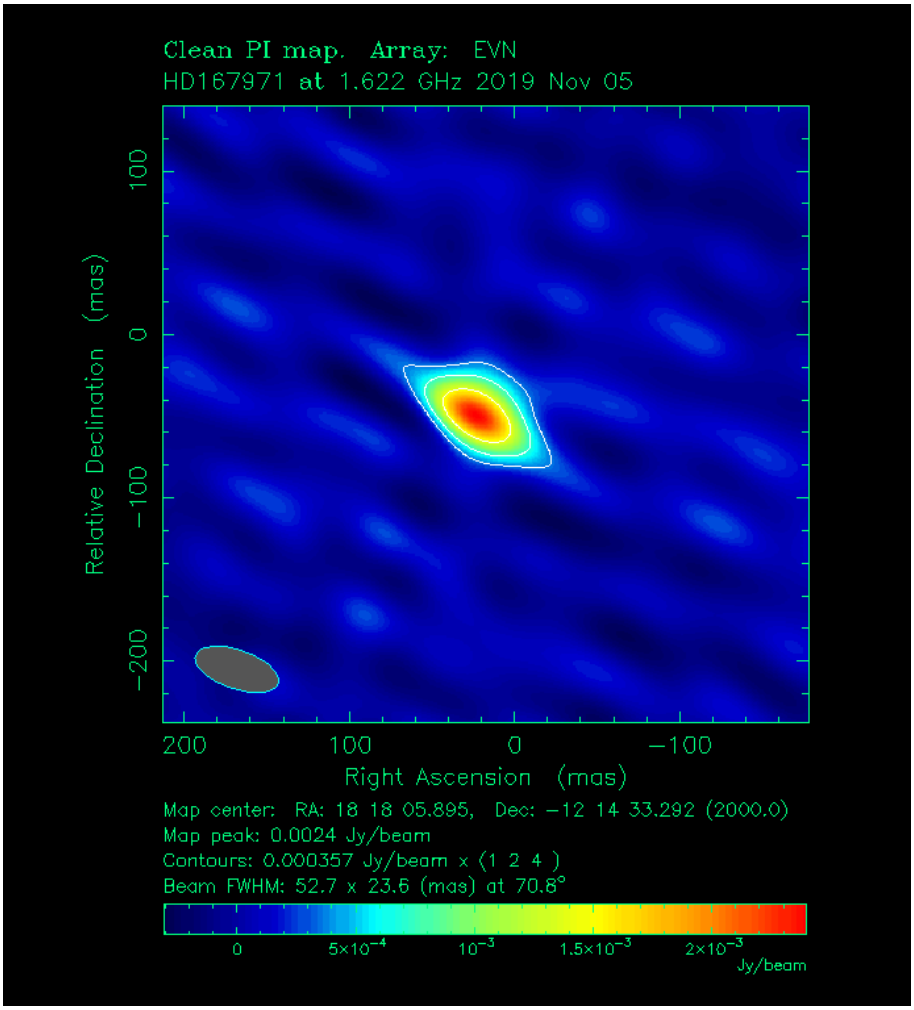


Figure 36: Zoom in the cleaned image of the second target, HD 167971.

Parameter	
RA (J2000)	18:18:05.896
DEC (J2000)	-12:14:33.342
Frequency (GHz)	1.62249
S_ν (mJy)	2.886
Δ S_ν (mJy)	0.381

Table 7: Parameters of HD 167971 obtained thanks to the cleaned image.

6 Discussion

6.1 Flux density - S_ν

6.1.1 Target 1 - HD 168112

The first important point to have in mind while discussing the results about HD 168112 is that, even if it is a known particle accelerator, this source has already been observed but not using VLBI measurements. However, it has been observed using the Very Large Array (VLA) (De Becker et al., 2004; Blomme et al., 2005). Therefore, it is already a valuable achievement to have detected this source using VLBI data and to be able to discuss the non-thermal nature of its emission given the low SNR associated to Colliding-Wind Binaries.

Concerning the obtained value of both the flux density and the error on this flux density (cfr. Table 6), the first conclusion is that the error is reasonable given the fact that there is an order of magnitude between the actual value of the flux density and its error. One can therefore stress again the efficiency of the self-calibration and the cleaning processes at this stage. As no other VLBI measurement is available for HD 168112, it is not possible to directly compare the value of the flux density obtained for this work with other values previously obtained. However, some VLA measurements are available. An important difference between VLA and VLBI data is that the resolution associated to VLA is of the order of $\sim$ arcseconds while the resolution associated to VLBI is of only a few milli-arcseconds, allowing to make the distinction between the thermal emission coming from the winds and the non-thermal emission coming from the interaction region in the case of Colliding-Wind Binaries. In the case of HD 168112, the semi-major axis of the synthesized beam is equal to 25.96 milli-arcsecond while the semi-minor axis is equal to 19.31 milli-arcsecond. It would translate as an overall resolution of $\sim$ 20 milli-arcsecond. It means that, if one wants to compare the flux density obtained inside this work to VLA measurements, one must take into account the fact that there are three contributions inside the VLA data as the interaction region is not resolved. Firstly, there is of course the non-thermal contribution coming from the interaction region. In addition, there are the two thermal contributions coming from the stellar winds of the binary system. In order to compare VLA and VLBI results, it is therefore interesting to estimate the thermal contribution coming from the winds for two different purposes that will be explained later.

The first purpose of an estimation of the radio thermal contribution of the stellar winds is concerning the detection threshold of VLBI data. By determining this contribution, it is possible to see if this emission will be detected in the VLBI data, given the detection threshold associated to VLBI measurements. One has to keep in mind that this thermal contribution is coming from the stellar winds of the two components of the system. Therefore, one can reasonably expect that this thermal contribution will be linked to the stellar winds parameters.

In order to estimate this contribution, one can make use of the development and the formula given in Leitherer et al. (1995) which is a rewriting of the radio flux of an expanding homogeneous stellar wind obtained by Panagia, Felli (1975) and Wright, Barlow (1975).

$$S_{\nu,th} = 2.32 \times 10^4 \left(\frac{\dot{M}Z}{v_\infty \mu}\right)^{4/3} \left(\frac{\gamma g_\nu \nu}{d^3}\right)^{2/3} \tag{38}$$

where $S_{\nu,th}$ is the thermal flux density of the wind in mJy, ν the frequency in Hz, d the distance to the source in kpc, $\dot{M}$ the mass loss rate in $M_\odot\ yr^{-1}$, v_∞ the terminal velocity of the stellar wind in km s^{-1}, μ the mean molecular weight of the wind, Z the rms ionic charge, γ the mean number of electrons per ion and g_ν the free-free Gaunt factor. The numerical factor in the right-hand side of equation 38 is coming from the units choice for all the parameters. Among all those parameters, the frequency, the mass losses and the terminal wind velocities are already known and are displayed inside Tables 1 and 2. Using A_i, M_i and Z_i as the fractional abundance, molecular weight and ionic charge of the i^{th} atomic/ionic species, one can define μ, Z and γ as

$$\mu = \frac{\sum A_i M_i}{\sum A_i}, \qquad Z = \frac{(\sum A_i Z_i^2)^{1/2}}{\sum A_i}, \qquad \gamma = \frac{\sum A_i Z_i}{\sum A_i} \tag{39}$$

By using the assumption that the stellar winds are made of 75% of hydrogen and 25% of helium, the value of the mean molecular weight of the wind (which is an often used standard value) is μ=1.75. By using the assumption that the helium is singly ionized among the stellar photospheres (Leitherer et al., 1995), it yields straightforward values of Z=1 and γ=1. The last parameter that must be computed is the free-free Gaunt factor that can be approximated by (Leitherer, Robert, 1991)

$$g_\nu = 9.77 \left(1 + 0.13 \log \frac{T_e^{3/2}}{Z\nu}\right) \tag{40}$$

where T_e is the electron temperature of the wind. This electron temperature is characterising the Maxwell-Boltzmann distribution of the electrons and is given in Kelvin. The frequency is once again given in Hz inside equation 40. This electron temperature can be approximated by (Drew, 1990)

$$T_e \sim 0.3\, T_{eff} \tag{41}$$

where T_{eff} is the effective temperature of the star that can be obtained thanks to the estimations of Muijres et al. (2012). The above relation is valid for a few stellar radii distance from the center of O-type stars. Therefore, it is valid in the external regions where the thermal radio emission is produced. Before computing all those parameters, a last point is important to mention.

The spectral types of the components inside the HD 168112 system are taken from Maíz Apellániz et al. (2019). In this article, two different spectral types are proposed. The first one is displayed in Table 1 and the second one is O4.5 III + O5.5 IV. Among the two possibilities of spectral types, the one with an O4.5 III primary star is likely to give rise to a larger thermal radio emission given the values of other parameters such as the mass loss or the terminal wind velocity. Therefore, this is the second spectral type option that will be used in the computation of $S_{\nu,th}$ in order to obtain the higher estimation of this thermal flux density. As no parameters are given for a spectral type of O4.5 inside Muijres et al. (2012), an average value in between the values for the O4 and O5 spectral types will be used. All the parameters used for the computation of the thermal flux of the two stellar winds of the system are gathered in Table 8.

