Spatial Filtering with Pinholes for MIDI

J. Meisner

[Draft, July 15, 2000]

1 Rationale

Assume an interferometer is observing an object which is completely unresolved by the individual light collectors (telescopes). Without degradation by the atmosphere, the wavefronts entering the telescopes would spatially be perfect plane waves, and temporally be noiselike with a cross-correlation corresponding to the fourier transform of the object structure evaluated at the spatial frequency B/λ, where B is the projected baseline. We refer to that correlation as the *underlying visibility* V. With perfect hardware, the interference of the perfect plane waves would produce an intensity at each point equal to the incoherent sum of the intensities provided by each telescope times the factor $\frac{1}{2}(1 + \Re\{Ve^{j\phi}\})$ where ϕ is the phase offset between the waves as they are interfered. The *measured* visibility is equal to the underlying visibility, since the light from each telescope reaches the photomixer (beam combiner) in the same spatial mode.

If the random effect of propagation through the turbulent atmosphere is taken into account, the complex amplitude profiles of the two beams will no longer be identical. The entire power of either beam will no longer be found in the spatial mode described when the atmosphere had not been present. A random amount of the optical power will be coupled into different modes. At a given time, that distribution will not generally be the same for the light entering each telescope. What's more, the phase relationship between the light interfered in different modes will generally be randomized. Depending on what basis set is used to describe the spatial modes, there may be no correlation whatsoever between the signals interfered in different modes. Thus, rather than the signals from the modes adding coherently, they will add in quadrature. Since, in practice, the central mode will usually contain much more power than the other modes, the resulting interfered signal will (after adding the signals from the other modes in quadrature) be only slightly larger than that obtained from the central mode alone. However an ideal spatial filter would pass *only* power from the central mode, which will therefore interfere with complete spatial coherence (however the magnitude of the instantaneous observed visibility will still also be proportional to the geometric mean of the powers in each of the interfering beams in the filtered mode at any point in time).

There are two advantages to be gained by such a process of spatial filtering. The first has been demonstrated by the FLUOR project, which spatially filters the light from each arm of the interferometer using optical fibers. By accepting only the light in the central mode, the calibration of the interferometer is enhanced, since the observed visibility will be given by the underlying visibility multiplied by the average geometric mean of the powers in the two optical signals, regardless of how much optical power had been coupled into other modes and rejected by the spatial filters. The calibration of visibility then requires only measuring the average power in the combined light (which is measured by default) and knowledge of the ratio of the average geometric mean of the powers to the arithmetic mean of the powers. In the case of FLUOR, the powers in each optical signal prior to the photomixer are monitored in real-time, so that this ratio can be determined directly. A somewhat less accurate system would possibly infer that ratio from measurement of the fluctuations in the *total* power accepted by the spatial filters. In either case, the contribution of light not present in the central mode contributes neither to the interference signal, nor to the photometric measurement, eliminating a major source of uncertainty in the calibration of the interferometer resulting from changes in the magnitude of atmospheric turbulence.

The second advantage of spatial filtering occurs whenever the system noise is not detector limited (as it usually is in the case of FLUOR). In that case, reduction or elimination of light not present in the central mode reduces the system noise level without (greatly) reducing the strength of the interference which (as discussed above) is overwhelmingly dominated by interference in the central mode. In the case of source noise limited operation, reduction of the total amount of light provides a decrease in noise according to the the square root of the portion of light accepted by the spatial filter(s). In the case of *background* limited detection (as we have with MIDI), the advantage is more profound, since a constant amount of background light is present in every mode accepted by the instrument, contributing a noise amplitude, therefore, proportional to the square root of the number of modes contributing to the detector current. For a given aperture, that is proportional to the effective area of an image plane mask, or *pinhole*.

MIDI will be the first interferometric instrument to take advantage of this enormous improvement in the SNR as a result of spatial filtering.

MIDI may also benefit from the aforementioned calibration enhancement, especially when the photometric pick-offs are used to monitor the power levels of the individual interefering beams prior to the photomixer. However at this longer wavelength, detection of absolute power levels is complicated by the presence of low-frequency noise (or "drift") sources either inherent in the detector (not yet measured) or due to small changes in the effective background temperature. The latter can in principle be monitored by the use of chopping, or concurrent observation off-source using additional detector pixels illuminated for this purpose.

