With Greatly Limited Resources
Part II: Surface Plasmon Random Scattering

Aaron Webster

March 4, 2011

1 Overview

In the following document I will give a brief account of some of the things we looked at last winter. It is in no way comprehensive, and focuses mainly on certain aspects of weirdospace I believe we have a good explanation for.

Section 2 deals with the numerical simulation, and to that accord two programs were written: a Monte Carlo simulation and an analytic counterpart. Both programs were written to some degree by Bert, so there is certainly overlap in the results, but there were some errors and limitations I hope to have addressed. This section should also serve to succinctly document the programs and the rationale behind their function.

The way plasmon scattering is implemented in the Monte Carlo simulation leads to some interesting statistical distributions of scatterers and their paths. Section 3 deals with this topic and some possible connections with other scattering theories.

Following, Section 5 attempts to give an explanation of the features found in both simulation and experiment. It also makes what I feel is a strong case for the observation of coherent backscattering.

The remaining sections are miscellany that didn't fit in anywhere else. They are presented mostly as results without a lot of commentary. Finally, the Appendix includes the larger figures and comparisons from previous sections.

2 Simulation

A parallel Monte Carlo simulation was written to model plasmon scattering. It assumes the geometry shown in Figure 1 consisting of

- An elliptical *illuminated region*, representing the incident evanescent field used to excite plasmons on the surface. This ellipse has radii $r_a = 5\,\mu\text{m}$ and $r_b = 3\,\mu\text{m}$.
- A *tip* - a single movable scatterer representing the STM tip.
- The *scan area*, a square area wherein the tip rasters, typically 128×128 points in a maximal area of $2.27 \times 2.27\,\mu\text{m}$.
- *Scatterers*, fixed point scatterers (usually 20) distributed randomly in a square of $8.8281 \times 8.8281\,\mu\text{m}$ where the plasmon can visit, and
- A *trajectory* or *path* meaning the ordered sequence of scatterers a plasmon visits before exiting the system. The trajectory has a typical maximum path length of $18\,\mu\text{m}$.

The plasmon is assumed to be a time independent monochromatic plane wave of the form $E(\mathbf{r}) = e^{ik_{\text{sp}}\mathbf{r}}$, where k_{sp} is the plasmon wavevector and $\mathbf{r}$ is some spatial variable. The plasmon trajectories are modeled as follows: First, a random scatterer with index n and coordinates (x_n, y_n) is chosen within the illuminated region as the first scatterer a plasmon visits. Since the illuminated region always encloses the scan area, we are guaranteed at least one scatterer (the tip) to choose from. The plasmon initializes with local phase $\varphi_{\text{l,n}} = k_{\text{sp}} x_n$ representing the local phase advance from the (arbitrary) point where the

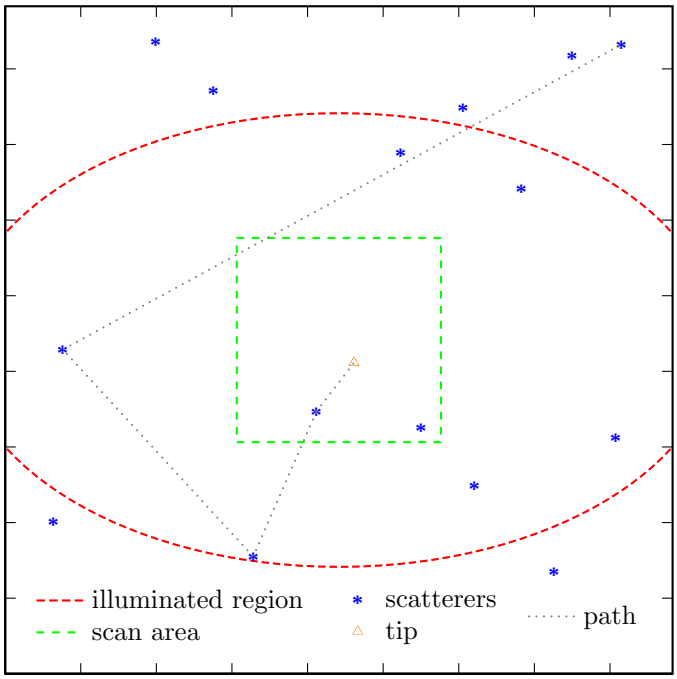


Figure 1: Definition of terms: geometry for the Monte Carlo scattering simulation

plasmon is excited by the (linear in either x or y) evanescent field to the first scatterer. Next, scatterers are randomly and sequentially chosen from amongst all possible scatterers including the tip and the total path length accumulated as the plasmon visits each scatterer. The trajectory ends when

- The same scatterer is visited twice in a row, or
- The total accumulated path length exceeds a pre-defined limit (18 µm).

This process is repeated many times for each tip position in the scan area. Usually 250 000 to 1 000 000 trajectories will be used for each point of the 128×128 grid. During this process the path length, and therefore the total phase from multiple scattering at K sites

$$\varphi_{\mathrm{ms,n}} = \sum_{k=0}^{K-1} \sqrt{(x_{k+1} - x_k)^2 + (y_{k+1} - y_k)^2} \tag{1}$$

is, along with the final scatterer, saved.

The ultimate output from the Monte Carlo simulation consists of *topology maps*, colloquially dubbed *weirdospace*, having the same pixel dimensions as the scan area. Each weirdospace image represents, for each pixel, the far field phase and intensity of a particular point at a certain angle along the plasmon ring for the respective tip position in the scan area. The weirdospace image for each angle ϕ along the ring is calculated by adding a final phase $\varphi_{\mathrm{ff,n}}$ equal to the propagation from the final scatterer to the far field along plasmon scattering angle $\theta_{\mathrm{sp}} \approx 44°$. This is

$$\varphi_{\mathrm{ff,n}} = k_0 \sin\theta_{\mathrm{sp}} \left(x_n \cos\phi + y_n \sin\phi\right) \tag{2}$$

Where $k_{\mathrm{sp}} = k_0 \sin\theta_{\mathrm{sp}}$. Each path therefore represents a wave of the form

$$E_n = e^{i(\varphi_{\mathrm{l}} + \varphi_{\mathrm{ms}} + \varphi_{\mathrm{ff}})} \tag{3}$$

And the total far field intensity is the linear superposition of these waves

$$E_{\mathrm{total}}(\phi) = \left| \sum_{n=0}^{N} E_n \right|^2 \tag{4}$$

In total, the program writes the following data

parameter	description
SX, SY	integer size of scan area
ELLIPSE_A	first radius of elliptical illuminated area
ELLIPSE_B	second radius of elliptical illuminated area
LAMBDA	wavelength of incident light, in microns
PATHLEN	maximum path length for a plasmon
ENTIREAREA	square boundary for random scatterer generation
SCANAREA	dimensions of the scan area in microns
EVENTS	events per point in the scan area
NSCAT	number of randomly distributed scatterers

Table 1: Explanation of variables used in the `parameters` file.

