# Model for a Distributed Radio Telescope

Patrick Fleckenstein

Department of Mathematics, Rochester Institute of Technology, Rochester, NY 14623, U.S.A.

*E-mail address*: `pat@csh.rit.edu`

Abstract. In this paper, a stochastic model for a single radio telescope is generated. The model takes into account interference from other radio sources (including the sky), errors in the measurement of the detector's position, errors in the measurement of time at the detector station, noise due to the quantum nature of light, and noise in the detector. This model is then used as the basis for a computer simulation of a distributed array of radio telescopes. The computer simulation successfully resolves a target signal using a few hundred very bad detectors. The simulation is further employed to analyze the model's sensitivity to various parameters such as the number of detectors and the magnitude of errors in time measurements.

# Contents

CHAPTER 1

# Summary and conclusions

This paper develops a stochastic model of the behaviour of a single radio telescope subject to incoming signals from various radio sources (including the sky). In the model, the time is expressed in terms of the time when the signal one is attempting to measure would pass the center of the Earth. The resulting stochastic model expresses the probability that the detector will detect $z$ photons at a particular time $t$:

$$p_d(z,t) = \sum_{x=0}^{\infty} \frac{p_{i\sum \mathbf{x}}(x,t) \cdot \int_{y=x-\frac{1}{2}}^{x+\frac{1}{2}} e^{-\frac{1}{2}\left(\frac{z-y}{\sigma_d}\right)^2} \cdot dy}{\sigma_d \sqrt{2\pi}} \tag{1}$$

This expresses the probability that $z$ photons are detected by the $i$-th detector at the time when the $i$-th detector should be receiving the signal from the primary radio source (the source being observed) which will pass the center of the Earth at time $t$. Here $\sigma_d$ is the standard deviation of the noise in the detector. The integral in this equation arises from the tenuous assumption that noise in the detector can be adequately modeled with a normal distribution.

In the above equation, the term $p_{i\sum \mathbf{x}}(x,t)$ gives the probability that $x$ photons from the various radio sources will reach the detector at time $t$. That probability is:

$$\begin{aligned} p_{i\sum \mathbf{x}}(x,t) = \int_{\epsilon_T=-\infty}^{\infty} & \frac{\left(\sum_j s_j\left(t+\frac{h_{ij}}{c}-\frac{h_{i1}}{c}-\epsilon_T\right)\right)^x e^{-\sum_j s_j\left(t+\frac{h_{ij}}{c}-\frac{h_{i1}}{c}-\epsilon_T\right)}}{x!} \\ & \cdot \frac{e^{-\frac{1}{2}\left(\frac{\epsilon_T}{\sqrt{(\sigma_h/c)^2+\sigma_t^2}}\right)^2}}{\sqrt{((\sigma_h/c)^2+\sigma_t^2)\cdot 2\pi}} \cdot d\epsilon_T \end{aligned} \tag{2}$$

Here, $s_j(t)$ is the strength of the $j$-th signal which would pass the center of the Earth during the sampling interval centered at time $t$. The $h_{ij}$ terms are the height of the $i$-th receiver measured parallel to a ray which starts at the center of the Earth and extends toward the $j$-th radio source. The $\epsilon_T$ term is the total error in the station's time measurements due to errors in its positional measurement $h_{ij}$ and errors in its direct time measurement. The standard deviation of the errors in positional measurements is $\sigma_h$, and the standard deviation of the errors in direct time measurements is $\sigma_t$. The speed of light is denoted $c$.

Because equations 1 and 2 are quite unwieldy, a computer simulation was developed based upon this single detector model to determine the performance of an array of these telescopes. The computer simulation shows that even faint signals can be detected with very poor detectors when the accuracy of positional and time measurements is within currently attainable limits.

Chapter 2 gives a summary of the variables used in the development of this model. Chapter 3 provides background information about radio telescope arrays and forms the base upon which this model is cast.

In chapter 4, the goals for this model are established. It goes on to highlight the problems in time and position measurements associated with distributing the telescopes in this array, and it mentions the problems of detecting signals which are incurred at each detector.

Chapter 5 goes on from there to derive equations 1 and 2. The analysis in this chapter depends heavily upon some proofs about the sums of normally distributed variables and the sums of Poisson distributed variables. Those proofs are in appendix A and appendix B.

Chapter 6 employs a computer simulation based upon the equations derived in chapter 5. It shows empirical results of running the simulation with various sources of radio interference, various magnitudes of time errors, and various numbers of telescopes in the array.

Chapter 7 goes on to suggest places for further exploration before one attempts to create this type of telescope array.

CHAPTER 2

# Glossary of variables

This table summarizes the variables used in the development of the distributed radio-telescope model. Throughout this paper, variables which are randomly distributed will be written in bold.[1]

| **Description** | **Variable** | **Units** |
|---|---|---|
| the strength of the signal from the $j$-th source that will pass the center of the Earth at time $t$ | $s_j(t)$ | $photons/s$ |
| the height of $i$-th observation point in the direction toward the $j$-th source with respect to the Earth's center | $h_{ij}$ | $m$ |
| the error in height measurement | $\epsilon_{\mathbf{h}}$ | $m$ |
| the standard deviation of height measurements | $\sigma_h$ | $m$ |
| the error in time measurement | $\epsilon_{\mathbf{t}}$ | $s$ |
| the standard deviation of time measurements | $\sigma_t$ | $s$ |
| the total error in time measurement caused by $\epsilon_{\mathbf{h}}$ and $\epsilon_{\mathbf{t}}$ | $\epsilon_{\mathbf{T}}$ | $s$ |
| the error in the detector | $\epsilon_{\mathbf{d}}$ | $photons/s$ |
| the standard deviation of the error in the detector | $\sigma_d$ | $photons/s$ |
| the speed of light | $c$ | $m/s$ |
| the radius of the Earth | $R$ | $m$ |

[1]Unfortunately, a bold epsilon $\epsilon$ is indistinguishable from a normal epsilon $\epsilon$ in LaTeX. However, the subscripts on them will still be bold.