Parameters	Unit	Value	Source
$\dot{M}$(A)	$M_\odot$ yr^{-1}	2.61 10^{-6}	[1]
$\dot{M}$(B)	$M_\odot$ yr^{-1}	1.36 10^{-6}	[1]
V_∞(A)	km s^{-1}	2702.5	[1]
V_∞(B)	km s^{-1}	2852	[1]
T_e(A)	K	12 148.95	[1]
T_e(B)	K	11 400.9	[1]
distance	kpc	1.7	[2]

Table 8: Parameters used during the computation of the thermal flux density of the stellar winds inside the HD 168112 system. The index [1] is referring to Muijres et al. (2012) while the index [2] is referring to Reipurth (2008),

Using all the parameters gathered in Table 8 as well as equation 38, one obtains a thermal flux density of 0.016 mJy for the primary component and a thermal flux density of 0.006 mJy for the secondary component considering an observation made with a wavelength of 18cm. The last point to consider is the fact that equation 38 is assuming an homogeneous stellar wind. However, this is not always the case and some denser clumps inside the wind are likely to increase the radio thermal flux density. Indeed, the *emission measure (EM)* can be defined as

$$EM = \int_{los} n_e^2 \, ds \tag{42}$$

where n_e is the number density of free electrons on a given distance along the line of sight. Therefore, a denser clumps inside the wind would increase this emission measure that would increase the flux density. As a result, it makes sense to introduce the *clumping factor*, f_{cl}, that will multiply the thermal flux density of an homogeneous stellar wind following the relation

$$S_{\nu,clumpy} \approx S_{\nu,homogeneous} \cdot f_{cl}^{2/3} \tag{43}$$

where the clumping factor is often approximated to $f_{cl} \approx 4$, which is valid inside the external part of the stellar wind (Runacres, Owocki, 2002). Taking into account the fact that the winds can be fragmented leads to a thermal flux density of 0.041 mJy for the primary component and a thermal flux density of 0.016 mJy for the secondary component considering an observation made with a wavelength of 18cm.

Those values can therefore be compared to the detection threshold of VLBI measurements in order to see if the thermal contributions coming from the winds should be detected or not. The typical detection threshold is $\sim 6\sigma$ for VLBI measurement. It means that a flux density of 6 times the background noise level is needed in order to claim a robust detection, if the observations are carried out without any prior knowledge on the position of the source. As in this case, there is some prior knowledge on the position of the targets, the threshold can be reduced down to $\sim 3\sigma$ (Deller, Middelberg, 2014). In this work, the level of the background noise has been used in order to obtain the error on the flux density of the target. It means that, in the case of HD 168112, the detection threshold would be at ~ 1.77 mJy.

Those values would mean that the radio thermal emission coming from the stellar winds of the two stars of the HD 168112 system would not be detected with the VLBI measurements as the sum of the thermal flux densities at $\lambda = 18$cm is equal to 0.057 mJy which is more than one order of magnitude below the detection threshold. This conclusion ensures that the emission which is detected in the frame of this work is the non-thermal synchrotron emission which is coming from the interaction region where the two stellar winds collide.

The second purpose of the estimation of the thermal flux density of the winds in the HD 168112 system is to estimate the non-thermal flux density on the basis of lower angular resolution measurements. Indeed, as mentioned previously, three components are contributing to the VLA measurements. Those three contributions are the non-thermal emission coming from the interaction region and the two thermal contributions coming from the stellar winds. As an estimation of those two thermal contributions has been obtained above, it is possible to obtain an estimation of the non-thermal contribution by subtracting the thermal contributions from the flux density obtained thanks to VLA measurements. In order to do so, the flux density values obtained in Bieging et al. (1989), Phillips, Titus (1990) and De Becker et al. (2004) can be used. Among those studies, only two measurements, at two different epochs (2010 and 2011), have been made using the same frequency as this work and those two measurements are not precise values but rather upper limits of 2.4 mJy including both the thermal and the non-thermal emissions. However, inside Bieging et al. (1989) and De Becker et al. (2004), observations at two different frequencies have been made for the same epoch. This allows to obtain a spectrum of the non-thermal emission and to see where does the measurement obtained in this work falls into the spectrum.

It is important to mention here that the estimated value of the thermal flux mentioned above is only valid for an observation made with $\lambda = 18$cm. Because the data points from Bieging et al. (1989) and De Becker et al. (2004) have not been obtained at the same wavelength, the thermal flux densities must be estimated once again, using the appropriate frequency. This is represented in Figure 37.

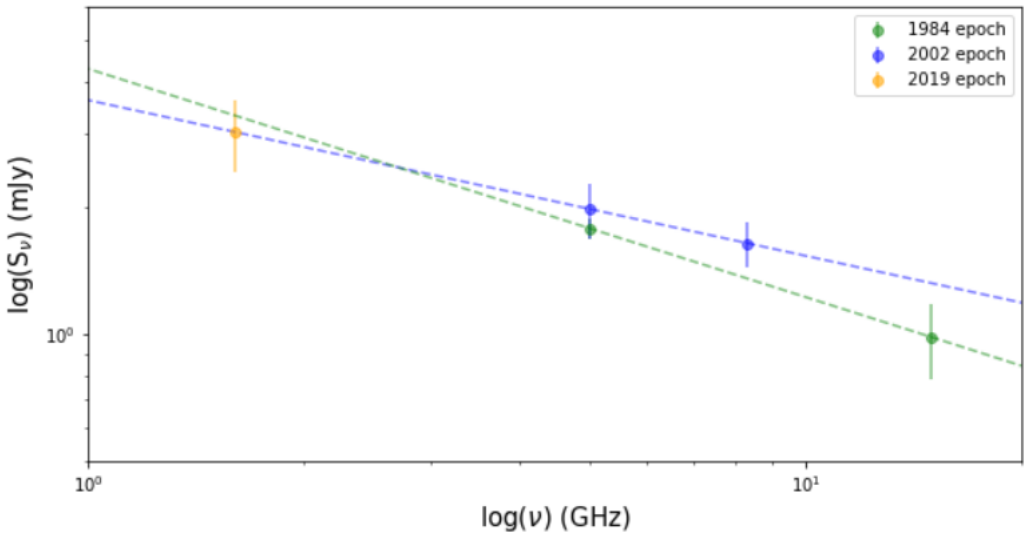


Figure 37: Flux density as a function of the frequency for HD 168112 obtained thanks to the estimation of the thermal flux density. The data points from the 1984 epoch are drawn from Bieging et al. (1989) while the data points from the 2002 epoch are drawn from De Becker et al. (2004).

Several comments are needed at this stage. Firstly, one must be careful with the representation of Figure 37. Indeed, all the data points displayed in this plot have not been taken at the same orbital phase. Because no orbital solution is yet available for the HD 168112 system, the orbital period is unknown and it is impossible to constrain the orbital phase of the observations. Therefore, the purpose of Figure 37 is to have an idea of the potential connection between the different observations rather than really comparing data points obtained at the same orbital phase. This limitation coming from the orbital phase is important to have in mind while inspecting Figure 37.

The second comment that can be made is the fact that, despite the approximations that have been made during the estimation of the thermal flux density, the measurement of the 2019 epoch falls, within the error bars, inside the range of values expected from previous epochs. This is still an important achievement to see that the value obtained in this work is consistent with the expected range of values, even if the orbital phase is not constrained.

6.1.2 Target 2 - HD 167971

Unlike HD 168112, this system has already been observed several times using VLBI measurements. As mentioned in section 4, the system has been observed at two different orbital phases and, thanks to observations at different frequencies, a spectral index has been computed (Sanchez-Bermudez et al., 2019). It will therefore be interesting to compare the results obtained in this work from a point of view of the obtained flux density value but also from the point of view of the convenience with the obtained spectral index.

As for the case of HD 168112, the self-calibration and the cleaning steps have been particularly efficient given the final cleaned image of the source (cfr. Figure 35). In this case, the value of the error on the flux density is even lower than for the first source which proves once again the efficiency of the data processing. However, in Figure 36, a number of bright and dark dots are still visible. Those dots are in fact representing some sinusoidal patterns that remain despite the data processing. As for the case of HD 168112, those patterns are the footprints of the poor uv-coverage associated to VLBI measurements.

In order to be able to compare the results obtained in the frame of this work with previous measurements, it is important to compare the various frequencies used during each measurements. Indeed, a simple comparison of the flux density value would not have any physical sense since the flux density is varying as a power law of the frequency. The comparison between the results of this work and the results of Sanchez-Bermudez et al. (2019) can be found in Table 9 and a visualization of the data is represented in Figure 38.