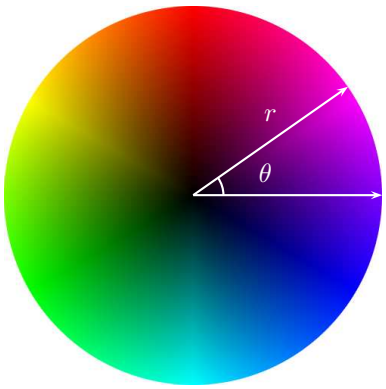


Figure 2: Weirdospace images are colored by using the above model based on a circular coordinate system. Intensity is mapped to the radial component r and phase the angle θ.

- Weirdospace images in a 64 bit floating point HDF5 files. The images are names `XXXXX-scatter.h5`, where `XXXXX` represents the angle in degrees of that particular frame along the ring. Both the phase and intensity are stored in each file.
- `ring.dat`, a tab separated values file containing the angle, mean intensity, and variance of intensity for each weirdospace image. This file is no longer created by the main program, but computed from the output by a supporting script `mkring.sh`.
- `parameters`, a file containing program variables, scatterer locations, and statistics. The parameters are as shown in Table 2. At the end there is a section called `STATS`. The first column of this section represents the number of scatterers visited, the second the percentage of all trajectories including exactly this many scatterers, and the third column the percentage of those trajectories included the tip in some way. The first row is the sum of all the following data.

The weirdospace images are colored in such a way as to show both phase and intensity information at the same time. This is done by mapping the information on to a circular coordinate system as shown in Figure 2, where the intensity is mapped to the magnitude of the radial component r and phase to the angle θ. This coding scheme is identical to that which can be produced in HSV color space with the saturation always being set to unity.

3 Statistical Properties

The way we choose scattering trajectories in the simulation leads to an interesting statistical distribution for the number of scatterers visited. Again, there are two criteria which can end a trajectory

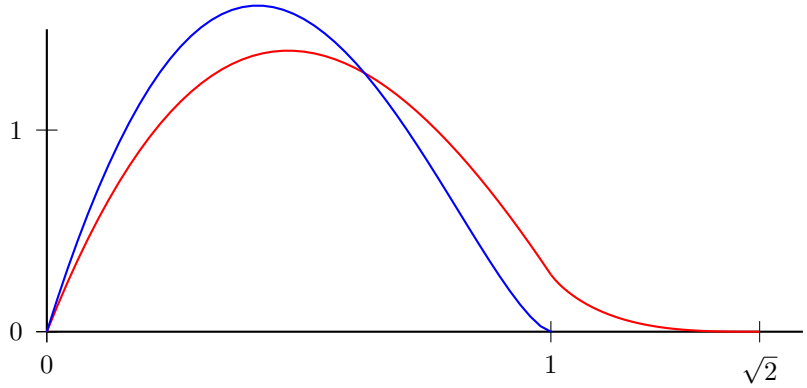


Figure 3: Probability distribution function for the square and disk line picking problems.

- The same scatter is chosen twice.
- The path length reaches a hard limit.

The first exit criteria was chosen ostensibly to provide for the strong presence of single scattering off the tip. Consider a rectangular area containing N scatterers. The probability at each scatterer of choosing pairwise sequential the same scatterer is $1/N$. The extension of this to the probability of the path ending on the nth scattering site, $P(n)$, is modeled by the familiar differential equation for exponential decay

$$\frac{dP(n)}{dn} = -\frac{1}{N}P(n) \tag{5}$$

which, when solved and the initial condition $P(0) = 1/N$ is applied, can be expressed as

$$P(n) = \frac{1}{N}e^{-n/N} \tag{6}$$

This is the lifetime of the plasmon trajectory considering its only exit possibility is due to choosing the same scatterer twice.

The second exit criteria comes when our path length has been exhausted. Since the scatterers are randomly distributed throughout the area, it is useful to know the mean free path for such a system. Consider a unit square. The average distance between any two points is given by the box integral

$$\int_0^1 \int_0^1 \int_0^1 \int_0^1 \sqrt{(x_1 - y_1) + (x_2 - y_2)}\, dx_1 dx_2 dy_1 dy_2 \tag{7}$$

This is the *square line picking* problem. Analytic evaluation of the above integral is complicated, but yields for its unit dimensions an average distance of

$$\overline{r} = \frac{\sqrt{2} + 2 + 5\log\left(1 + \sqrt{2}\right)}{15} \tag{8}$$

and the probability distribution

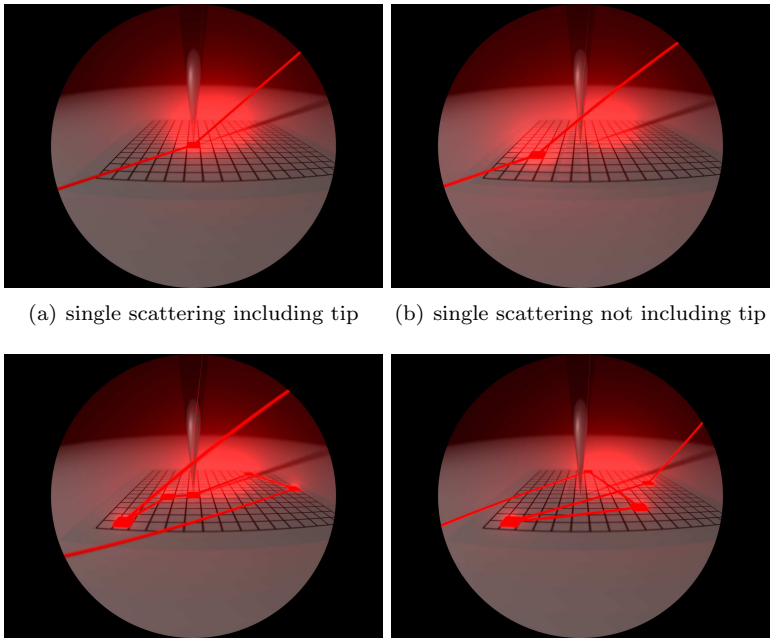
$$P_{\mathrm{s}}(l) = \begin{cases} 2l\left(l^2 - 4l + \pi\right) & 0 \le l \le 1 \\ 2l\left[4\sqrt{l^2 - 1} - \left(l^2 + 2 - \pi\right) - 4\arctan\left(\sqrt{l^2 - 1}\right)\right] & 0 \le l \le \sqrt{2} \end{cases} \tag{9}$$

If we instead suppose a unit disk instead of a unit square, the mean distance can be found to be

$$\overline{r} = \frac{128}{45\pi} \tag{10}$$

and probability distribution for a disk of radius R is

$$P_{\mathrm{d}}(l) = \frac{4l}{\pi R^2}\arccos\left(\frac{l}{2R}\right) - \frac{2l^2}{\pi R^3}\sqrt{1 - \frac{l^2}{4R^2}} \tag{11}$$



(a) single scattering including tip (b) single scattering not including tip

(c) multiple scattering including tip (d) multiple scattering not including tip

Figure 4: Types of scattering events we include or exclude from the simulation.

It is assumed that the superposition of the two probabilities should produce the statistical results found in 19, but as yet I have not found a way to do so.

These line picking problems are useful because they characterize how many scatterers on average a plasmon will visit until exiting in the Monte Carlo simulation. They also reveal a limitation: as the scattering density is increased, a plasmon as simulated does not visit more scatterers, nor does its mean free path decrease. Rather, the result of increasing the number of scatterers is only to suppress single scattering off the tip. This may or may not be physical.