2 Noise Performance: A Basic Analysis

A complete analysis of the SNR expected as a function of pinhole size (or more generally, the spatial filter transfer function) would need to take into account the details of the statistics of atmospheric turbulence in coupling power into the spatial modes defined according to some ortho-normal basis set, and any expected correlation between the electric fields coupled into different modes. For our initial analysis, we will take advantage of the fact that the diameters of the VLTI UT's (8 meters) are not much larger than the expected value of the Fried Parameter r_0 (typically 5 meters at 10 microns at Paranal), and that near-perfect tip-tilt correction has been applied by the STRAP unit. In this case, the vast majority of optical power can be found in the central mode, and the PSF of a star image can be roughly approximated as an Airy disk. Thus we can proceed, as a reasonable approximation, as if the atmosphere were not present, and that the PSF results from the diffraction of light by a circular pupil in an otherwise ideal optical system.

2.1 Airy PSF

In this case, the electric field of the light observed at the image plane (i.e. pinhole position) will be radially symmetric and in phase, with an rms amplitude given by the square root of the (radially symmetric) intensity profile:

$$I(r) = I_0 \left[2 \frac{J_1(\pi r)}{\pi r} \right]^2 \tag{1}$$

In the above equation, I_0 is the intensity at the center of the image. J_1 is the first order ordinary Bessel function, and r is the *normalized* radial coordinate, defined as:

$$r = 2A \frac{\sqrt{x^2 + y^2}}{\lambda} \tag{2}$$

where x and y are the *physical* coordinates relative to the optical axis, λ is the vacuum wavelength of light, and A is the numerical aperture of

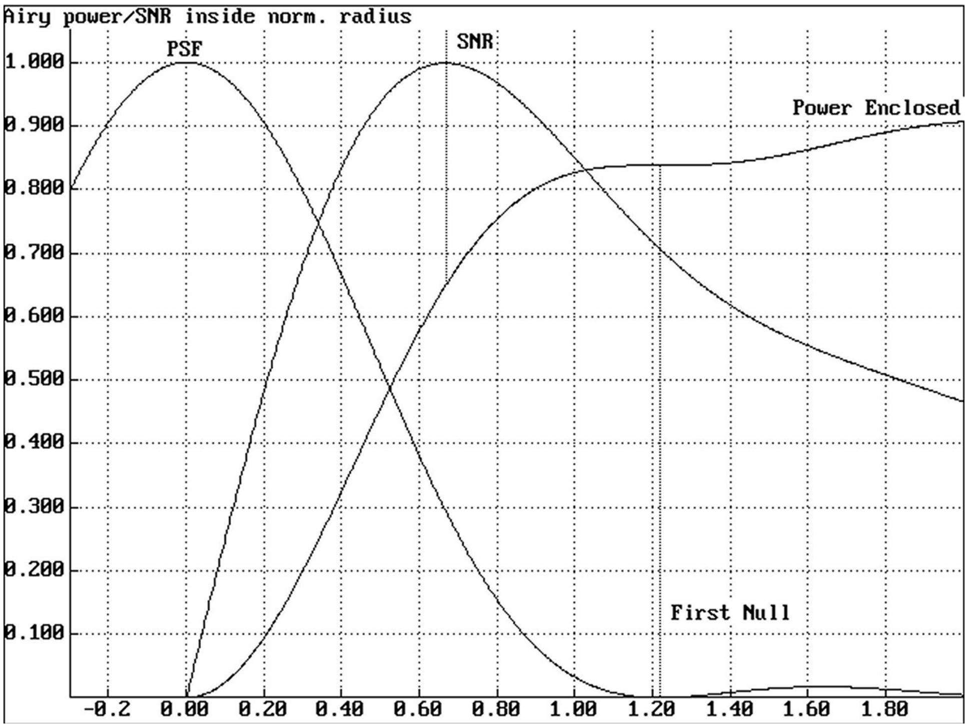

Figure 1: Intensity profile, Signal enclosed, and resulting SNR, for normalized pinhole radius, assuming a signal spatial profile given by the Airy function, and uniform background.

the converging beam. Numerical aperture from a pupil corresponding to a uniform disk focussed in vacuum is defined as the sine of the half-angle of the light cone.