	2006 epoch		2016 epoch		2019 epoch
Parameters	X-band	C-band	X-band	C-band	L-band
Frequency (GHz)	8.4	5	8.4	5	1.6
λ (cm)	3.5	6	3.5	6	18
S_ν (mJy)	5.324	12.314	2.817	7.316	2.886
ΔS_ν (mJy)	0.201	0.294	0.16	0.18	0.092

Table 9: Comparison between the results of Sanchez-Bermudez et al. (2019) (2006 and 2016 epochs) and the results of this work (2019 epoch) for HD 167971.

It is important to mention here that the errors on the flux densities for the 2006 and 2016 epochs were not given as such by Sanchez-Bermudez et al. (2019). Those errors were given in Jy.beam^{-1}.

Therefore, in order to obtain an approximated value of the error on the flux density, the errors given in Jy.beam^{-1} have been multiplied by a certain factor, given the size of the synthesized beam and the size of the emission regions displayed in the images of Sanchez-Bermudez et al. (2019). Those factors have been chosen in such a way that, once multiplied by this factor, the synthesized beam was approximately covering the full extension of the emission region.

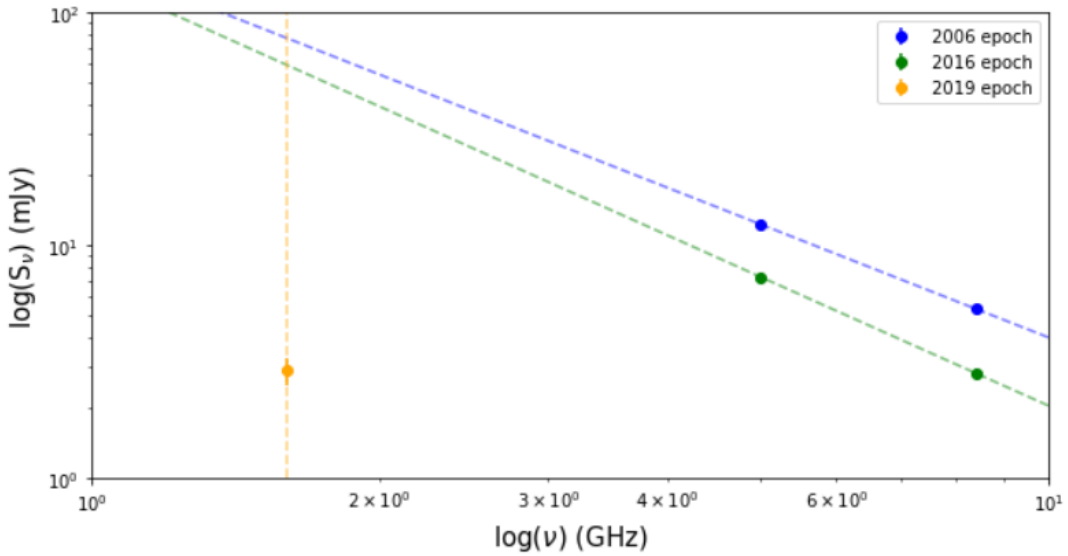


Figure 38: Flux density as a function of the frequency in logarithmic scales for observations of HD 167971 at different epochs.

Several information can be drawn from Figure 38. First of all, error bars are included for the three epochs but are barely visible because of their low amplitude in the logarithmic scale. Concerning the two epochs mentioned in Sanchez-Bermudez et al. (2019), one can indeed see that two different frequencies have been used during the observations. In addition, it is important to stress that the expected negative slope is observed. Indeed, the flux density is following a power law of negative index as a function of the frequency, as it can be seen in Figure 3. The following relation, for an optically thin synchrotron emission, can be written as

$$S_\nu \propto \nu^{-\alpha} \tag{44}$$

where α is the spectral index. By using this relation, it is easy to understand that, when using logarithmic scales as in Figure 38, the expected trend is a linear decrease. By using the 2006 and 2016 epochs used in Sanchez-Bermudez et al. (2019), two slightly different slopes (i.e. two different spectral indices) can be obtained.

Nevertheless, at first sight, the most striking information that can be drawn from Figure 38 is the fact that the data point obtained in the framework of this work is absolutely not compatible with the expected power law trends extrapolated from the two previous epochs. However, in order to understand those results, one must realise that all those measurements are corresponding to a different orbital phase of the system. In addition, a variable radio emission has been detected by Blomme et al. (2007). In other words, different orbital phases will correspond to different locations inside the radio lightcurve obtained by Blomme et al. (2007) and a different radio flux is expected for each orbital phase. It would be therefore interesting to compute the orbital phase associated to the measurement of the 2019 epoch to see where does this measurement stand inside the radio lightcurve and the representation of the orbit (cfr. Figure 39). It can be done using the orbital parameters gathered in Table 10.

Parameter	Unit	Value
T_0	MJD	54 736
P	days	7 806
ϕ		0.52

Table 10: Orbital parameters used in order to compute the orbital phase, ϕ, associated to the observation of the 2019 epoch. The time of the reference periastron passage, T_0, and the period, P, are taken from Le Bouquin et al. (2017).

By combining the computed orbital phase and the Figure 39, one obtains that the epoch is nearly consistent with the apastron, i.e. the position of maximum stellar separation. In order to see if this orbital phase is consistent with a smaller flux density, it is interesting to use the radio lightcurve that has been obtained by Blomme et al. (2007). In fact, several lightcurves at different wavelengths of observation have been computed. The λ=6cm lightcurve is used for the sake of this discussion as this is the lightcurve containing the largest number of data points, therefore being more reliable for the discussion. In the frame of Colliding-Wind Binaries, there are two processes that are acting against each other, the intrinsic intensity of the synchrotron emission modulated by the stellar separation and the free-free absorption of this emission modulated by the column density of the stellar wind material along the line of sight. Concerning the former, the production of synchrotron emission depends on the strength of the magnetic field as this field is needed to produce particle acceleration. However, for an increasing stellar separation, the magnetic fields coming from the two stellar components will be submitted to a geometric dilution and are therefore expected to lead to a smaller synchrotron flux density.

Concerning the latter process, i.e. the free-free absorption, it is likely to obtain a more efficient free-free absorption in the vicinity of periastron. Indeed, in that part of the orbit, the synchrotron emission might be embedded inside dense stellar winds which are likely to give rise to an efficient absorption of the emission. It is important to be careful with this statement as the efficiency of the absorption is notably dependent on the inclination of the plane of the orbit. Some configurations could give rise to a significant detection of synchrotron emission even at the periastron, if the angle in between the line of sight and the plane of the orbit is near to 90 degrees for example. Because of this effect, the maximum efficiency of the free-free absorption is not necessarily expected exactly at periastron passage, but rather in the vicinity of periastron. This is where the use of the 6 cm radio lightcurve is of major importance. By subtracting the estimated period of the system (∼21.4yrs) to the date of the observation used in this work, one obtains that the observation is equivalent with an observation that would have been made in 1997. Concerning the value of the flux density, as the wavelength used for the computation of the lightcurve (6 cm) and the wavelength used for this work (18cm) are not the same, an extrapolation is needed. By using the fact that the flux density is proportional to the frequency to a given power which is equal to the opposite of the spectral index, one can write

$$S_{\nu_{18}} \propto \nu_{18}^{-\alpha} \tag{45}$$
$$S_{\nu_6} \propto \nu_6^{-\alpha} \tag{46}$$

where $S_{\nu_{18}}$ and S_{ν_6} are respectively the flux densities measured at a wavelength of 18 cm and 6 cm, ν_{18} and ν_6 the frequencies associated to a wavelength of, respectively, 18 cm and 6cm and where α is the spectral index. By dividing those two relations by each other, it allows to get rid of the identical proportionality factors and to obtain

$$\frac{S_{\nu_{18}}}{S_{\nu_6}} = \left(\frac{\nu_{18}}{\nu_6}\right)^{-\alpha} \tag{47}$$

in which $S_{\nu_{18}}$ is the flux density obtained for this work, ν_{18} is equal to 1.6 GHz and ν_6 is equal to 5 GHz. It is therefore possible to obtain the value of the flux density at 6 cm, which is an extrapolation of the measure obtained in this work, provided that the value of the spectral index is known. The theoretical value for synchrotron emission in the optically thin regime is $\alpha = 0.5$. However, in Sanchez-Bermudez et al. (2019), a value of $\alpha = 1.1$ has been obtained. It is therefore possible to compute the flux density at 6 cm using two approaches, each one considering a different value of the spectral index. By doing so, it is possible to place two data points inside the radio lightcurve of Figure 40, corresponding to the observation of the 2019 epoch.