4 Scattering Permutations

Unlike the experiment, the Monte Carlo simulation allows one to exclude certain events from ever being generated. In this case, it is important to note that these trajectories aren't removed from the simulation - rather they are never generated in the first place. Therefore, if you run a simulation specifying N events, the result will be N events of only the type allowed.

Though any arbitrary event can be excluded from the simulation, we are categorically interested in four main types (Figure 4). They are

I Single scattering involving the tip - a trajectory including only the tip.

II Single scattering **NOT** involving the tip - a trajectory including only one scatterer which is not the tip.

III Multiple scattering involving the tip - a trajectory including two or more scatterers where at least one of them is the tip.

IV Multiple scattering **NOT** involving the tip - a trajectory including more than one scatterer where none of those scatterers is the tip.

If we treat each of the four event types as a binary number (1 = included in the simulation, 0 = excluded from the simulation), there are $2^4 = 16$ distinct possibilities amongst them. To this accord, and in the figures to follow, they are arranged with the convention in Figure 5. In each cell, the color gray represents the type of scattering event which has been excluded.

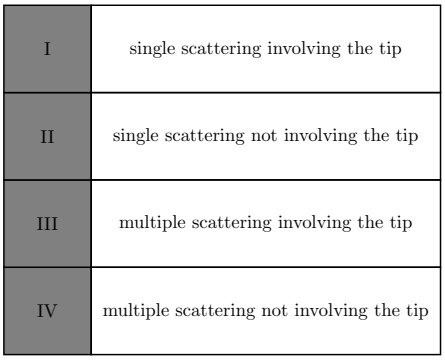


Figure 5: The figures to follow have an accompanying sequence of four vertical squares. These squares represent which events have been excluded from the simulation. Squares in gray represent excluded events.

5 Explanation of Features in Weirdospace

5.1 Type I Events: Single Scattering Involving the Tip

The most dominant features of weirdospace are known as *primary stripes*. Primary stripes are thought to arise from single scattering off the tip. Consider the phase accumulation for this event type. It consists of only two terms: the local phase $\varphi_{\mathrm{l,tip}}$ sampled by the tip position $(x_{\mathrm{tip}}, y_{\mathrm{tip}})$, and the phase to the far field $\varphi_{\mathrm{ff,tip}}$. The intensity in weirdospace for primary stripes is

$$E_{\mathrm{I}} = A_{\mathrm{I}}\, e^{i(\varphi_{\mathrm{l,tip}} + \varphi_{\mathrm{ff,tip}})} \tag{12}$$

where an argument $ae^{i\Phi}$, representing is a fixed background phase required to produce any structure at all is always assumed to be present. As outlined in Bert's thesis, there is excellent agreement between experiment and theory regarding the mathematical description of this feature.

Observe Permutations 0 to 7 in Figure 16. All of these include single scattering off the tip and off of them exhibit primary stripes. The marginal exception is Permutation 5, where primary stripes appear as a phase rather than an intensity feature. Nonetheless, primary stripes appear in all permutations where single scattering off the tip is present, and conversely are suppressed in all where it is not.

5.2 Type II Events: Single Scattering not Involving the Tip

Single scattering not involving the tip can be thought of as the superposition of all single scattering events from all scatterers.

$$E_{\mathrm{II}} = \sum_{n=0}^{N} A_{\mathrm{II},n}\, e^{i(\varphi_{\mathrm{l},n} + \varphi_{\mathrm{ff},n})} \tag{13}$$

In the far field the superposition of many waves with phase terms $\varphi_{\mathrm{l},n}$ and $\varphi_{\mathrm{ff},n}$, themselves essentially uncorrelated and uniformly random tend to cancel each other out. The result of this is a random fixed phase (white noise), shown singly in Permutation 11.

I don't quite have all the math down on this one yet. If we assume the total phase of the argument for this event type is modeled by a uniform random variable, then the real or imaginary part of the exponential will also be a random variable $\in (-1, 1)$. This of course is a random walk with an expectation value of 0 (or very nearly such for a finite number of waves).

5.3 Type III Events: Multiple Scattering Involving the Tip

This is where things get interesting. Consider a multiply scattered path involving the tip where the tip is **not** the last scatterer. The phase to the far field φ_{ff} is determined by the superposition of all the randomly placed scatterers, so it should behave like the phase from a Type II event – an essentially

constant background. The local phase consists of many multiply scattered paths involving the tip, the result of which is like adding a constant to the phase of the tip – the phase from all the paths not involving the tip average out and all you get is the local tip phase. This manifests itself in weirdospace as a fixed background at the plasmon frequency (the "horizontal bars") as a wave of the form

$$E_{\mathrm{IIIa}} = A_{\mathrm{IIIa}}\, e^{i(k_{\mathrm{sp}}x)} \tag{14}$$

This can easily be seen in the 180° scattering direction for all permutations which include multiple scattering involving the tip. Also, note that only in cases where multiple scattering involving the tip has been suppressed does this kind of background disappear.

Again, here I haven't reached the result analytically, but numerical estimates confirm that, say for an infinite sum of cosines and a random variable p_n

$$\lim_{N\to\infty} \sum_{n=0}^{N} \frac{\cos(r+p_n)}{N} = \cos(r+\vartheta) \tag{15}$$

where ϑ is a fixed phase.

Next, consider a multiply scattered path involving the tip where the tip is the last scatterer. The statistical result is similar to that of the previous case, but here the local phase becomes a constant fixed phase and the phase to the far field is all which is preserved. This manifests itself as a set of stripes of fixed wavelength, making a complete rotation every rotation around the ring.

$$E_{\mathrm{IIIb}} = A_{\mathrm{IIIb}}\, e^{ik_{\mathrm{sp}}(x_n\cos\phi + y_n\sin\phi)} \tag{16}$$

At 0 and π rad these are indistinguishable from the previous case, and orthogonal at $\pi/2$ and $3\pi/4$ rad. It should be noted here that if the path involved the tip twice where the tip occurs in both the local phase and the phase to the far field, this is identical to a Type I event.

The combination of the two Type III events with the Type I and II events explain much of the structure around the ring, as are shown in the figures in the Appendix.

5.4 Type IV Events: Multiple Scattering Not Involving the Tip

Somewhat anticlimactically, this results in another fixed background. Both the local phase and the phase to the far field are determined statistically from the group as a whole in the same way as Type II events.