CHAPTER 3

# Background information

Many of the best radio observatories use multiple radio telescopes in tandem for an observation. The outputs of the individual telescopes are combined to form a clearer image of the target object than any one of the telescopes could provide. There is a tradeoff between the number of radio telescopes required and the sensitivity of each telescope. The Very Large Array[1] uses 27 $25m$-telescopes, the Very Long Baseline Array[2] uses 10 $25m$-telescopes, and the proposed Atacama Large Millimeter Array[3] will use 64 $12m$-telescopes. If a much larger array were possible with much smaller telescopes (ones that individuals could afford), one would hope to be able to harness some of the 400 members of the Society of Amateur Radio Astronomers[4] and some of the two million users who have donated computer time to the SETI@Home[5] project. With wide enough support, one may be able to create an array large enough to make useful observations.

Radio telescope arrays link several telescopes together. Signals from the radio source hit the different telescopes at slightly different times because the signals have to travel different distances to each of the telescopes. The signals from these telescopes are then recombined. The position of each telescope with respect to the radio source is taken into account in order to synchronize the signals.

The recombined signals reinforce each other. Noise in the telescopes and from other stellar sources have the same probability of being in-phase as out of phase while the synchronized signals will be in-phase. By carefully placing multiple telescopes, scientists can select particular frequencies from particular sources despite large amounts of background interference.

The Very Long Baseline Array takes a different approach than most other arrays. The 27 telescopes of the Very Large Array, for example, are all on the same plot of land in New Mexico. The 10 telescopes of the Very Long Baseline Array are not in close proximity. One telescope is in St. Croix, another in Hawaii, others in California, etc. Rather than precisely positioning the telescopes to some precalculated formation, the scientist adjusts for the actual positions of the radio telescopes to synchronize their data streams. This compensation can be accomplished so long as the location of each telescope is precisely known and the data stream for each telescope is precisely timestamped.

Global Positioning Satellite (GPS)[6] receivers are becoming commonplace. Amateurs can precisely know their latitude, longitude, and elevation. The Network

[1] `http://www.aoc.nrao.edu/vla/html/VLAhome.shtml`
[2] `http://www.aoc.nrao.edu/vlba/html/VLBA.html`
[3] `http://www.alma.nrao.edu/`
[4] `http://www.bambi.net/sara.html`
[5] `http://setiathome.ssl.berkeley.edu/`
[6] `http://www.aero.org/publications/GPSPRIMER/index.html`

Time Protocol[7] is nearing nanosecond accuracy on new computers. Computer users will be able to accurately measure time intervals and synchronize time with standard clocks. The technology to synchronize data streams from amateur telescopes will soon be widespread.

The only missing technology is cheap radio telescopes. This, however, is largely due to demand. There has never been a use for cheap radio telescopes. If a large, distributed array of cheap radio telescopes were possible, cheap radio telescopes could readily be constructed.

[7] http://www.eecis.udel.edu/~mills/precise.htm

CHAPTER 4

# Formulation of the model

Radio signals from pulsars and other sources are constantly hitting the Earth. Figure 1 shows a signal from one pulsar approaching several detectors on Earth. The signal will hit the detectors at different times because the detectors are different distances away from the source (the pulsar).

For the purposes of this model, it is assumed that the Earth and the sources of radio signals are all motionless. The length of time for any single point of observation will be under 1/100-th of a second for this array. This model assumes that any motion of the source or the Earth during that 1/100-th of a second will be negligible and that the coordinate transformations to adjust for position over a series of observation points can easily be added later. The coordinate transformations

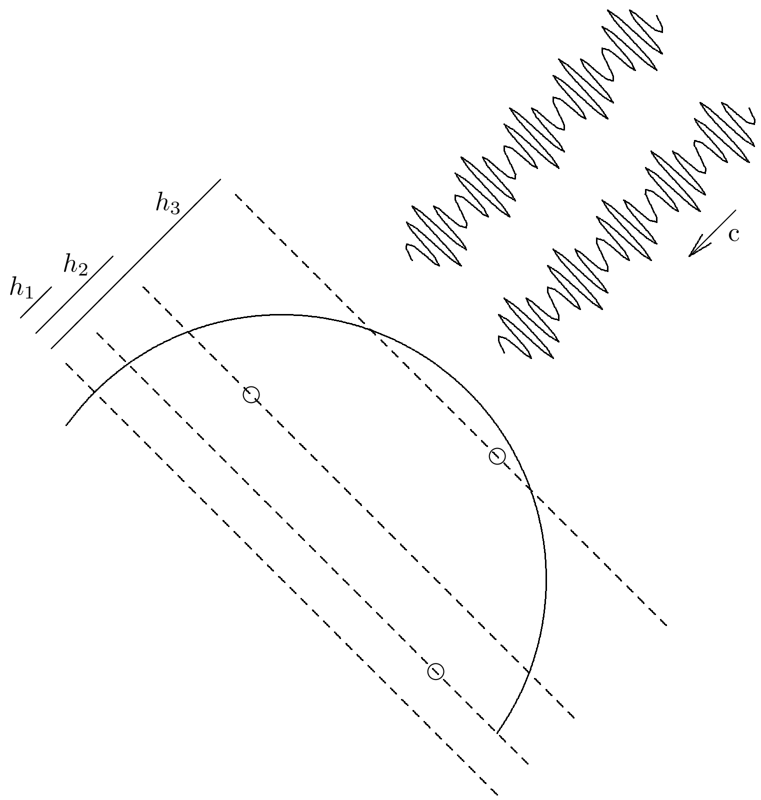


FIGURE 1. Signal hitting sensors on the Earth

would just obfuscate the main line of this development if they were to be included from the outset.