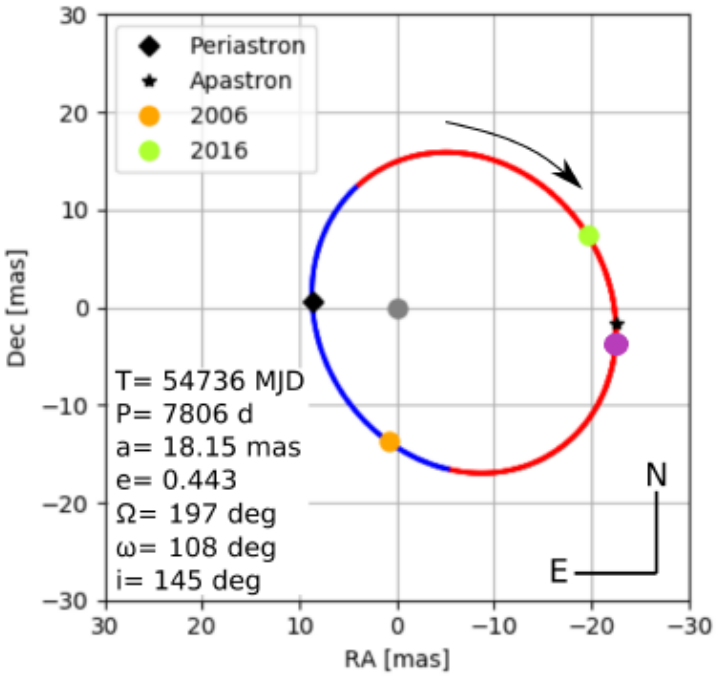


Figure 39: Orbit of the outer star with respect to the two central stars of the HD 167971 system projected in the plane of the sky. The purple dot is representing the orbital phase of the 2019 epoch. Figure modified from Sanchez-Bermudez et al. (2019).

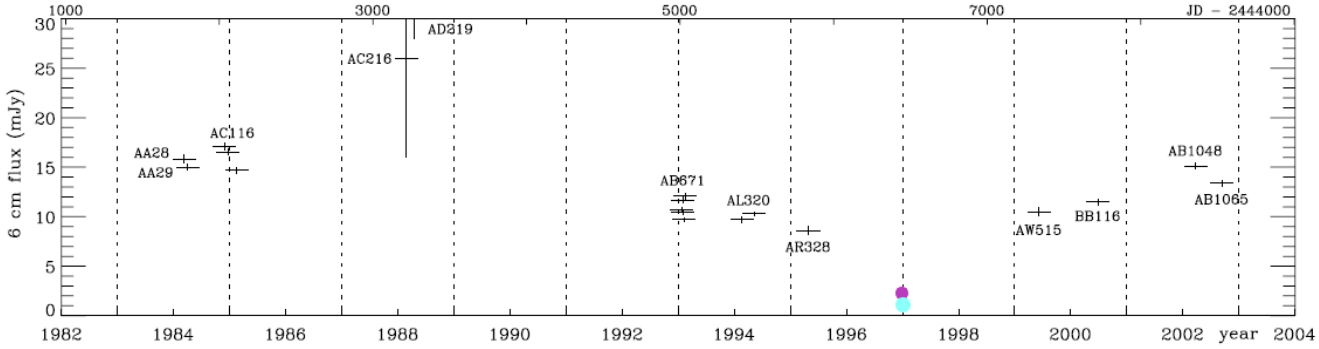


Figure 40: Representation of the 2019 epoch inside the radio lightcurve obtained by Blomme et al. (2007). The purple dot has been obtained by considering a theoretical value of 0.5 for the spectral index while the light blue dot is associated to the value of 1.1 obtained by Sanchez-Bermudez et al. (2019). The purple dot is at a flux density of 1.63 mJy while the light blue dot is at a flux density of 0.82 mJy.

Several information can be drawn from Figure 40. Firstly, the two colored dots, which correspond to an orbital phase consistent with the apastron, fall clearly in a minimum of the lightcurve. Secondly, by subtracting half a period to the 1997 epoch, one obtains the 1987 epoch. It is visible that this epoch, which is therefore consistent with the periastron, is associated to a maximum of the radio lightcurve. This second observation leads to a crucial information. Because the maximum of the radio lightcurve is observed at an orbital phase consistent with the periastron (i.e. the position of minimum angular separation), it means that, despite the very dense stellar winds present at this orbital phase given the short stellar separation, the free-free absorption is not dominating the intrinsic intensity of the synchrotron emission. In other words, despite the competition between the free-free absorption and the intrinsic synchrotron emission of the interaction region, the maximum flux density is expected in the vicinity of the periastron while the minimum flux density is expected in the vicinity of the apastron.

This conclusion is of great importance as it is fully consistent with the results of this work. Indeed, the fact that the flux density measured in the frame of this work is significantly lower than the expected values by the 2006 and the 2016 epochs of Figure 38 can be explained by the fact that the 2019 epoch is consistent with the apastron passage. A significantly lower flux density is therefore completely consistent with the previous VLBI measurements taken in between the periastron and the apastron.

The following interesting result that can be obtained is concerning the thermal flux density of the stellar winds. As it has been done for HD 168112, the thermal flux density of the stellar winds inside the system can be estimated using equations 38 and 40. The various parameters that have been used in order to obtain the thermal flux density are gathered inside Table 11. In this case, three stellar components have to be taken into account as there are two stars in the center of the system around which a third star is orbiting.

By using the parameters gathered inside Table 11 and the exact same methodology as for the HD 168112 system, the thermal flux densities of the three stellar components can be computed. For an observation made at $\lambda = 18$cm, a flux density of 0.9 10^{-3} mJy is obtained for the Aa component, 0.86 10^{-3} mJy is obtained for the Ab component and 2.21 10^{-3} mJy for the B component. By summing all those three contributions, one obtains that the overall thermal contribution of the stellar winds amounts to 3.98 10^{-3} mJy. One can see that there is an order of magnitude of difference between the estimated thermal flux density for HD 168112 and HD 167971. However, both equidistant systems are made of O-type stars and one can therefore expect the same order of magnitude for the thermal flux density of the two systems. The explanation here lies in the mass loss rates displayed in Table 8 and Table 11. Indeed, by considering the order of magnitude of difference for $\dot{M}$ as well as equation 38, the difference between the obtained thermal flux densities can be explained.

By using the same method to estimate the detection threshold in this case, one obtains that this threshold amounts to $\sim$ 1.143 mJy. Therefore, the conclusion is the same than for HD 168112. It means that, as the thermal contribution is nearly three orders of magnitude smaller than the detection threshold, this thermal emission is not detected. This ensures that the emission which has been detected in this work is the synchrotron emission coming from the interaction region.

Parameters	Unit	Value	Source
$\dot{M}$(Aa)	$M_\odot$ yr^{-1}	2.48 10^{-7}	[1]
$\dot{M}$(Ab)	$M_\odot$ yr^{-1}	1.42 10^{-7}	[1]
$\dot{M}$(B)	$M_\odot$ yr^{-1}	3.26 10^{-7}	[1]
V_∞(Aa)	km s^{-1}	4408	[1]
V_∞(Ab)	km s^{-1}	2594	[1]
V_∞(B)	km s^{-1}	2921	[1]
T_e(Aa)	K	10 046.1	[1]
T_e(Ab)	K	9069.3	[1]
T_e(B)	K	8529	[1]
distance	kpc	1.7	[2]

Table 11: Parameters used during the computation of the thermal flux density of the stellar winds inside the HD 167971 system. The index [1] is referring to Muijres et al. (2012) while the index [2] is referring to Reipurth (2008),

6.2 Imaging

A major result concerning the imaging of HD 168112 is the shape that can be observed in Figure 34. Indeed, some horizontal shape can be seen. However, it is possible that this shape can be created by instrumental effects. In order to verify if it is the case, the shape of the synthesized beam can be analysed. If one looks at the shape of the synthesized beam in Figure 34, one can see that the shape is a bit elongated in the vertical direction. This elongation is created by the uv-coverage available for the observation. At this stage, one can therefore stress that the shape of the emission region cannot be created by the shape of the synthesized beam as the two elongations are more or less perpendicular. Therefore, the shape of the emission cannot be a footprint of the poor uv-coverage of the observation.

As an additional check, it is interesting to compute the clean image of the source by randomly flagging one antenna of the network. By doing so, it allows to verify that the shape of the emission region is not created by the signal of a given antenna that can still be corrupted. If the horizontal elongation is still present inside the image of the source for which an antenna has been flagged, it is a strong hint, in addition to the above criteria, that the horizontal shape is coming from the emission region itself. This verification has been performed by flagging various antennas of the network. The computed images obtained after having flagged the Onsala antenna in Sweden (O8) in the first place and the Zelenchukskaya antenna in Russia (Zc) in the second place can be found in Figure 41. The conclusion that can be drawn from those two cleaned images is the fact that the horizontal elongation is still visible while no horizontal elongation is visible for the synthesized beam. In addition, the level of the noise as well as the value of the map peak is changing inside those two plots. It is expected since when an antenna is flagged, the uv coverage is changing, which is changing the properties of the final image.