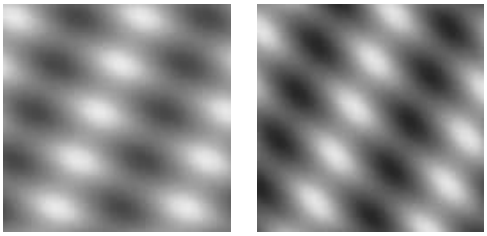
5.5 Coherent Backscattering or Complex Exponentials and the Error of your Ways or How I Learned to Stop Worrying and Love the Bomb

Since coherent backscattering (CBS) is a general multiple scattering phenomenon, it was though that it would be seen in the experimental weirdospace data - manifest as so called *secondary stripes*. This would be a wave identical to that of single scattering off the tip, but excited by an evanescent field traveling in the opposite direction, $\varphi_{\mathrm{l}} \to -\varphi_{\mathrm{l}}$.

$$E_{\chi} = A_{\chi}\, e^{i(-\varphi_{\mathrm{l}}+\varphi_{\mathrm{ff}})} \tag{17}$$

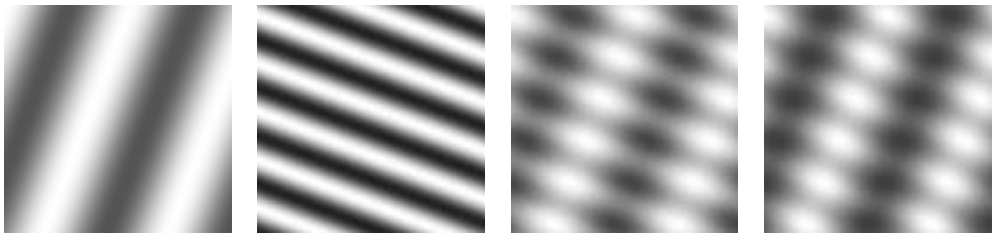
In Bert's thesis, the idea that secondary stripes were connected to CBS was rejected based on analysis of the experimental data showing secondary stripes were at the wrong angle. I intend here to give evidence to the contrary.

Figure 6 shows the analytically predicted orientation of secondary stripes along with the best fit orientation of such stripes to an experimental weirdospace image for a particular angle along the ring. If you look carefully, you'll notice something strange about the figure: the secondary stripes almost perfectly "line up". This is consistent with one computing the intensity as the sum of the squares rather then the square of the sums, the later befitting for wave superposition. This is detailed in Figure 7. Here we show the approximate orientation of primary and secondary stripes for the experimental data



(a) expected (b) best fit from experiment

Figure 6: Two figures from Bert's thesis. To the left, the expected orientation of secondary stripes and to the right, the best fit from the experimentally measured orientation.



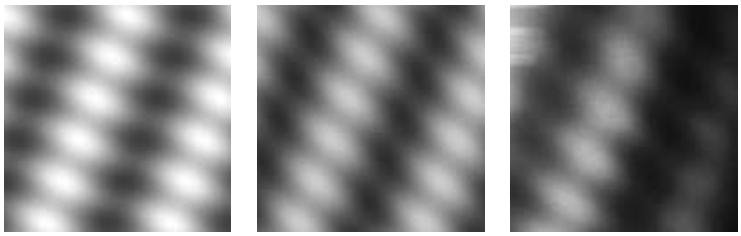
(a) $e^{i\alpha}$ (b) $e^{i\beta}$ (c) $\left|e^{i\alpha}\right|^2 + \left|e^{i\beta}\right|^2$ (d) $\left|e^{i\alpha} + e^{i\beta}\right|^2$

Figure 7: Adding phases from primary and secondary stripes results in a characteristic interference pattern not consistent with the addition of intensities.

in question along with the resultant sum of squares and square of sums. In this image, the difference between the two is marginal. Indeed, a priori interpretation of the images would not lead one to think they were anything but inconsistent with the experiment. However, by adjusting the relative amplitudes of these waves as in Figure 8, assuming E_χ to be much less than E_I, the analytic result comes to match the experiment nearly perfectly. It should be noted that the figure in Bert's thesis shows a CBS signal about 1/3 the amplitude of single scattering off the tip, which doesn't seem to fit well with what one would assume for such an effect.

There are many more examples of secondary stripes and their relation to experiment in the Appendix.

We've actually seen this before in the Monte Carlo simulations, but didn't know what to make of it without an explanation of the weirdospace features caused my multiple scattering. Figure 13 shows the difference between analytic, with E_χ removed, and Monte Carlo simulations with time reverse paths removed. The results are almost exactly what one would expect if secondary stripes were truly a coherent backscattering phenomenon. Amazing.



(a) $|0.29 \cdot E_\mathrm{I} + 0.08 \cdot E_\chi|^2$ (b) $|2.44 \cdot E_\mathrm{I} + 0.23 \cdot E_\chi|^2$ (c) experiment

Figure 8: By adjusting the relative amplitudes of the waves predicted for primary and secondary stripes, good agreement can be reached with the experimental data.

6 Comparisons Between Experiment, Simulation, and Theory

The following are comparisions between wierdospace images produced in the experiment, through the Monte Carlo simulation, and by analytic means. Noise has been added to some of the analytic images to "get you in the right mood".

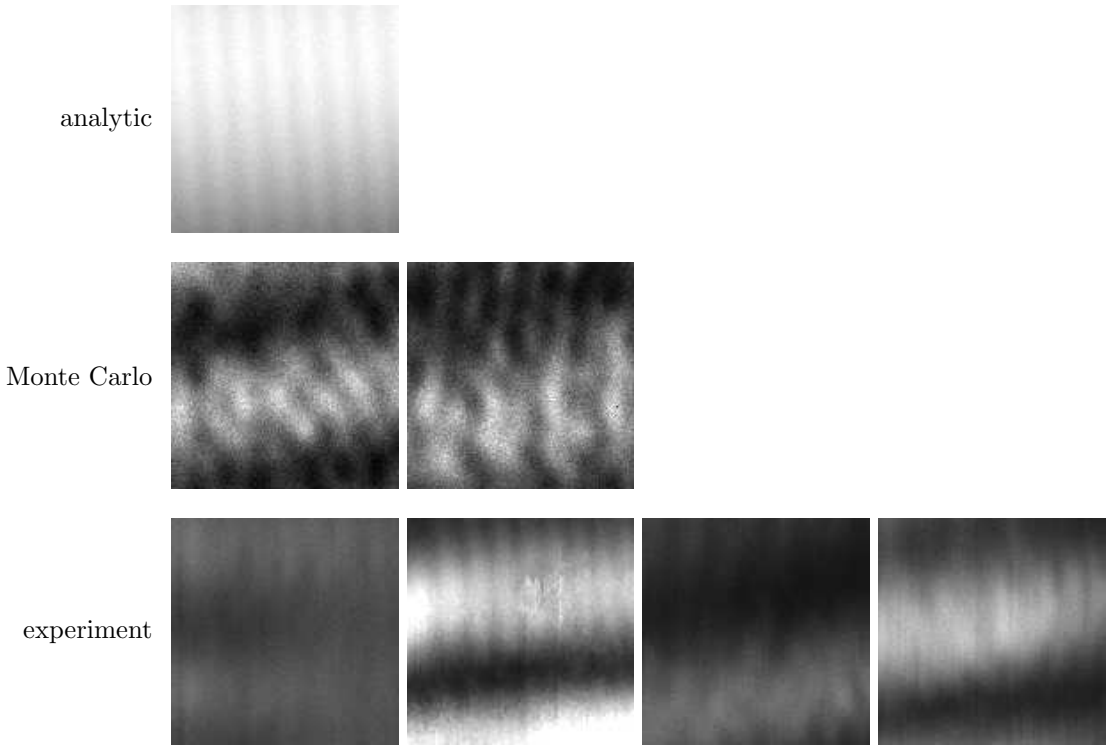
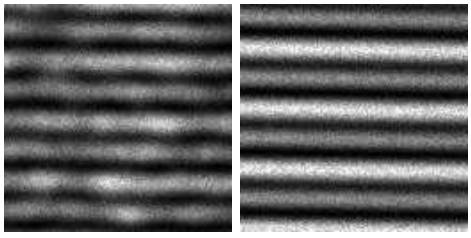


Figure 9: Comparison of analytic, Monte Carlo, and experimental weirdospace in the vicinity of the 180° scattering direction. The vertical features are secondary stripes produced by Equation 17.