## 1. Measuring time

The relative distances of the detectors from the sources are critical factors in this model. This model uses $h_{ij}$ to represent the height of the $i$-th detector with respect to the $j$-th source. The quantity $h_{ij}$ is the height of the detector measured parallel to a ray originating at the center of the Earth and extending toward the $j$-th radio source. For example, if the $j$-th radio source were directly over the north pole, then $h_{ij}$ would be

$$h_{ij} = R \sin \theta_i$$

where $R$ is the radius of the Earth and $\theta_i$ is the degrees latitude north of the equator for the $i$-th detector. An implicit assumption here is that each radio source is far enough away that one can assume that the incoming radio waves are parallel without incurring significant error.

The center of the Earth is the point of reference for all measurements in this model. In order to determine the strength of the signal from the $j$-th source that would have hit the center of the Earth at time $t$, one must compensate for the fact that the signal hit the $i$-th detector sooner because the $i$-th detector was $h_{ij}$ meters closer to the $j$-th source than the center of the Earth is. Because light travels at $c$ meters per second, one must check the output of the $i$-th detector from time $t_{ij}$ where

$$t_{ij} = t - \frac{h_{ij}}{c}$$

However, because there is some error in the measurement of position of the detector and in the measurement of time at the detector station, $\mathbf{t_{ij}}$ is randomly distributed.

$$\mathbf{t_{ij}} = t - \left( \frac{h_{ij}}{c} + \frac{\epsilon_{\mathbf{h}}}{c} + \epsilon_{\mathbf{t}} \right) \tag{3}$$

where $\epsilon_{\mathbf{h}}$ is the error in the measurement of position and $\epsilon_{\mathbf{t}}$ is the error in the measurement of time.

This model assumes that the error in the measurement of position $\epsilon_{\mathbf{h}}$ and the error in the measurement of time $\epsilon_{\mathbf{t}}$ are each normally distributed with zero mean. The standard deviation of the error in position measurement is denoted $\sigma_h$, and the standard deviation of the error in time measurement is denoted $\sigma_t$. In practice, these errors may not be normally distributed. But, hopefully the number of telescopes involved will be large enough to make it reasonable to use normal distributions as approximations of the true distributions. Also, these errors may not have zero means. But, because the model only depends on the relative positions and times of the detectors, a bias in these distributions would not affect the results. The key here is to synchronize the signals from the different detectors not to determine the exact time the signal would pass the center of the Earth.

## 2. Measuring signals

The actual signal that reaches a detector has noise in it already. Radio waves are not emitted at a constant rate from a source. Due to the quantum nature of light, a "constant" source has a constant probability of emitting photons as opposed to a constant emission of photons. Thus, the number of photons actually released by a source follows a Poisson distribution whose parameter is the ideal emission rate for the source.[1].

*Note: Radio waves are electromagnetic waves. They are of a lower frequency than visible light waves, but they are fundamentally the same thing. Radio waves are composed of photons just as visible light waves are.*

If the source would ideally emit $s_j(t)$ photons for time $t$, then the probability of $x$ photons from the source arriving at the $i$-th detector at time $\mathbf{t_{ij}}$ is given by:

$$p_{\mathbf{x}}(x) = \frac{(s_j(t))^x \, e^{-s_j(t)}}{x!} \tag{4}$$

Various sources contribute to noise in the detector—thermal noise, g-r noise and $1/f$ noise.[2] Rather than model each of these individually, this model assumes that the net effect will be additive noise that is normally distributed with zero mean and standard deviation $\sigma_d$. This is the most tenuous assumption thus far and must certainly be taken up in future work.

[1] [**Kitchin**] p. 41

[2] [**Kitchin**] p. 40

CHAPTER 5

# Analysis of the model

In this model, the terms $\frac{\epsilon_\mathbf{h}}{c}$ and $\epsilon_\mathbf{t}$ are combined into a total error in time due to the errors in positional measurement and time measurement. This total error in time is denoted $\epsilon_\mathbf{T}$. Thus, equation 3 from page 12 becomes:

$$\mathbf{t_{ij}} = t - \left(\frac{h_{ij}}{c} + \epsilon_\mathbf{T}\right) \tag{5}$$

Given that $\epsilon_\mathbf{h}$ and $\epsilon_\mathbf{t}$ are normally distributed with zero means and standard deviations of $\sigma_h$ and $\sigma_t$ respectively, one would like to know the distribution of the total error in time $\epsilon_\mathbf{T}$. The total error in time $\epsilon_\mathbf{T}$ is normally distributed with zero mean and a standard deviation of $\sqrt{(\sigma_h/c)^2 + \sigma_t^2}$. The full derivation of this is given in appendix A.