Therefore, it is reliable to stress that this elongation is coming from the emission region itself. In fact, this elongation is coming from the collision between the stellar winds of the two components of the system. Indeed, and as showed inside Table 1, the two stellar wind terminal velocities and mass loss rates of the components are not equal. It means that one of the two wind is more energetic than the other. This quantity can be defined using the *stellar wind kinetic power* defined by

$$P_{kin} = \frac{1}{2}\dot{M}V_{\infty}^{2} \tag{48}$$

where $\dot{M}$ is the mass loss rate and V_{∞} is the stellar wind terminal velocity. Therefore, the two stellar winds involved inside the collision will not be characterized by an equal kinetic power. As a consequence, the interaction region will take the shape of a bow shock, which is what can be observed in Figure 34. Therefore, in addition to the achievement of detecting the non-thermal emission of HD 168112 using the VLBI technique, the shape of the interaction region is visible inside the cleaned image, which has not been achieved up to now for this source.

Unfortunately, no such observation concerning the shape of the interaction region can be made for the second target, HD 167971. Indeed, if one looks at Figure 36, there is an elongation in between the horizontal and the vertical direction. However, this elongation is fully consistent with the elongation that can be observed in the representation of the synthesized beam. Therefore, no conclusion regarding the detection of the shape of the interaction region for the HD 167971 system can be made as it is very likely to be a consequence of the uv-coverage that was available during the observation. It is also important to mention that the same verification by flagging some antennas has been made for HD 167971. When some antennas of the network are flagged, no shape which deviates from the shape of the synthesized beam can be observed, which ensures that no conclusion concerning the shape of the interaction region can be made in this case.

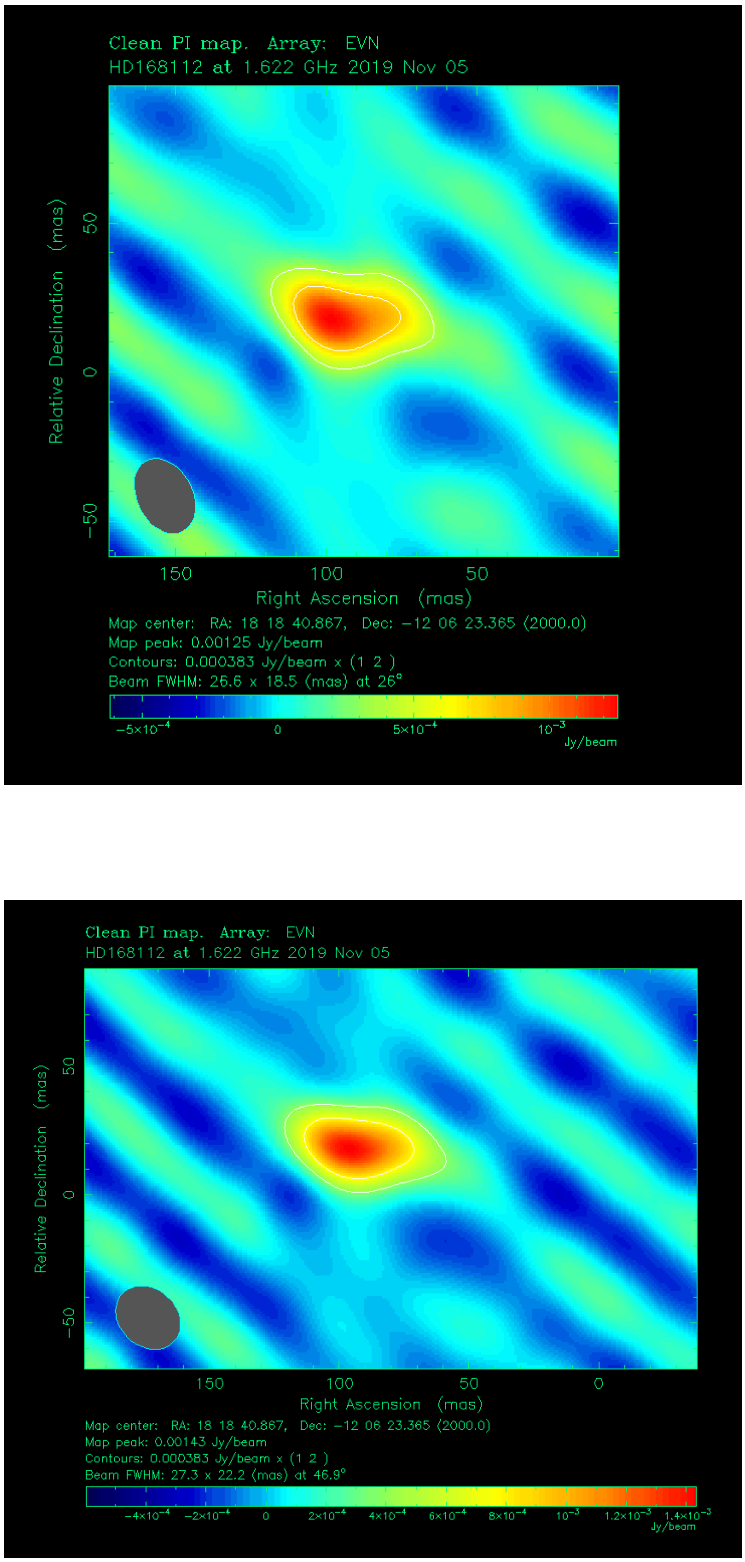


Figure 41: *Top*: Cleaned image of HD 168112 after the flagging of the O8 antenna. *Bottom*: Cleaned image of HD 168112 after the flagging of the Zc.

7 Conclusion

To conclude, several objectives have been achieved in this work. Firstly, the capacity of the European VLBI Network to detect and image the synchrotron radio emission region in PACWBs has been proved. Indeed, the emission has been detected for both systems. In addition, several conclusions allowed to enhance the reliability on the fact that the emission which has been detected is the non-thermal emission coming from the interaction region. For HD 168112, the thermal flux density of the system has been estimated and it has been shown that this estimation is way below the detection threshold (one order of magnitude below). In addition to this conclusion, the elongation of the emission that can be seen in Figure 34 and the discussion that is associated to this image showed that the emission is coming from a collision between the stellar winds in which particle acceleration occurred. For HD 167971, the same discussion concerning the estimation of the thermal flux density and the detection threshold is valid. Furthermore, there are 3 orders of magnitude between the thermal flux density and the detection threshold, enhancing the reliability of the detection of non-thermal emission. In addition to this discussion, the fact that the obtained value of the flux density is consistent with previous results obtained for this system helps to enhance again the reliability of the non-thermal detection. Therefore, the main goal of this work, which was to demonstrate the capacity of the EVN to detect (in both systems) and image (in one of them) such non-thermal emission in PACWBs has been greatly accomplished.

In addition to this main objective, the results of this work allowed to obtain several particular achievements for both systems. Concerning HD 168112, the obtained flux density accounts for the first measurement of the non-thermal flux density of this system without detecting the thermal contribution at the same time. This result has been made possible thanks to one indicator mentioned in section 1.4, the VLBI method. Then, the image obtained in this work accounts for the first radio image of this system using VLBI data. In addition, an other indicator mentioned in section 1.4 which is the elongation of the emission region that is visible in Figure 34, has never been obtained previously.

Finally, concerning HD 167971, the results of this work allowed to improve a bit the sampling of the long period orbit of the system from the point of view of the high angular resolution imaging. Thanks to the EVN observation of this system at the 2019 epoch, an observation of the system in the vicinity of the apastron is now available while the two previous VLBI observations were located in between the periastron and the apastron. This specific orbital phase is especially important as it coincides with a lower expected synchrotron emission. The clear detection presented in this work clarifies that the synchrotron emission region can be efficiently imaged at any orbital phase of the system. The results of this work are fully consistent with previous results (either at lower or high angular resolution) and are in very good agreement with the expectations from the general model for particle acceleration and non-thermal emission of radiation from PACWBs.

8 Prospects

8.1 Astrometry

The first prospect that arises after this work is considering the astrometry. Indeed, the topic of the astrometry has not been tackled in the specific framework of this work. The main idea would be to compute the precise astrometric positions of the emission regions detected in this work in order to verify that this emission is located in between the astrometric positions of the stars, which is expected for an emission region coincident with the wind-wind interaction region. However, some astrometric uncertainties can already be considered at this stage. Firstly, there is a theoretical astrometry uncertainty which is present when using interferometric data. Thompson et al. (1988) derived that this uncertainty can be written as

$$\sigma_{\alpha,\delta} = \frac{1}{2\pi}\frac{1}{\mathrm{SNR}}\frac{\lambda}{B} \tag{49}$$

where α and δ are the right ascension and declination of the source of interest, $\sigma_{\alpha,\delta}$ the error on the position, SNR the signal to noise ratio of the observation, λ the wavelength used during the observation and B the baseline length projected on the sky. Depending on the uncertainty which is considered (i.e. on the right ascension or on the declination), the uncertainty of equation 49 writes σ_α or σ_δ. However, the term λ/B in the above equation is basically describing the size of the Airy disk. In other words, it is describing the resolution that can be reached during the observation. In the frame of this work, the resolution was not fixed. Indeed, some weighting (natural or uniform) has been used during the imaging of the sources and it therefore altered the resolution. What can be done instead of directly using the ratio between the wavelength and the projected baseline is to use the size of the synthesized beam. As one has to obtain an astrometric uncertainty for both the right ascension and the declination, one must project the synthesized beam onto both directions. When describing the synthesized beam, three parameters are needed. The first two are the *semi-major axis* and the *semi-minor axis* that describe the ellipse which constitutes the beam. The last parameter, which is more a positional parameter, is the *position angle*. This angle is defined as the angle in between the North direction and the semi-major axis of the ellipse where the angle is increasing from the North direction towards the East direction. Therefore, by using this angle, one can project the two axes of the synthesized beam in order to obtain the size of this beam in the direction of the right ascension and the declination. The different parameters and obtained values after this projection are gathered in Table 12 for both targets.