(a) Monte Carlo (b) analytic

Figure 10: Comparison of Monte Carlo and analytic weirdospace in the vicinity of the 0° scattering direction. The alternating intensity pattern is an interference between primary stripes and the two Type III events.

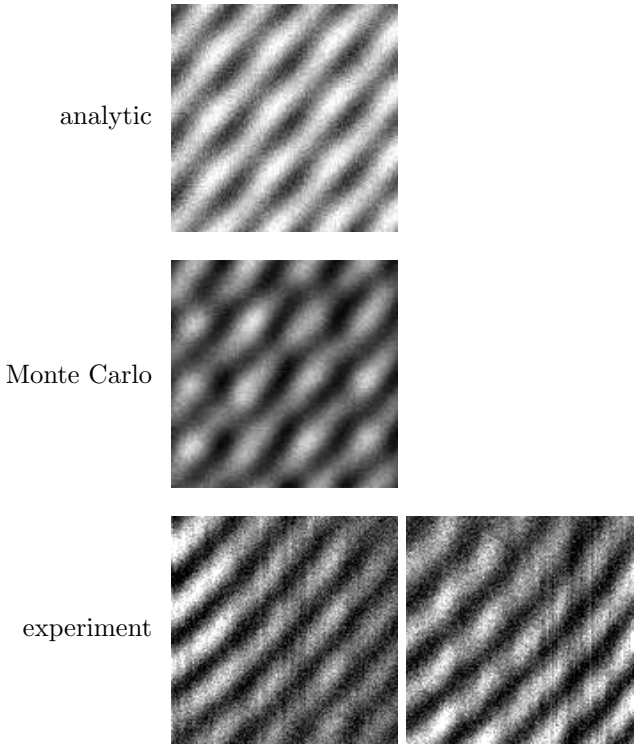


Figure 11: Comparison of analytic, Monte Carlo, and experimental weirdospace in the vicinity of the 85° scattering direction. Note the distinct shape of the dark pockets predicted by the analytic and Monte Carlo simulations. This shape is due to interference of the two Type III events.

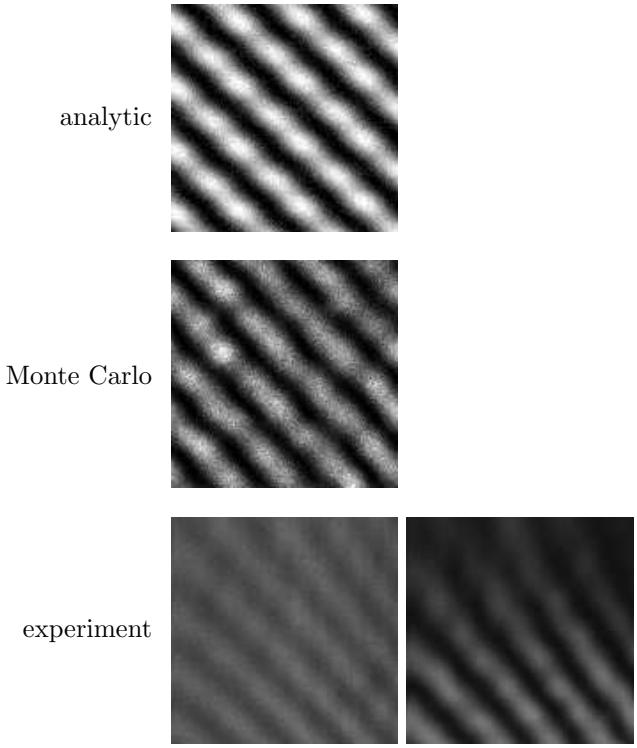
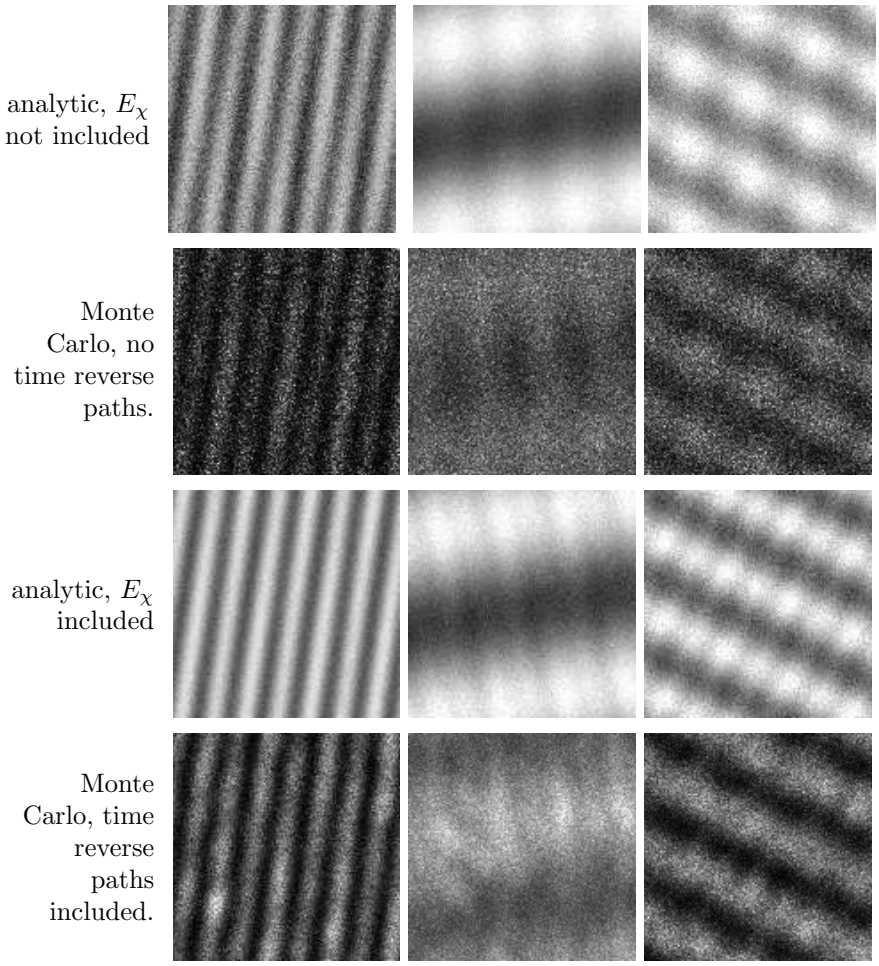
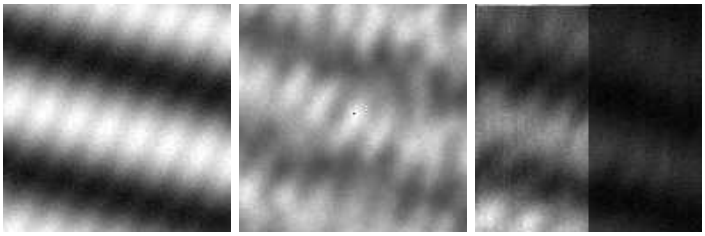


Figure 12: Comparison of analytic, Monte Carlo, and experimental weirdospace in the vicinity of the 270° scattering direction. Note the orientation of the secondary stripes and that they come in intensity groups of two.