There will be more than one source whose radio signal will strike the detector. Because of this, one must modify equation 4 to take into account multiple sources. Taking $s_1(t)$ to be the signal one wishes to observe, one will synchronize the time signals of the $i$-th detector using $\mathbf{t_{i1}}$ from equation 5. In other words, the output of the $i$-th detector should be checked at time $t - \frac{h_{i1}}{c}$. One must take into account the strength of the other signals which would hit that detector at the same time. To do this, one must check the strength of the $j$-th signal at time $t + \frac{h_{ij}}{c} - \frac{h_{i1}}{c}$. One can readily verify that if one is checking the strength of the signal from the primary source, one would be checking the strength of the signal that would reach the center of the Earth at time $t$.

Given multiple variables which are each Poisson distributed, their sum is Poisson distributed. The full derivation of this is found in appendix B. Because of this equation 4 becomes:

$$\overline{p_{i\sum\mathbf{x}}}(x,t) = \frac{\left(\sum_j s_j\left(t + \frac{h_{ij}}{c} - \frac{h_{i1}}{c}\right)\right)^x e^{-\sum_j s_j\left(t + \frac{h_{ij}}{c} - \frac{h_{i1}}{c}\right)}}{x!} \tag{6}$$

for multiple radio signal sources.

In order to take into account the errors in time measurement, one must sum the probability from equation 6 times the probability of a particular error in time measurement over all possible values of the error in time measurement. That is:

$$p_{i\sum\mathbf{x}}(x,t) = \int_{\epsilon_T=-\infty}^{\infty} \overline{p_{i\sum\mathbf{x}}}(x, t-\epsilon_T)\cdot p_{\epsilon_\mathbf{T}}(\epsilon_T)\cdot d\epsilon_T \tag{7}$$

Because the error in time measurement $\epsilon_\mathbf{T}$ is normally distributed with zero mean and a standard deviation of $\sqrt{(\sigma_h/c)^2 + \sigma_t^2}$ (see page 15), the function $p_{\epsilon_\mathbf{T}}(\epsilon_T)$

is given by:

$$p_{\epsilon_{\mathbf{T}}}(\epsilon_T) = \frac{e^{-\frac{1}{2}\left(\frac{\epsilon_T}{\sqrt{(\sigma_h/c)^2+\sigma_t^2}}\right)^2}}{\sqrt{((\sigma_h/c)^2+\sigma_t^2)\cdot 2\pi}} \tag{8}$$

It simplifies things greatly if one assumes that that the total error in time measurement $\epsilon_T$ is the same for each source. In reality, the error in time due to the error in position measurement $\epsilon_{\mathbf{h}}$ will be different for sources which are in different directions. However, because the speed of light $c$ is very large compared to the standard deviation of the error in position measurement $\sigma_h$, the whole value $\epsilon_{\mathbf{h}}$ has a very minor effect on $\epsilon_T$ anyway. With this assumption, equation 7 can be written out completely:

$$\begin{aligned} p_{i\sum\mathbf{x}}(x,t) = \int_{\epsilon_T=-\infty}^{\infty} & \frac{\left(\sum_j s_j\left(t+\frac{h_{ij}}{c}-\frac{h_{i1}}{c}-\epsilon_T\right)\right)^x e^{-\sum_j s_j\left(t+\frac{h_{ij}}{c}-\frac{h_{i1}}{c}-\epsilon_T\right)}}{x!} \\ & \cdot \frac{e^{-\frac{1}{2}\left(\frac{\epsilon_T}{\sqrt{(\sigma_h/c)^2+\sigma_t^2}}\right)^2}}{\sqrt{((\sigma_h/c)^2+\sigma_t^2)\cdot 2\pi}}\cdot d\epsilon_T \end{aligned} \tag{9}$$

This expresses the probability that $x$ photons reach the $i$-th detector at the time when the $i$-th detector should be receiving the signal from the primary radio source (the source being observed) which will pass the center of the Earth at time $t$.

The equation for the probability that the $i$-th detector detects $z$ photons at that time must also take into account the noise in the detector. In the previous chapter, this was assumed to be additive noise which was normally distributed with zero mean and a standard deviation $\sigma_d$. Thus, the probability that the $i$-th detector detects $z$ photons is obtained by multiplying the probability that there were $x$ photons present and $z-x$ noise in the detector summed over all possible values of $x$ and $z-x$. This is:

$$p_d(z,t) = \sum_{x=0}^{\infty} \frac{p_{i\sum\mathbf{x}}(x,t)\cdot\int_{y=x-\frac{1}{2}}^{x+\frac{1}{2}} e^{-\frac{1}{2}\left(\frac{z-y}{\sigma_d}\right)^2}\cdot dy}{\sigma_d\sqrt{2\pi}} \tag{10}$$

The integral in the above equation raises immediate warning flags about the choice of a normal distribution to model the detector noise. The integral serves to mesh the continuous normal distribution with the discrete summation by rounding off the noise in the detector to the nearest whole photon.

However, the hopes of simplifying equation 10 are slim. On the brighter side, it is clear that when all of the various random variables obtain their mean values and when the primary source is the only source, the detected quantity is precisely the one sought. So, while unwieldy, the model is on track.

Unfortunately, this equation only models a single detector. Taking into account all of the detectors in the array involves determining how the average of variables distributed as in equation 10 is distributed. Rather than tackle this directly, the author wrote a computer simulation employing the above model. This simulation is used in the following chapter to make predictions with this model and to explore its sensitivities.