Parameters	Unit	HD 168112	HD 167971
semi-major axis (a)	arcsec	0.02596	0.05271
semi-minor axis (b)	arcsec	0.01931	0.02358
position angle (θ)	deg	20.84	70.85
beam_α	mas	27.28	57.53
beam_δ	mas	31.13	39.57

Table 12: Parameters and results obtained after the projection of the synthesized beam onto the right ascension ($beam_\alpha$) and declination ($beam_\delta$).

To complete equation 49, one can use a value of 5 for the SNR of HD 168112 and a value of 7.5 for the SNR of HD 167971. By combining the values of those SNR and the values obtained in Table 12, the astrometric error due to interferometry can be computed. The values of this error for both targets are gathered in Table 13.

	Unit	HD 168112	HD 167971
$\sigma_{interferometry,\alpha}$	mas	8.57	12.05
$\sigma_{interferometry,\delta}$	mas	9.78	8.29

Table 13: Error on the astrometry due to interferometry on the right ascension and declination of both targets used in this work.

Secondly, there is a systematic uncertainty about the position of the phase calibrator within the International Celestial Reference Frame that propagates during the calibration steps. Some database such as the *Astrogeo database*[10] are providing the reference position of such sources as well as the error on this position. In the case of J1825-0737, the systematic error amounts to 0.23 mas. The error on the astrometric position of the sources induced by this systematic error on the position of the calibrator comes from the fact that, in interferometry and especially with VLBI, relative positions are considered. Indeed, during the calibration steps, the phase calibrator is assumed to be fixed in the sky at the position given in the Astrogeo database. Then, based on this position, the calibrations are extrapolated to the targets. Therefore, the relative positions between the position of the calibrator and the positions of the targets are considered. As a result, if there is an initial error on the position of the calibrator (which will be used as the reference position), it will give rise to an error on the position of the targets. This systematic uncertainty therefore translates to an uncertainty of $\sigma_{systematic} = 0.23$ mas on both the right ascension and declination for the two targets.

[10] http://astrogeo.org/vlbi-images/

Finally, the phase referencing technique which is used to observe some trends inside the targets' data despite the low SNR is creating a third uncertainty. During the phase referencing technique, an alternation of observation of the phase calibrator and of the targets is performed in order to extrapolate the trends visible in the calibrator's data as represented in Figure 14. During those observations, the phase calibrator is assumed to be so close to the targets that the atmospheric response will be the same during the observations of the calibrator and the observations of the targets. However, this hypothesis is not totally valid. Because the positions of the calibrator and the targets are different, one may expect the Earth's atmosphere to behave differently when observing one source or the other, even if all the sources are relatively close to each other. Using the notations of Pradel et al. (2006), the astrometric error induced by the phase referencing technique used on VLBI data is, in μas, given by

$$\Delta_{\alpha cos\delta,\delta} = \left(\Delta^{1^\circ}_{\alpha cos\delta,\delta} - 14\right) \times d + 14 \tag{50}$$

in which $\Delta_{\alpha cos\delta,\delta}$ is the astrometric error induced by phase referencing in μas, $\Delta^{1^\circ}_{\alpha cos\delta,\delta}$ is the astrometric error for a 1° source separation and d is the source separation in degrees that can be obtained thanks to

$$d = \sqrt{((\alpha - \alpha_0)cos\delta_0)^2 + (\delta - \delta_0)^2} \tag{51}$$

where α_0 and δ_0 are the right ascension and declination of the reference source, respectively and where α and δ are the right ascension and the declination of the target, respectively. One can therefore conclude that the larger the separation between the phase calibrator and the targets, the larger the uncertainty on the astrometric position that is derived for the targets. If one wants to compute the value of $\Delta_{\alpha cos\delta,\delta}$ in the frame of this work, given the values of $\Delta^{1^\circ}_{\alpha cos\delta,\delta}$ listed in Pradel et al. (2006) depending on the declination of the source of interest, one obtains that the errors on the astrometric position induced by the phase referencing technique are the values listed in Table 14.

	Unit	HD 168112	HD 167971
$\Delta_{\alpha cos\delta}$	mas	0.213	0.219
Δ_δ	mas	0.259	0.266

Table 14: Error induced by the phase referencing technique on the right ascension ($\Delta_{\alpha cos\delta}$) and declination ($\Delta_\delta$) of HD 168112 and HD 167971 using the values of $\Delta^{1^\circ}_{\alpha cos\delta,\delta}$ and the methodology listed in Pradel et al. (2006).

Some comments are needed concerning the values displayed in Table 14. Firstly, in equation 50, the parameter $\Delta^{1^\circ}_{\alpha cos\delta,\delta}$ has some pre-defined value that are listed in Pradel et al. (2006). Those different values of $\Delta^{1^\circ}_{\alpha cos\delta,\delta}$ depend on the declination of the source of interest and have been computed for a few declinations. However, $\Delta^{1^\circ}_{\alpha cos\delta,\delta}$ has not been computed for the exact values of the declination of either HD 168112 or HD 167971. Therefore, the purpose of the values listed inside Table 14 is more to obtain an order of magnitude on the possible uncertainty on the astrometric position rather than to obtain the precise values. Secondly, those errors are smaller than 300 μas, which is the upper limit defined by Pradel et al. (2006) for low or high declinations. Those results, even if they are approximated, are consistent with the expected results when using the methodology of Pradel et al. (2006).

All three different errors on the astrometric positions mentioned above can be added in quadrature as

$$\sigma_{tot} = \sqrt{\sigma^2_{interferometry} + \sigma^2_{systematic} + \Delta^2_{\alpha cos\delta,\delta}} \tag{52}$$

If one computes this overall uncertainty, one obtains the values gathered in Table 15. Therefore, it will be of major importance to consider those uncertainties in a future project that will have for objective to constrain the astrometry of the interaction region of HD 168112 and HD 167971.

	Unit	HD 168112	HD 167971
$\sigma_{tot,\alpha}$	mas	8.577	12.053
$\sigma_{tot,\delta}$	mas	9.786	8.294

Table 15: Total estimated astrometric uncertainty on both the right ascension ($\sigma_{tot,\alpha}$) and the declination ($\sigma_{tot,\delta}$) for both targets used in this work.

One can see that those astrometric errors are quite significant. By looking at the Figure 39, one can see that the astrometric error is nearly covering the entire distance in between the two central stars (grey dot) and the outer star (purple dot). Therefore, it is not possible to constrain the position of the interaction region because of the significant astrometric errors. However, by looking carefully to the values of the different errors, one can see that this is the error defined by equation 49 which dominates clearly inside the quadrature addition. While a value of $\lambda = 18$cm has been chosen in the frame of the pilot project in order to ensure the detection of non-thermal emission and the non-detection of thermal emission, an observation performed at e.g. $\lambda = 6$cm would already decrease by a factor 3 this astrometric error.

Indeed, the objective of the pilot project being to demonstrate the capacity of the EVN to detect such non-thermal emission, a wavelength of 18cm is convenient in two different ways. Firstly, as the flux density of an optically thin synchrotron emission is decreasing with an increasing frequency, it is more likely to detect such non-thermal emission at a low frequency, i.e. a high wavelength (cfr. Figure 3). Secondly, as this is the opposite for the flux density of an optically thin free-free radiation, it is more likely that the thermal emission will be negligible at low frequencies (cfr. Figure 2). However, in the frame of a project that has for objective to characterize the astrometry of such systems, the emphasis can be put on the reduction of the astrometric error by observing with a shorter wavelength, i.e. a higher frequency, in order to significantly decrease the error defined by equation 49.

In addition, as it can be seen in Figure 32, the phase calibrator does not appear in the center of the image. This problem, which is likely due to some ionospheric effects, has to be solved if one wants to better constrain the astrometry of the two systems analyzed in this work as the position of the phase calibrator will be used as a reference position in order to determine the positions of the targets. Therefore, if such a shift is observed for the position of the calibrator, the positions of the targets must be corrected for this shift.

8.2 Observations at other frequencies

The second possible prospect is concerning new observations of the targets analyzed in this work. The goal of new observations would be to be able to compare the results with previous ones or with the results obtained thanks to this work. As seen in the case of HD 167971, the radio flux variability is a strong hint of the presence of non-thermal emission and the orbital phase is modulating the detected radio flux density. By observing HD 168112 and HD 167971 at different orbital phases and different frequencies than what has been done in this work or previously, it will allow to obtain or complete the radio lightcurve associated to the systems. This will be useful in order to better constrain the variation of the radio flux density.