(a) 16° (b) 160° (c) 233°

Figure 13: Comparison of analytic and Monte Carlo simulations in the case where time reverse paths (Equation 17) have been removed from either.



(a) analytic (b) Monte Carlo (c) experiment

Figure 14: Comparison of analytic and experimental weirdospace for some angle around the ring. Note the relationship between intensity maxima on each primary stripe and their "hexagonal packing" relationship. Also, the way these maxima connects with another with a thin, curvy intensity. This interference effect is from primary and secondary stripes.

(a) scan538 (b) scan645 (c) scan645

Figure 15: Exhibition of some of the stripe features found in the experimental data. Note the (alternating) relationship between intensity maxima on each primary stripe and their "hexagonal packing". This interference effect is from primary and secondary stripes.

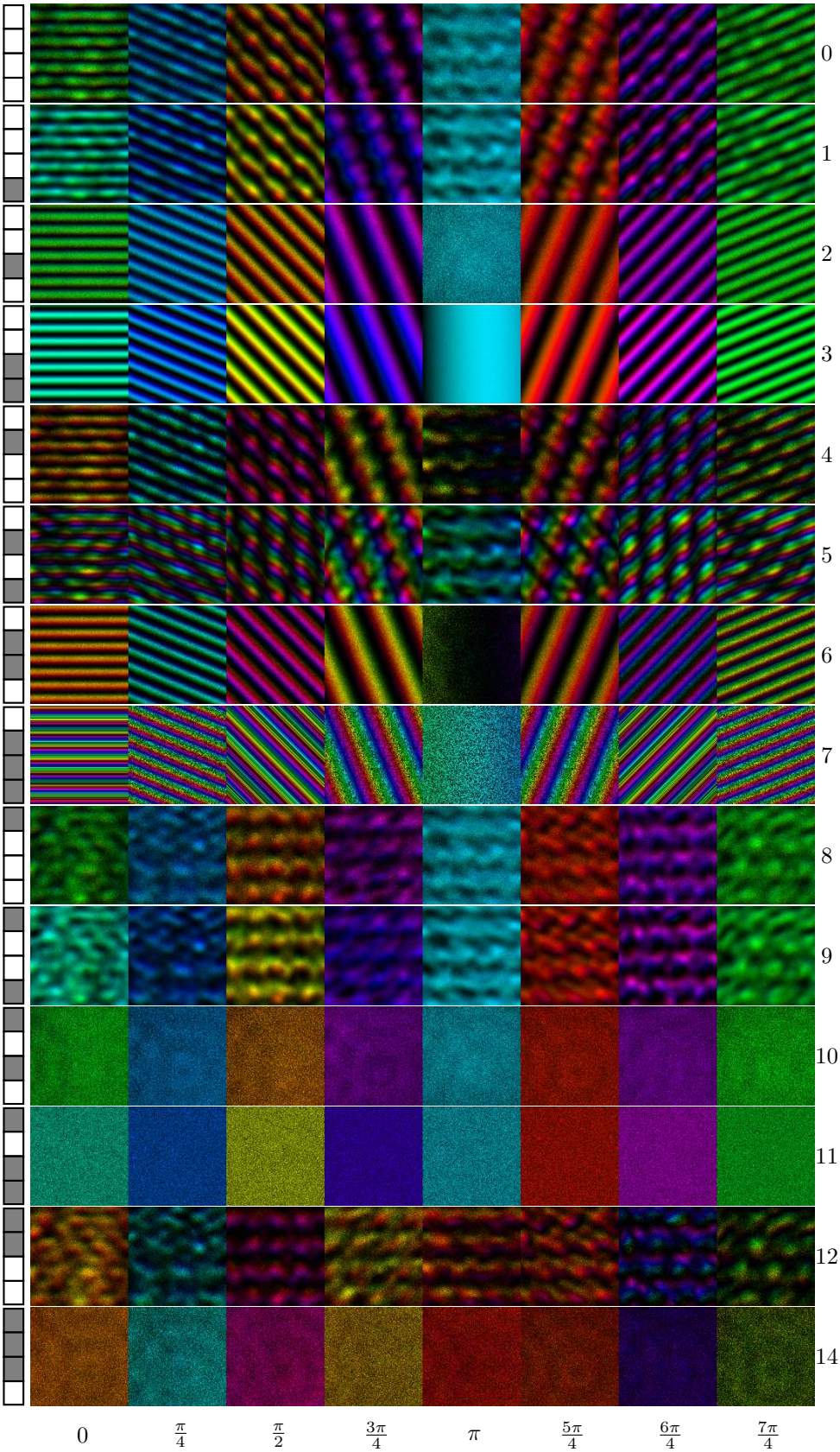


Figure 16: Permutations of the Monte Carlo simulation around the ring: intensity and phase. Intensity values have been normalized for each frame while phase values normalized across the entire dataset.