CHAPTER 6

# Interpretation of the model

Figure 1 shows the mean squared error of a distributed array of radio telescopes as a function of the number of telescopes. In this plot, the only radio sources were the primary target and the sky. The primary source signal was taken to be a sine wave of frequency 100 Hz. *Note: Young nascent pulsars rotate at about 100 Hz, and they slow down as they age.*[1] The source was taken to be 1/1000-th the brightness of the sky. The standard deviation in position measurement $\epsilon_{\mathbf{h}}$ was taken to be 0.8 meters (reasonable under current GPS performance specifications). And, the standard deviation of time measurement $\epsilon_{\mathbf{t}}$ was taken to be 0.01 seconds (well within the current claims of the Network Time Protocol). The detectors were sampled at 1024 Hz.

The detector was taken to be 20 times as noisy as the sky. By noise in the sky, one means the noise caused by the Poisson distribution of photons emitted by the sky. The standard deviation of the Poisson distribution is the square root of its mean. For this simulation, the sky's mean brightness was taken to be 10000 photons per second. Thus, the standard deviation of the sky noise is 100 photons

[1][**UCSC**] second paragraph

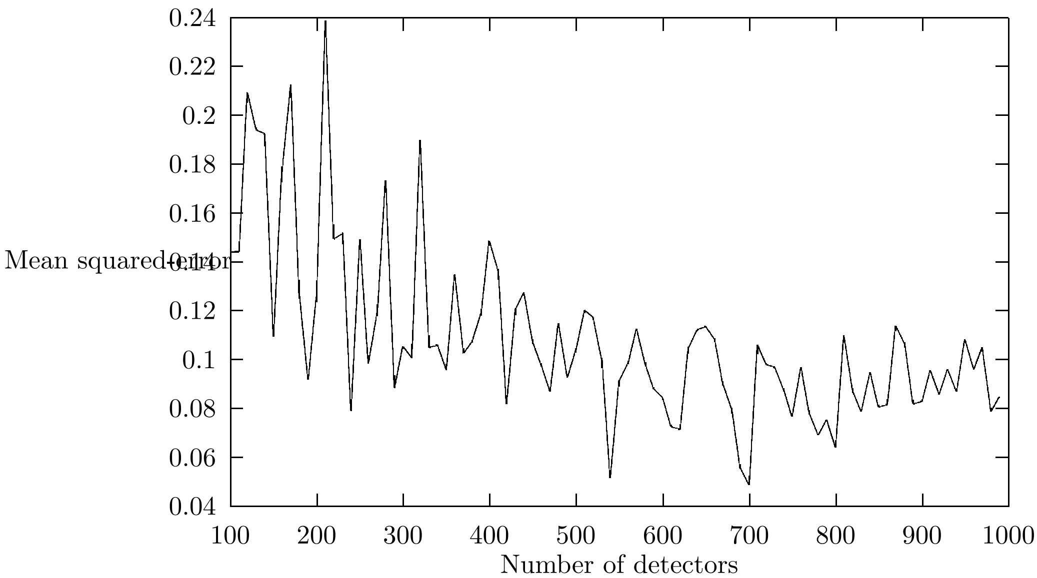


FIGURE 1. Mean squared-error as a function of the number of detectors

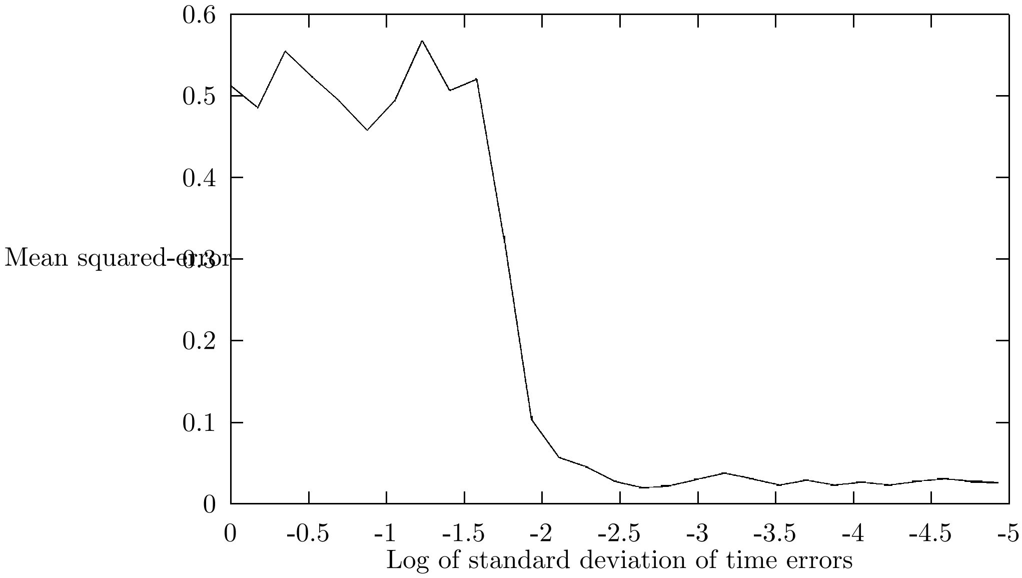


FIGURE 2. Mean squared-error as a function of the $\log_{10}$ of the standard deviation of the time error $\sigma_t$

per second. The detector was taken to have a standard deviation of 2000 photons per second.

From figure 1, it is clear that if only a single radio source (plus sky noise) is being detected by more than 350 telescopes, the mean squared-error easily stays below 0.02 photons per second on a signal whose amplitude is 10 photons per second. That is a mean squared-error less than 1/500-th of the signal against a noisy sky with a very poor detector.