A cross-section of this intensity function is plotted in Figure 1, labelled "PSF"; the actual function is a figure of rotation using this radial profile. As is well known, the first null of the Airy function occurs at $R = 1.22$ as indicated on the graph.

Neglecting scale factors which are constant given A and λ, we can write the signal power admitted by a pinhole of radius R as:

$$P = \int_0^R I(r) 2\pi r \, dr = \frac{8}{\pi} I_0 \int_0^R \frac{J_1(\pi r)^2}{r} dr \tag{3}$$

Expressed as a proportion of the total power in the incident beam $P(\infty)$, we then find:

$$\frac{P(R)}{P(\infty)} = 2\int_0^R \frac{J_1(\pi r)^2}{r} dr \tag{4}$$

This function is also plotted in Figure 1, labelled "Power Enclosed." Note that the horizontal axis of this graph variously denotes a *point* on a radius of the PSF for which $I(r)$ is plotted, and the *outer radius* of a pinhole to which the other two plots refer.

The SNR is proportional to $P(R)$ and inversely proportional to the square root of the total background light admitted. That amount is proportional to the area of the opening, πR^2. Thus the resulting SNR follows:

$$SNR \sim \frac{P(R)}{R} \tag{5}$$

That function is also plotted as "SNR" in Figure 1, normalized relative to its peak value. It can be seen that the SNR reaches its peak at R=.66 and an SNR exceeding 90% of the peak is obtainable over a 2:1 range of pinhole radius. At the point of peak SNR, the pinhole size is such that about 66% of the signal power is admitted by the pinhole.

2.2 Gaussian PSF

The above model using the Airy function is probably appropriate for the case of MIDI whose tip-tilt corrected wavefront will be close to diffraction limited. However it will be instructive to apply the same analysis using a different simple model. In the case of an atmospheric PSF much larger than the diffraction limit, the long-term average intensity profile will be gaussian. Gaussian profiles are also obtained from light exiting an optical fiber. For this case we use the radial coordinate s which is normalized such that the intensity of the PSF is $1/e$ at the point that $s = 1$. Note that this radial coordinate, therefore, bears no fixed relationship to the normalized coordinate r used above, which was determined relative to a specific optical model.

The integration of the power inside a disk of normalized radius R is then found to be:

$$P(R) = \frac{1}{\pi}\int_0^R e^{-s^2} 2\pi s \, ds = \cdots = 1 - e^{-R^2} \tag{6}$$

The SNR is likewise found to be:

$$SNR \sim \frac{P(R)}{R} = \frac{1 - e^{-R^2}}{R} \tag{7}$$

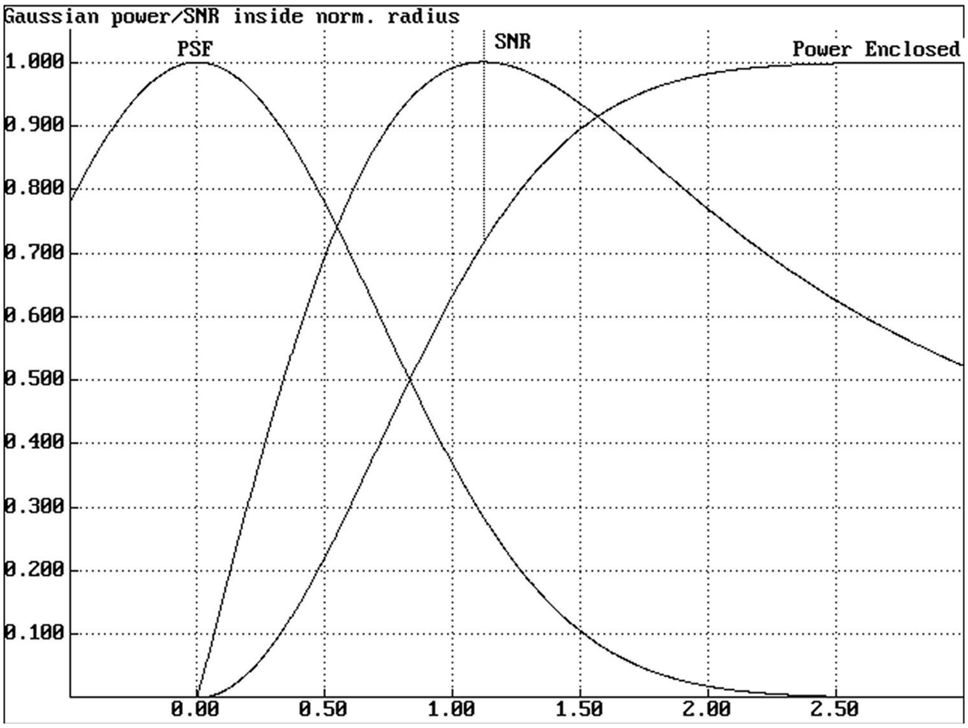

Figure 2: Intensity profile, Signal enclosed, and resulting SNR, for normalized pinhole radius, assuming a gaussian signal, and uniform background.

Again, the radially symmetric gaussian intensity profile, enclosed signal power, and resulting SNR, are plotted (as normalized) in Figure 2. The SNR finds its peak value at $R = 1.121$ and we again see that it achieves 90% of its peak value over a 2:1 range of pinhole radius.

Note that while this analysis would be appropriate for a true gaussian beam in the presence of a uniform background flux, it is *not* at all applicable for analysis of an atmospherically disturbed PSF since the gaussian shape in that case applies to the *long-term* average of the squared amplitude profile, and does not take into account the instantaneous complex PSF (or, equivalently, the random distribution of light into various spatial modes included within the radius R). However we are pleased to note that the point of maximum SNR occurs when 72% of the light is allowed through the pinhole, not far from the 66% figure for the Airy function. Remember that the "R" values plotted on the abscissas of these graphs are unlike, each corresponding to a separate model.

3 Efficiency Compared to an Ideal Spatial Filter

We wish to evaluate the efficiency of a specific spatial filter relative to the performance achievable by an ideal spatial filter. An ideal spatial filter will have an acceptance profile (in amplitude *and* phase) which matches the beam profile of the incoming beam so that its entire power is transmitted. Since it is a single mode device, this ideal spatial filter will transmit the amount of background radiation created in one spatial mode. The *non-ideal* device will accept a different amount of background radiation and less than all of the signal. If η_{sig} of the signal power is accepted by such a device, and it accepts an amount of background radiation contained in N modes, then the resulting SNR using the non-ideal device will equal the SNR obtainable with an ideal spatial filter times η_{net} where:

$$\eta_{net} = \frac{\eta_{sig}}{\sqrt{N}} \tag{8}$$

3.1 Efficiency when filtering a beam from a circular aperture

For a circular pupil function described by an acceptance angle whose sine is A (the so-called "numerical aperture") then an image plane mask of physical area a_0 will accept a number of modes N given by:

$$N = \frac{\pi A^2}{\lambda^2} a_0 \tag{9}$$

In the case of a pinhole of physical radius r_p, this reduces to:

$$N = \frac{\pi^2 \, r_p^2 \, A^2}{\lambda^2} \tag{10}$$

Using the previous definition of "normalized" radius r (in which the numerical aperture of a circular pupil function has been taken into account) we then find the following number of modes accepted by the pinhole whose radius in normalized form is specified to be R_n, we obtain:

$$N = \frac{\pi^2}{4} R_n^2 \tag{11}$$

We had found that the optimum SNR using a circular mask was obtained when $R_n = .66$, at which point we find $N = 1.075$. In other words, the system will accept slightly more power than contained in a single spatial mode. We found that the optimum pinhole accepts $\eta_{sig} = .66$ of the incident light (note that the similarity of the values of R_n and η_{sig} should be regarded as a coincidence). The resulting net efficiency is found to be:

$$\eta_{net} = \frac{\eta_{sig}}{\sqrt{N}} = .64 \tag{12}$$

This is the amount by which the SNR has been reduced relative to an ideal spatial filter using a pinhole of optimum size illuminated by a diffraction limited pattern from a circular aperture.

3.2 Efficiency when filtering a gaussian beam

Although this does not apply to the light received from a physical telescope aperture (or which could be derived from one, since the extent of a gaussian profile mathematically extends to infinity) it is still interesting to perform the above calculation for the case of a gaussian beam. Then we can likewise compare the SNR that would be achieved in the case of a received signal having a gaussian profile in the presence of uniform background radiation using a pinhole, to the theoretically maximum SNR obtainable using a matched filter. Recall that a gaussian aperture is special inasmuch as its effective rms radius (found by weighting the radial coordinate by its power contribution) multiplied by the effective radius of the image plane PSF (likewise weighted) is a minimum relative to all possible transform pairs.