In addition, observing those systems at a different orbital phase and at a different frequency will allow to verify the possible variation of the shape of the emission region. This verification will help to better constrain the elongation of the interaction region. Finally, by observing at a different frequency than what has been done for this work, it will be possible to derive a spectral index. Indeed, at least two observations at different frequencies are needed to derive such spectral index as it corresponds to the slope of the synchrotron spectrum. No spectral index could therefore have been derived from the results of this work. Obtaining the spectral index of the non-thermal emission will allow to obtain information on the distribution of the relativistic electrons resulting from the particle acceleration process.

8.3 New detection

A last possible prospect is concerning the observation of other systems. Indeed, approximately 50 objects are now known to be particle accelerators and are listed inside the Catalogue of Particle-Accelerating Colliding-Wind Binaries. However, among those 50 systems, only a few of them have already been imaged as what has been done for HD 168112 and HD 167971 in this work. Observing systems listed inside the catalogue will allow to increase the number of PACWBs for which a radio image has been obtained. Hence, VLBI arrays such as the Australian Long Baseline Array (LBA) could allow to observe such systems located in the Southern hemisphere, as most of the systems located in the Northern hemisphere have already been studied. Obtaining such images will help to constrain several physical parameters such as the non-thermal flux density, the elongation of the interaction region or even the astrometric position of the emission region.

Finally, as the feasibility of the detection of such non-thermal emission has been demonstrated by this work as well as by several previous studies, it would be interesting to use VLBI arrays in order to try to detect non-thermal emission in systems which are not specifically known to be particle accelerators, both inside the Northern and Southern hemisphere. Indeed, as the detection does not require the observation of the entire orbit but rather a snapshot at a given orbital phase, it would be possible to observe such long-period systems and perhaps to complete the list of known PACWBs at an affordable observation time cost.

Annexe 1

The content of this annexe is highly based on a collection of lectures (Perley et al., 1989) and on Thompson et al. (2017). One can consider the radio source to be located in $\vec{R}$ and the observer to be located in $\vec{r}$. Because the time interval of observation is finite, it is interesting to first consider the development in Fourier series of the emitted electric field.

$$\vec{E}(\vec{R},t) = \sum_{k=-\infty}^{\infty} \vec{E}_\nu(\vec{R}) \exp(2i\pi kt/\tau) \tag{53}$$

where the components $\vec{E}_\nu(\vec{R})$ of this series are not functions of time anymore. It is then more convenient to work with those components instead of working directly with the time-varying electric field. It is important to note that those coefficients will be complex. The received field at the location $\vec{r}$ can then be written as

$$\vec{E}_\nu(\vec{r}) = \iiint P_\nu(\vec{R},\vec{r})\vec{E}_\nu(\vec{R})\, dx\, dy\, dz \tag{54}$$

where $\mathrm{P}_\nu(\vec{R},\vec{r})$ is called the *propagator* and is accounting for the modifications of the emitted electric field along the travel from the source to the observer.

This is now convenient to express the *spatial coherence function*. Considering that radio interferometry will be performed, the electric field will be measured at two different spatial positions that are noted $\vec{r}_1$ and $\vec{r}_2$. The spatial coherence function is defined as the product

$$V_\nu(\vec{r}_1,\vec{r}_2) =< \vec{E}_\nu(\vec{r}_1), \vec{E}^*_\nu(\vec{r}_2) > \tag{55}$$

where the * symbol is referring to the complex conjugate. One can then consider several assumptions. The first one relies on the fact that as the source might be very far away from the Earth, it is impossible to describe the three-dimensional structure of the emitting region (i.e. the depth cannot be observed). Therefore, one can consider a large sphere of radius $\vec{R'}$ around the observer called the *celestial sphere* inside which there is no other radiation than the one emitted by the source of interest. In order to describe the field emitted by the source, one can focus on the distribution of the the electric field onto the surface of the celestial sphere. This particular distribution will be noted $\xi_\nu(\vec{R'})$ and will be useful to introduce the relation between the spatial coherence function and the intensity of the radiation. The remaining assumption that needs to be taken into account is that, one can consider that the interior of the celestial sphere defined above is empty. Because of the fact that the celestial sphere is considered to have a large radius, the surface of this sphere and therefore the evaluation of the electric field on its surface can be considered to be in the far-field. In addition, because the empty space is considered inside the sphere, the *Huygens-Fresnel principle* can be used.

This principle states that every point on a wave front will act as an intermediate source of spherical waves. In addition, all those waves may interfere with each other. The resulting wave front will be the sum of all the secondary waves created by the intermediate sources. Using this principle allows to express the propagator of equation 54. One can also use the expression of $\xi_\nu(\vec{R'})$ so that equation 54 can be written as

$$E_\nu(\vec{r}) = \int \xi_\nu(\vec{R'}) \frac{e^{2\pi i\nu|\vec{R}-\vec{r}|/c}}{|\vec{R}-\vec{r}|}\, dS \tag{56}$$

where the electric field is now considered as a scalar value evaluated at the point of interest $\vec{r}$ (i.e. no polarisation is considered) while dS is accounting for an element of surface of the celestial sphere. One can now use this expression of the electric field at the position of interest to re-express the spatial coherence function given above. It yields

$$V_\nu(\vec{r}_1,\vec{r}_2) = \left\langle \iint \xi_\nu(\vec{R}_1)\xi^*_\nu(\vec{R}_2) \frac{e^{2\pi i\nu|\vec{R}_1-\vec{r}_1|/c}}{|\vec{R}_1-\vec{r}_1|} \frac{e^{-2\pi i\nu|\vec{R}_2-\vec{r}_2|/c}}{|\vec{R}_2-\vec{r}_2|}\, dS_1\, dS_2 \right\rangle \tag{57}$$

One can assume that, as an astronomical object is considered, the radiation is not spatially coherent for $\vec{R}_1 \neq \vec{R}_2$ which yields

$$< \xi_\nu(\vec{R}_1)\xi^*_\nu(\vec{R}_2) >= 0 \tag{58}$$

Therefore, the spatial coherence function can be given by

$$V_\nu(\vec{r}_1,\vec{r}_2) = \int < |\xi_\nu(\vec{R'})|^2 > |\vec{R'}|^2 \frac{e^{2\pi i\nu|\vec{R'}-\vec{r}_1|/c}}{|\vec{R'}-\vec{r}_1|} \frac{e^{-2\pi i\nu|\vec{R'}-\vec{r}_2|/c}}{|\vec{R'}-\vec{r}_2|}\, dS \tag{59}$$

Writing the vector $\vec{R}/|\vec{R}|$ as $\vec{s}$ allows to define the *intensity* of the radiation as

$$I_\nu(\vec{s}) =< |\xi_\nu(\vec{R'})|^2 > |\vec{R'}|^2 \tag{60}$$

and because the value $|\vec{R'}|$ is large, the terms in $|\vec{r}/\vec{R'}|$ are small and can be neglected. Using the definition of a solid angle, one can replace the element of surface dS by $|\vec{R}|^2 \mathrm{d}\Omega$ and, with some geometrical considerations, the spatial coherence function is now given by

$$V_\nu(\vec{r}_1,\vec{r}_2) \simeq \int I_\nu(\vec{s}) e^{-2\pi i\nu\vec{s}.(\vec{r}_1-\vec{r}_2)/c}\, d\Omega \tag{61}$$

where the spatial coherence function is finally related to the intensity of the radiation that needs to be obtained. It is important to note that this relation is in fact a Fourier transform that will be inverted in order to have access to the intensity of the radiation.

References

Atri Pikky. Constraining black hole natal kicks with high-precision radio astrometry. 2020.

Beasley AJ, Conway JE. VLBI phase-referencing // Very Long Baseline Interferometry and the VLBA. 82. 1995. 327.

Bieging JH, Abbott David C, Churchwell EB. A survey of radio emission from Galactic OB stars // The Astrophysical Journal. 1989. 340. 518–536.

Blomme Ronny, De Becker Michaël, Runacres MC, Van Loo S, Gunawan DYA Setia. Non-thermal radio emission from O-type stars-II. HD 167971 // Astronomy & Astrophysics. 2007. 464, 2. 701–708.

Blomme Ronny, Van Loo Sven, De Becker Michaël, Rauw Grégor, Runacres MC, Gunawan DYA Setia, Chapman JM. Non-thermal radio emission from O-type stars-I. HD 168112 // Astronomy & Astrophysics. 2005. 436, 3. 1033–1040.

Briggs DS. High fidelity interferometric imaging: Robust weighting and NNLS deconvolution // American Astronomical Society Meeting Abstracts. 187. 1995. 112–02.

Burns Ross Alexander. The structure and kinematics of the Milky Way Galaxy and the formation of its massive stars as seen with maser VLBI // Ph. D. Thesis. 2016.

Condon James J, Ransom Scott M. Essential radio astronomy. 2016.

Cornwell Tim, Fomalont Ed B. Self-calibration // Synthesis imaging in radio astronomy II. 180. 1999. 187.