Figure 2 shows the model's sensitivity to errors in time measurements. It plots the mean squared-error as a function of the $\log_{10}$ of the standard deviation of time measurements $\epsilon_{\mathbf{t}}$ for a system of 500 detectors scattered randomly over the surface of the Earth. In this plot, the only radio sources were the primary target and the sky. The primary source signal was taken to be a sine wave of frequency 100 Hz as it was for figure 1. The source was taken to be 1/1000-th the brightness of the sky. The standard deviation in position measurement $\epsilon_{\mathbf{h}}$ was again taken to be 0.8 meters. The detectors were again taken to be 20 times as noisy as the sky as they were for figure 1.

Because the frequency of the source was taken to be 100 Hz, it is unsurpising that the mean squared-error of the array drops off dramatically once the standard deviation of the time error drops below 1/100-th of a second.

For multiple radio sources, the results are still promising. Assuming that each detector is an omni-directional radio antenna (that is, it has a good response to radio signals coming from any direction), the mean squared-error still drops below 2 photons per second with around 150 detectors. Figure 3 shows the situation where there are fifty radio sources in addition to the sky. The mean squared-error here is with respect to the primary source which has an amplitude of 10 photons per second. As with the earlier plots, the detectors are taken to be 20 times as

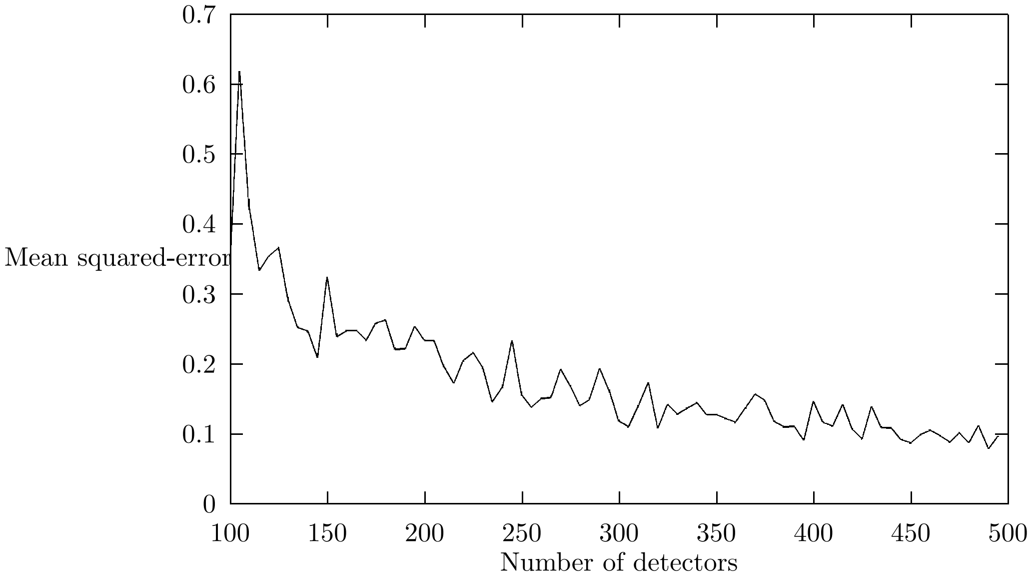


FIGURE 3. Mean squared-error as a function of the number of detectors with multiple sources of interference

noisy as the sky. The amplitude of the primary source is taken to be 1/1000-th the brightness of the sky. The primary source has a frequency of 100 Hz. The other radio sources are scattered about the sky and have similar amplitudes and wavelengths as the primary source.

Because of this, one need not point the individual telescopes in any direction. The array can image the entire sky at one time. One can choose a direction from which to analyze the data by selecting height values for the detectors $h_{i1}$'s appropriate to that direction. Thus, the array of telescopes is quite robust. Even with really poor telescopes, the array can distinguish faint sources against a bright sky amid interference from other sources.

CHAPTER 7

# Further work

Most of the interpretation of this model is based upon a computer simulation. As such, more analysis should be done to verify that the computer simulation faithfully reproduces the model and correctly outputs data. All of the code which the author wrote for this computer simulation is freely available at:

`http://www.nklein.com/products/scope`

The computer simulation assumes also that the detectors attempted to synchronize their sample times (as best they can given the total error in time measurement $\epsilon_T$). In reality, it is more likely that each detector's sampling will have its own time offset and that one would like to know the output of a detector for a time period that is partially in one sample and partially in another sample. For example, if a telescope sampled every 1/1000-th of a second beginning at time zero, but the adjustment factor $t_{i1}$ for that telescope is 1/2500-th of a second, then one would rather the telescope had started recording 1/2500-th of a second earlier. It would be nice to model the situation where the values from a particular telescope must be interpolated from the values that the telescope recorded.

Another assumption in the computer simulation is that the detectors are distributed with equal probability over the entire surface of the Earth and that the pulsar radio sources are evenly distributed over the sky. In reality, chances are good that the detectors will be clustered in high population areas and that the pulsars will most likely be in the direction of the galactic center. A more detailed model of the Earth and the locations and strengths of various pulsars would benefit the simulation greatly.

As mention on page 13, the assumption that the noise in the detector can be modeled with a normal distribution of zero mean is questionable. A more detailed model of the detector noise which takes into account thermal noise, g-r noise, and $1/f$ noise is desirable.

However, despite these concerns, the model paints a promising picture for a distributed observatory. It is hopeful that an array such as this may prove viable.