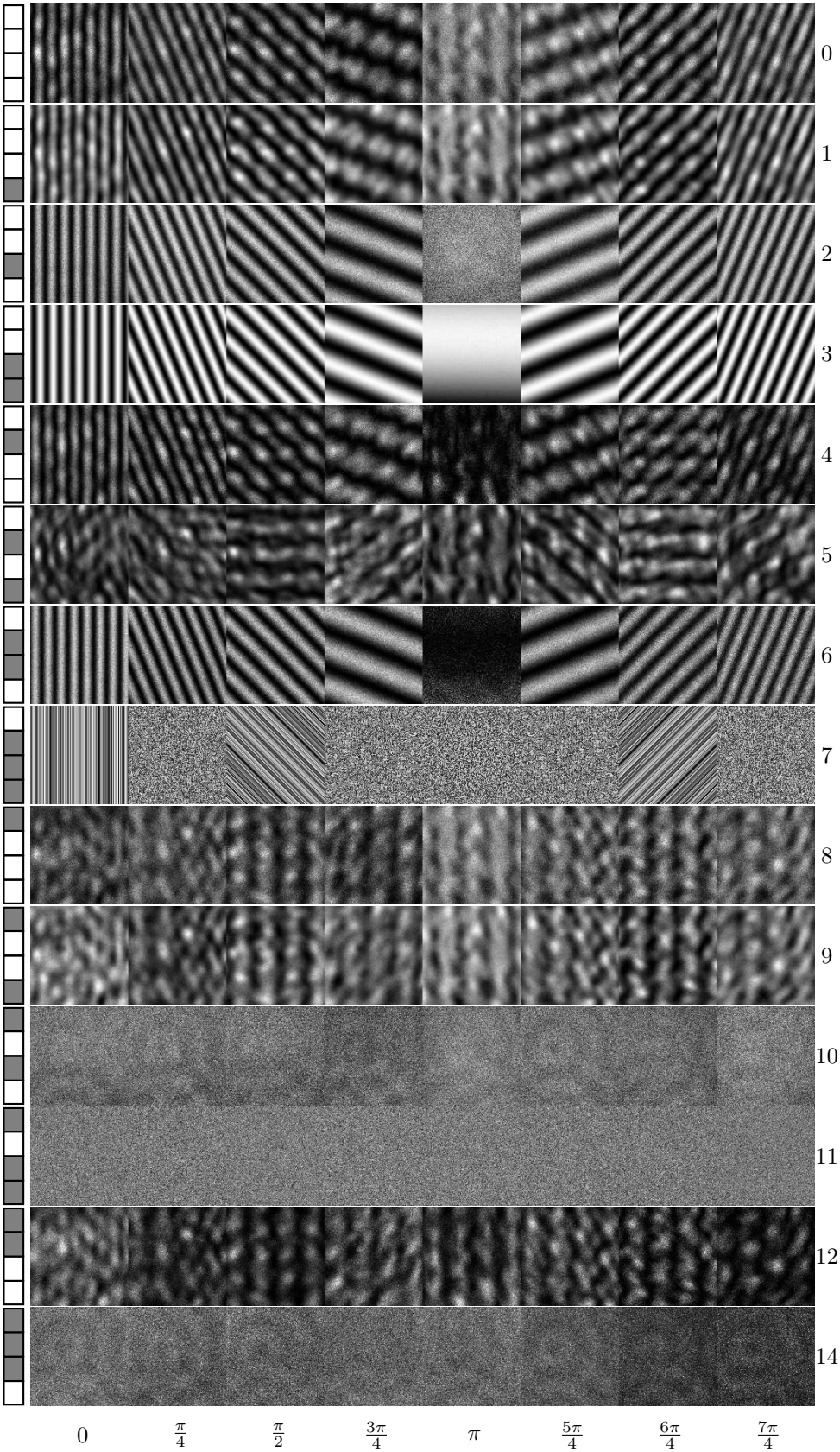


Figure 17: Permutations of the Monte Carlo simulation around the ring: intensity only. Intensity values have been normalized for each frame individually.

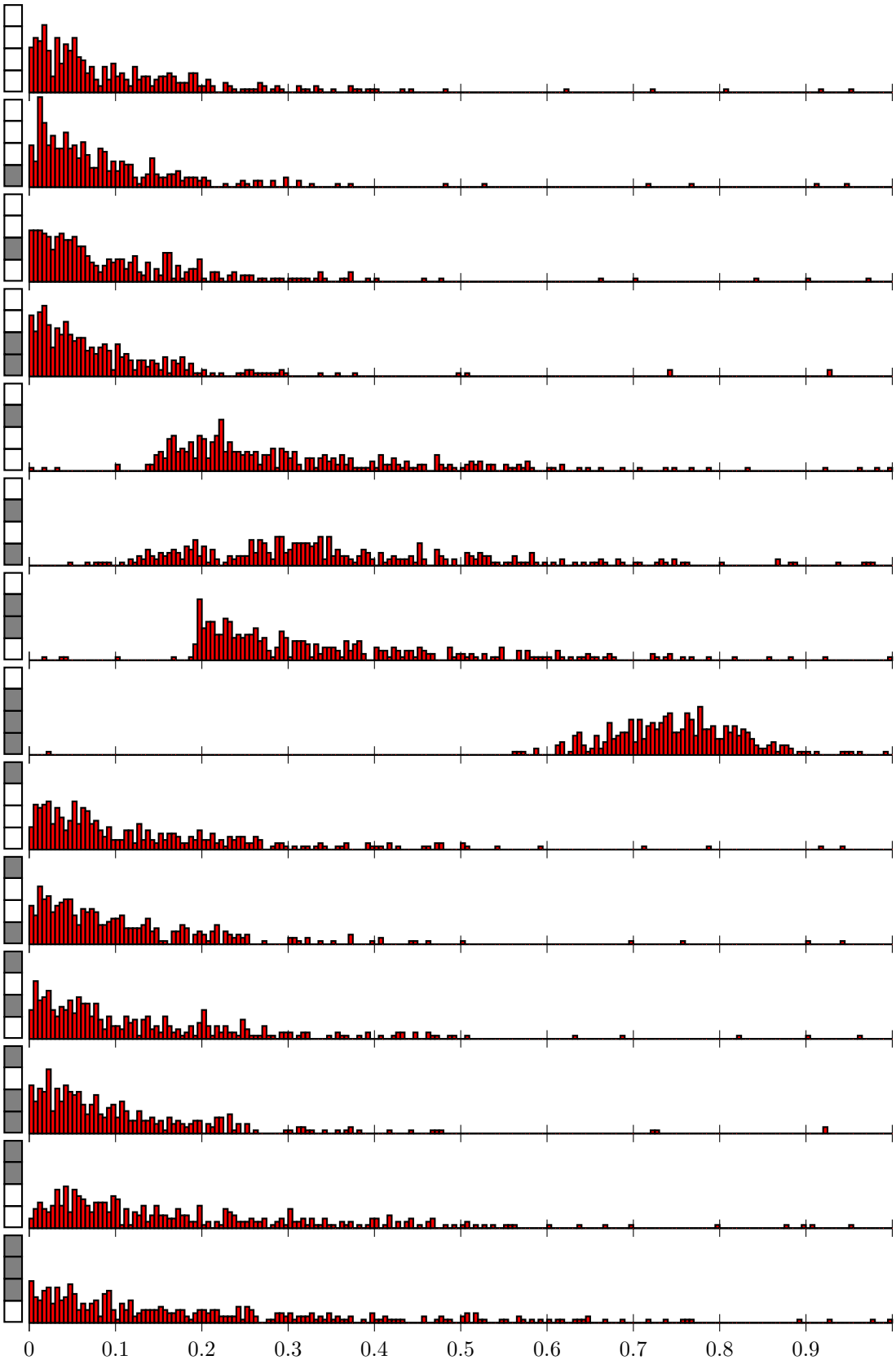


Figure 18: Probability distribution function, plotted here as a histogram, of the normalized binned averaged intensities of weirdospace images around the ring.