De Becker M. Determination of physical parameters of particle-accelerating colliding-wind binaries based on generalized partition considerations // A&A. XII 2018. 620. A144.

De Becker M., Sana H., Absil O., Le Bouquin J. B., Blomme R. The long-period eccentric orbit of the particle accelerator HD 167971 revealed by long baseline interferometry // MNRAS. VII 2012. 423, 3. 2711–2717.

De Becker Michaël. Long-term XMM–Newton investigation of two particle-accelerating colliding-wind binaries in NGC 6604: HD 168112 and HD 167971 // Monthly Notices of the Royal Astronomical Society. 2015. 451, 1. 1070–1080.

De Becker Michaël. SPAT0069 : Radio Astrophysics. 2021.

De Becker Michaël, Benaglia Paula, Romero Gustavo Esteban, Peri Cintia Soledad. An investigation into the fraction of particle accelerators among colliding-wind binaries-Towards an extension of the catalogue // Astronomy & Astrophysics. 2017. 600. A47.

De Becker Michaël, Raucq Françoise. Catalogue of particle-accelerating colliding-wind binaries // Astronomy & Astrophysics. 2013. 558. A28.

De Becker Michaël, Rauw Grégor, Blomme Ronny, Waldron WL, Sana Hugues, Pittard JM, Eenens Philippe, Stevens Ian Robert, Runacres MC, Van Loo Sven, others . Quasi-simultaneous XMM-Newton and VLA observation of the non-thermal radio emitter HD 168112 (O5. 5III (f)) // Astronomy & Astrophysics. 2004. 420, 3. 1061–1077.

De Becker Michaël, Van Grootel Valérie. SPAT0008 : Interstellar Medium. 2021.

Deller A. T., Middelberg E. mJIVE-20: A Survey for Compact mJy Radio Objects with the Very Long Baseline Array // AJ. I 2014. 147, 1. 14.

Drew JE. On star wind temperatures, SI IV 1397-A absorption and HI H-alpha emission // Properties of Hot Luminous Stars. 7. 1990. 230–241.

Drury L O'C. An introduction to the theory of diffusive shock acceleration of energetic particles in tenuous plasmas // Reports on Progress in Physics. 1983. 46, 8. 973.

Garrett M. VLBI imaging : principles and practice. presentation // NATO VLBI school. 2001.

Hessels Jason. The Techniques of Radio Interferometry I: Basics. April 2013.

Högbom JA. Aperture synthesis with a non-regular distribution of interferometer baselines // Astronomy and Astrophysics Supplement Series. 1974. 15. 417.

Isella Andrea. Interferometry basics // Science. nrao. edu. Proc. of Caltech CASA Radio Analysis Workshop, CA, Pasadena. National Radio Astronomy Observatory. 19. 2011.

Jamet Luc, Pérez Enrique, Cervino Miguel, Stasińska G, Delgado RM González, Vílchez JM. On the importance of the few most massive stars: the ionizing cluster of NGC 588 // Astronomy & Astrophysics. 2004. 426, 2. 399–413.

Keimpema A, Kettenis MM, Pogrebenko SV, Campbell RM, Cimó G, Duev DA, Eldering B, Kruithof N, Langevelde HJ van, Marchal D, others . The SFXC software correlator for very long baseline interferometry: algorithms and implementation // Experimental Astronomy. 2015. 39, 2. 259–279.

Le Bouquin J. B., Sana H., Gosset E., De Becker M., Duvert G., Absil O., Anthonioz F., Berger J. P., Ertel S., Grellmann R., Guieu S., Kervella P., Rabus M., Willson M. Resolved astrometric orbits of ten O-type binaries // A&A. V 2017. 601. A34.

Leitherer Claus, Chapman Jessica M, Koribalski Baerbel. Radio continuum measurements of southern early-type stars // The Astrophysical Journal. 1995. 450. 289.

Leitherer Claus, Robert Carmelle. Observations of stellar winds from hot stars at 1.3 millimeters // The Astrophysical Journal. 1991. 377. 629–638.

Longair Malcolm S. High energy astrophysics. 2010.

Maíz Apellániz J., Sota A., Arias J. I., Barbá R. H., Walborn N. R., Simón-Díaz S., Negueruela I., Marco A., Leão J. R. S., Herrero A., Gamen R. C., Alfaro E. J. The Galactic O-Star Spectroscopic Survey (GOSSS). III. 142 Additional O-type Systems. // ApJS. V 2016. 224, 1. 4.

Maíz Apellániz J., Trigueros Páez E., Negueruela I., Barbá R. H., Simón-Díaz S., Lorenzo J., Sota A., Gamen R. C., Fariña C., Salas J., Caballero J. A., Morrell N. I., Pellerin A., Alfaro E. J., Herrero A., Arias J. I., Marco A. MONOS: Multiplicity Of Northern O-type Spectroscopic systems. I. Project description and spectral classifications and visual multiplicity of previously known objects // A&A. VI 2019. 626. A20.

Marcote B. Non-thermal emission from high-energy binaries through interferometric radio observations // arXiv preprint arXiv:1603.09276. 2016.

Massey Philip. Massive stars in the galaxies of the Local Group // New Astronomy Reviews. 2013. 57, 1-2. 14–27.

Muijres LE, Vink Jorick S, Koter Alex de, Müller PE, Langer N. Predictions for mass-loss rates and terminal wind velocities of massive O-type stars // Astronomy & Astrophysics. 2012. 537. A37.

Panagia Nino, Felli Marcello. The spectrum of the free-free radiation from extended envelopes // Astronomy and Astrophysics. 1975. 39. 1–5.

Pannuti Thomas G. The Physical Processes and Observing Techniques of Radio Astronomy: An Introduction. 2020.

Perley Richard A, Schwab Frederic R, Bridle Alan H. Synthesis imaging in radio astronomy: a collection of lectures from the third NRAO synthesis imaging summer school. 6. 1989.

Petrachenko B. Radio antennas, feed horns, and front-end receivers. presentation // EGU and IVS Training School on VLBI for Geodesy and Astrometry. 2013.

Phillips RB, Titus MA. Synchrotron radiation from the winds of O supergiants-TB= 10 to the 7.6 th K at 60 stellar radii // The Astrophysical Journal. 1990. 359. L15–L18.

Pittard JM, Dougherty SM. Radio, X-ray, and γ-ray emission models of the colliding-wind binary WR 140 // Monthly Notices of the Royal Astronomical Society. 2006. 372, 2. 801–826.

Pradel N, Charlot P, Lestrade J-F. Astrometric accuracy of phase-referenced observations with the VLBA and EVN // Astronomy & Astrophysics. 2006. 452, 3. 1099–1106.

Reipurth Bo. The Young Cluster NGC 6604 and the Serpens OB2 Association // Handbook of Star Forming Regions, Volume II. 2008. 5. 590.

Runacres MC, Owocki SP. The outer evolution of instability-generated structure in radiatively driven stellar winds // Astronomy & Astrophysics. 2002. 381, 3. 1015–1025.

Sana Hugues, Le Bouquin J-B, Lacour S, Berger J-P, Duvert G, Gauchet L, Norris B, Olofsson J, Pickel D, Zins G, others . Southern massive stars at high angular resolution: observational campaign and companion detection // The Astrophysical Journal Supplement Series. 2014. 215, 1. 15.

Sana Hugues, Rauw Grégor, Gosset Eric. The young open cluster NGC 6231: five years of investigations // Massive Stars and High-Energy Emission in OB Associations. 2005. 107–110.

Sanchez-Bermudez J., Alberdi A., Schödel R., Brandner W., Galván-Madrid R., Guirado J. C., Herrero-Illana R., Hummel C. A., Marcaide J. M., Pérez-Torres M. A. A VLBI study of the wind-wind collision region in the massive multiple HD 167971 // A&A. IV 2019. 624. A55.

Shepherd MC, Pearson TJ, Taylor GB. DIFMAP: an interactive program for synthesis imaging. // Bulletin of the American Astronomical Society. 26. 1994. 987–989.

Stevens IR. X-ray and gamma-ray emission from single and binary early-type stars // Massive Stars and High-Energy Emission in OB Associations. 2005. 3–12.

Stevens Ian R, Blondin John M, Pollock AMT. Colliding winds from early-type stars in binary systems // The Astrophysical Journal. 1992. 386. 265–287.

Thompson A Richard, Moran James M, Swenson George W. Interferometry and synthesis in radio astronomy. 2017.

Thompson AR, Moran JM, Swenson GW. Book Review: Interferometry and synthesis in radio astronomy./Wiley-Interscience, 1986 // Space Science Reviews. 1988. 46. 380.

Wright Alan E, Barlow Michael J. The radio and infrared spectrum of early-type stars undergoing mass loss // Monthly Notices of the Royal Astronomical Society. 1975. 170, 1. 41–51.

Zinnecker Hans, Yorke Harold W. Toward understanding massive star formation // Annu. Rev. Astron. Astrophys. 2007. 45. 481–563.