APPENDIX A

# Distribution of $\epsilon_{\mathbf{h}}/c + \epsilon_{\mathbf{t}}$

Given that $\epsilon_{\mathbf{h}}$ is normally distributed with mean 0 and standard deviation $H$ and that $\epsilon_{\mathbf{t}}$ is normally distributed with mean 0 and standard deviation $T$, we want to find the distribution of $\frac{\epsilon_{\mathbf{h}}}{c} + \epsilon_{\mathbf{t}}$. We will show that the sum is distributed normally with mean 0 and standard deviation $\sqrt{(H/c)^2 + T^2}$.

PROOF. Under the assumptions, the probability that $\epsilon_{\mathbf{h}}$ takes on any particular value $h$ is given by:

$$p_{\epsilon_{\mathbf{h}}}(h) = \frac{e^{-\frac{1}{2}\left(\frac{h}{H}\right)^2}}{H\sqrt{2\pi}} \tag{11}$$

and the probability that $\epsilon_{\mathbf{t}}$ takes on any particular value $t$ is given by:

$$p_{\epsilon_{\mathbf{t}}}(t) = \frac{e^{-\frac{1}{2}\left(\frac{t}{T}\right)^2}}{T\sqrt{2\pi}} \tag{12}$$

We now concern ourselves with the probability that $\frac{\epsilon_{\mathbf{h}}}{c} + \epsilon_{\mathbf{t}}$ takes on any particular value $z$. This probability is given by the following integral which sums the probabilities of each possible combination of $h$ and $t$ that result in a particular $z$:

$$p_{\mathbf{z}}\left(z = \frac{h}{c} + t\right) = \int_{-\infty}^{\infty} p_{\epsilon_{\mathbf{h}}}(h) \cdot p_{\epsilon_{\mathbf{t}}}\left(z - \frac{h}{c}\right) \cdot dh \tag{13}$$

We shall substitute into equation 13 the values of $p_{\epsilon_{\mathbf{h}}}$ and $p_{\epsilon_{\mathbf{t}}}$ given by equations 11 and 12.

$$\begin{aligned} p_{\mathbf{z}}(z) &= \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}\left(\frac{h}{H}\right)^2}}{H\sqrt{2\pi}} \cdot \frac{e^{-\frac{1}{2}\left(\frac{z-h/c}{T}\right)^2}}{T\sqrt{2\pi}} \cdot dh \\ &= \frac{1}{HT \cdot 2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{h^2}{H^2} + \frac{z^2 - 2zh/c + h^2/c^2}{T^2}\right)} \cdot dh \\ &= \frac{1}{HT \cdot 2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left[\frac{\left(H^2/c^2 + T^2\right)h^2 - 2zH^2h/c + H^2z^2}{H^2T^2}\right]} \cdot dh \end{aligned} \tag{14}$$

Equation 14 looks hopeless at first glance. But, because

$$\int_{-\infty}^{\infty} e^{\frac{1}{2}\left(\frac{h-\mu}{\sigma}\right)^2} \cdot dh = \sigma\sqrt{2\pi} \tag{15}$$

it remains only to massage the integral into this form. This is easier than it may appear because it doesn't matter what $\mu$ turns out to be. It doesn't figure into the result of the integral.

To massage the integral into the form of equation 15, we will complete the square in the exponent of equation 14 without introducing any new terms containing $h$.

$$
\begin{aligned}
p(z) &= \frac{\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left[\frac{\left(H^2/c^2+T^2\right)h^2-2H^2zh/c+\frac{H^4z^2}{H^2+T^2c^2}}{H^2T^2}\right]} \cdot e^{-\frac{1}{2}\left[\frac{\frac{H^4z^2}{H^2+T^2c^2}+H^2z^2}{H^2T^2}\right]} \cdot dh}{HT\cdot 2\pi} \\
&= \frac{e^{-\frac{1}{2}\frac{\left(1-\frac{H^2}{H^2+T^2c^2}\right)z^2}{T^2}}}{HT\cdot 2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{h\sqrt{(H/c)^2+T}-\frac{H^2z}{c\sqrt{(H/c)^2+T^2}}}{HT}\right)^2} \cdot dh \\
&= \frac{e^{-\frac{1}{2}\frac{\left(1-\frac{H^2z}{H^2+T^2c^2}\right)}{T^2}}}{HT\cdot 2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{h-\frac{H^2z}{c\left[(H/c)^2+T^2\right]}}{\frac{HT}{\sqrt{(H/c)^2+T}}}\right)^2} \cdot dh
\end{aligned}
$$

Now, the integral is in the form of equation 15. The value of $\sigma$ is $\frac{HT}{\sqrt{(H/c)^2+T^2}}$. The value of $\mu$ is not very pretty, but it does not enter into the value of the integral. Now, we can replace the integral with its value.

$$
\begin{aligned}
p(z) &= \frac{e^{-\frac{1}{2}\left(\frac{z}{\sqrt{(H/c)^2+T^2}}\right)^2}}{HT\cdot 2\pi} \cdot \frac{HT\sqrt{2\pi}}{\sqrt{(H/c)^2+T^2}} \\
&= \frac{e^{-\frac{1}{2}\left(\frac{z}{\sqrt{(H/c)^2+T^2}}\right)^2}}{\sqrt{(H/c)^2+T^2}\sqrt{2\pi}}
\end{aligned}
\tag{16}
$$

Equation 16 shows that $\mathbf{z} = \epsilon_{\mathbf{h}}/c + \epsilon_{\mathbf{t}}$ is normally distributed with mean 0 and standard deviation $\sqrt{(H/c)^2+T^2}$. □

By extension (if one takes $c$ to be 1), one can see that the sum of normally distributed variables $\mathbf{x_1}, \mathbf{x_2}, \ldots \mathbf{x_n}$ with zero means and standard deviations $X_1, X_2, \ldots X_n$ respectively is normally distributed with zero mean and standard deviation $\sqrt{\sum_{i=1}^{n} X_i}$. Particularly, if $X_1 = X_2 = \cdots = X_n = X$, then the sum is normally distributed with zero mean and standard deviation $\sqrt{n} \cdot X$.