Let us assume that light is being focussed from a far-field gaussian profile described by the half-angle ϕ_0 at which the intensity is reduced to $1/e$ of its on-axis value. Such an aperture will image to a spot also with a gaussian profile as described by (6) where the *normalized* radial coordinate s is related to the physical radial coordinate r_p by:

$$s = \frac{\pi\, \phi_0}{\lambda} r_p \tag{13}$$

The *effective area* of a gaussian function whose amplitude has fallen to $1/e$ at the physical radial parameter R_1 is given by:

$$A_{eff} = \int_0^{2\pi} \int_0^{\infty} r\, dr\, d\theta \;\; e^{-\frac{r^2}{R_1^2}} = \pi\, R_1^2 \tag{14}$$

In the above case, the effective area of the diffraction limited gaussian spot formed by the far-field pattern described will thus be:

$$A_{eff} = \pi\, R_1^2 = \pi \left(\frac{\lambda}{\pi\, \phi_0} \right)^2 = \frac{\lambda^2}{\pi\, \phi_0^2} \tag{15}$$

Now, using a *physical pinhole* whose radius is R_p we would have an effective area given, of course, by πR_p^2. We had earlier calculated that the pinhole radius needed to optimize the SNR was 1.12 in normalized coordinates, with a physical area therefore of:

$$A_{eff} = \pi\ R_1^2 = \pi \left(1.12 \frac{\lambda}{\pi\ \phi_0}\right)^2 = 1.25 \frac{\lambda^2}{\pi\ \phi_0^2} \tag{16}$$

which therefore accepts 1.25 times the background light that would be accepted by a single spatial mode from this aperture (recall that the comparable number in the case of a pinhole optimized for a uniform circular aperture was 1.075).

Again, we can calculate the net efficiency of this sub-optimum filter relative to an ideal matched filter using (8). We had found that .72 of the signal power would be accepted by such a pinhole illuminated by a gaussian beam. Thus we find the relative efficiency to be:

$$\eta_{net} = \frac{\eta_{sig}}{\sqrt{N}} = .64 \tag{17}$$

This turns out to be almost exactly the ratio found for the case of filtering an Airy function using an optimized pinhole.

4 Conclusion

The actual PSF obtained with MIDI will be the squared magnitude of the Airy function (unsquared) convolved with the instantaneous atmospheric OTF. We know that the long term average of the squared magnitude of the atmospheric OTF is expected to yield a gaussian function corresponding to the "seeing disk." At the wavelength of MIDI (and with tip-tilt correction applied to the degraded wavefront), that seeing disk will not be much larger than the diffraction limited PSF, so we could conjecture that the overall PSF will be similar to the squared Airy function itself (as we have previously modelled) but somewhat broadened. Thus it could be conjectured that the optimum pinhole size would still be one that admitted about 2/3 of the unfiltered power. A proper analysis would have to take into account the actual power coupled into various spatial modes by the atmospheric transfer function and the correlation (or lack thereof) between the phases of those components.

However we can state qualitatively, that the *first* motivation for the use of spatial filtering, enhanced calibration of fringe visibility, will favor smaller pinholes, since this provides a greater rejection of non-central modes relative to the central mode. Thus one would tend to favor pinhole sizes which are at or a little smaller than the point simply giving the maximum SNR. This would then perhaps correspond to pinholes accepting only 40 - 50% of the light of the incoming wavefront focussed around the pinhole. In the case of brighter sources (where the SNR is already high) and observations requiring the best calibration of fringe visibility, further reduction of the pinhole size would be called for in order to more completely confine the output power to a single mode.

The wide peak of the SNR peak observed as a result of normalized pinhole size is also fortunate in view of the proposed use of such spatial filters in a broad-band optical instrument such as MIDI whose wavelength range is almost 2:1. By optimizing the center wavelength, the entire wavelength band will still have an SNR within .9 of that obtainable with a pinhole. However it still must be considered that the rejection of source light not in the central mode would be much more complete at low optical frequencies (long wavelengths) than at the high end of the optical bandpass.