APPENDIX B

# Distribution of sum of Poisson distributed variables

Given two variables $\mathbf{x}$ and $\mathbf{y}$ which are Poisson distributed with parameters $X$ and $Y$ respectively, we will show that the sum $\mathbf{z} = \mathbf{x} + \mathbf{y}$ is Poisson distributed with parameter $X + Y$.

PROOF. Because $\mathbf{x}$ is Poisson distributed, the probability that it takes on any particular value $x$ is given by:

$$p_{\mathbf{x}}(x) = \frac{X^x e^{-X}}{x!} \tag{17}$$

Similarly, the probability that $\mathbf{y}$ takes on any particular value is:

$$p_{\mathbf{y}}(y) = \frac{Y^y e^{-Y}}{y!} \tag{18}$$

The probability that $\mathbf{z}$ takes on any particular value $z$ is the probability that $\mathbf{x}$ takes on a particular value $x$ times the probability that $\mathbf{y}$ takes on a value of $z - x$. There are many combinations of $x$ and $y$ that accomplish this. Thus, we must sum the probabilities over each combination. Since neither $x$ nor $y$ can be negative, $x$ can only range from 0 to $z$. The probability then that the sum $\mathbf{x} + \mathbf{y}$ takes on the value $z$ is given by:

$$p_{\mathbf{x}+\mathbf{y}}(z) = \sum_{x=0}^{z} p_{\mathbf{x}}(x) \cdot p_{\mathbf{y}}(z - x) \tag{19}$$

where $p_{\mathbf{x}}$ is given by equation 17 and $p_{\mathbf{y}}$ is given by equation 18.

Substituting equations 17 and 18 into equation 19, we obtain:

$$\begin{aligned} p_{\mathbf{x}+\mathbf{y}}(z) &= \sum_{x=0}^{z} \frac{X^x e^{-X}}{x!} \cdot \frac{Y^{(z-x)} e^{-Y}}{(z-x)!} \\ &= e^{-(X+Y)} \sum_{x=0}^{z} \frac{X^x}{x!} \cdot \frac{Y^{(z-x)}}{(z-x)!} \end{aligned}$$

We can simplify this greatly by multiplying the whole equation by $z!/z!$.

$$p_{\mathbf{x}+\mathbf{y}}(z) = \frac{e^{-(X+Y)}}{z!} \sum_{x=0}^{z} \frac{z!}{x!(z-x)!} \cdot X^x \cdot Y^{(z-x)}$$

The expression $\frac{z!}{x!(z-x)!}$ can be rewritten $\binom{z}{x}$. Because $\binom{z}{x}$ is the coefficient of $X^x Y^{(z-x)}$ in the binomial expansion of $(X+Y)^z$, we can get rid of the summation in the above equation.

$$
\begin{aligned}
p_{\mathbf{x}+\mathbf{y}}(z) &= \frac{e^{-(X+Y)}}{z!} \sum_{x=0}^{z} \binom{z}{x} \cdot X^{x} \cdot Y^{(z-x)} \\
&= \frac{(X+Y)^{z} e^{-(X+Y)}}{z!}
\end{aligned} \tag{20}
$$

This shows that $\mathbf{x}+\mathbf{y}$ is Poisson distributed with parameter $X+Y$. □

By extension, the sum of Poisson distributed variables $\mathbf{x_1}, \mathbf{x_2}, \ldots \mathbf{x_n}$ with respective parameters $X_1, X_2, \ldots X_n$ is Poisson distributed with parameter $\sum_{i=1}^{n} X_i$.

# Bibliography


[Carr] Carr, Joseph J. *Secrets of RF Circuit Design*, Second Edition. TAB Books; New York; 1997. ISBN 0-0701-1673-3.

[Christiansen] Christiansen, W. N. and J. A. Högbom. *Radiotelescopes*. University Press, Cambridge, 1969.

[Kitchin] Kitchin, C. R. *Astrophysical Techniques*, Third Edition. Institute of Physics Publishing; Bristol; 1998. ISBN 0-7503-0498-7.

[SPIE98] Phillips, Thomas G. *Advanced Technology MMW, Radio, and Terahertz Telescopes*; Proceedings of the SPIE; Volume 3357. SPIE—The International Society for Optical Engineering; Bellingham, Washington; 1998. ISBN 0-8194-2804-3.

[Shannon] Shannon, Claude E. "A Mathematical Theory of Communication". *Bell Systems Technical Journal*, vol. 27; July and October; 1948. Downloaded from `http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html` on 2000-10-19.

[Spiegel] Spiegel, Murray R. *Shaum's Outline of Theory and Problems of Statistics*, Second Edition. McGraw-Hill, Inc.; New York; 1988. ISBN 0-0706-0234-4.

[Swenson] Swenson, Jr. G. W. *An Amateur Radio Telescope*. Pachart Publishing House; Tuscon; 1980. ISBN 0-9129-1806-3.

[UCSC] UCSC Compact Object Group. "Radio Pulsars: An Introduction". `http://pulsar.ucolick.org/cog/pulsars/intro.html` viewed 2000-11-05.