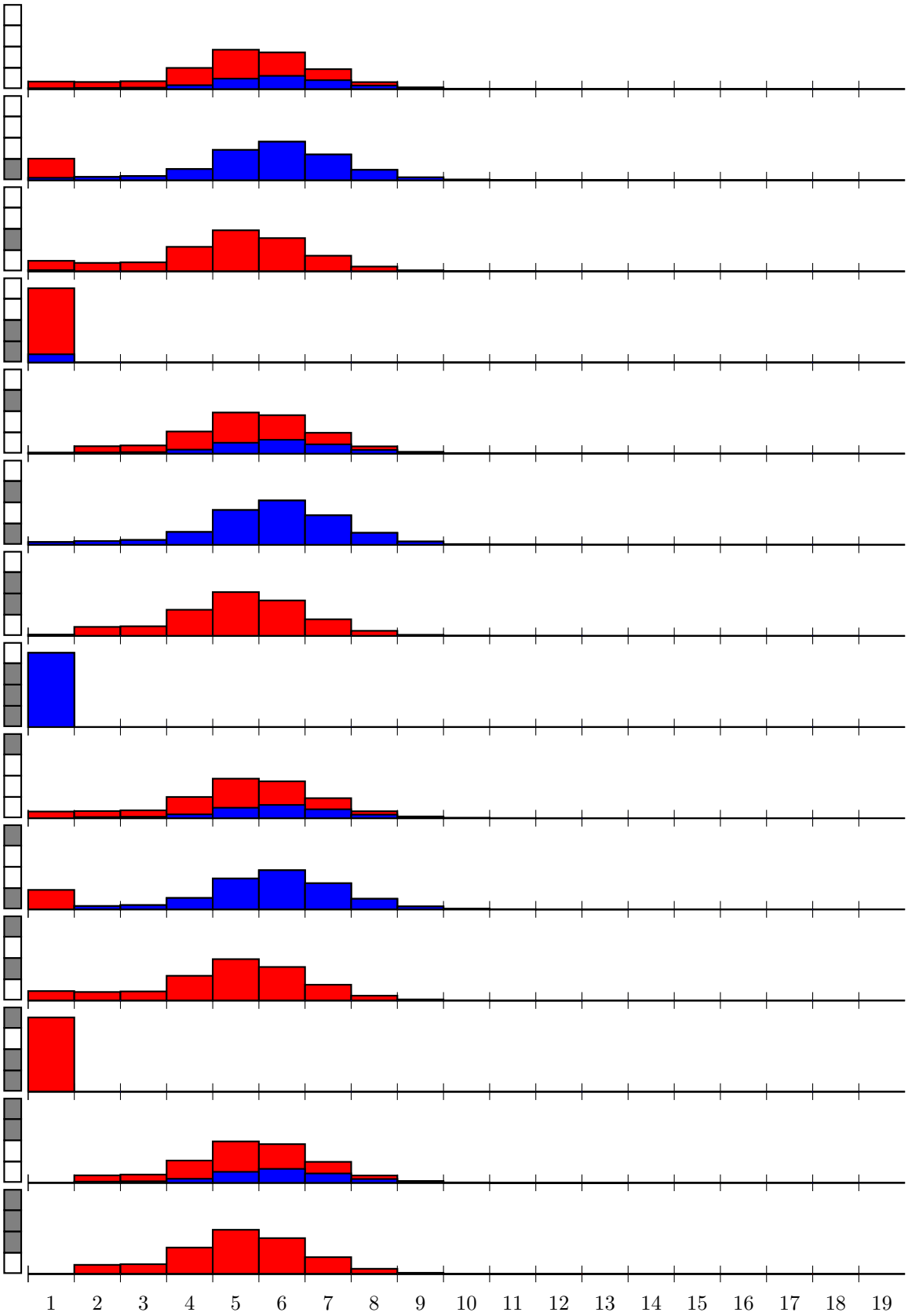


Figure 19: Histogram showing the normalized distribution of scattering paths having visited a certain number of scattering sites. All paths are shown in red, and the proportion of those paths including the tip in some way are shown in in blue.

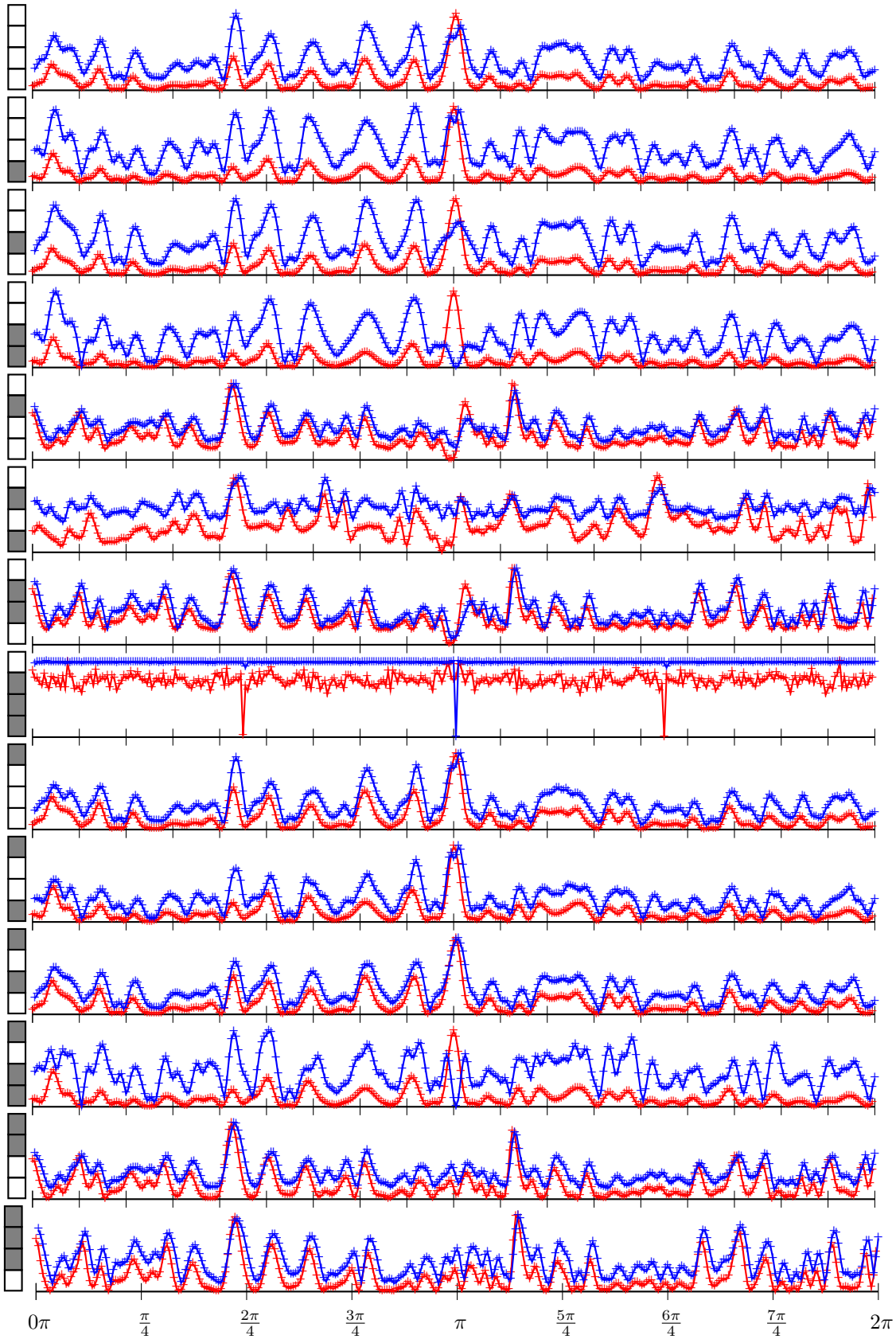


Figure 20: Normalized intensity (red) and normalized contrast (blue) of weirdospace images around the ring. Contrast here is defined as the average of the top 1 % minus the average of the bottom 1 % of the intensity values.

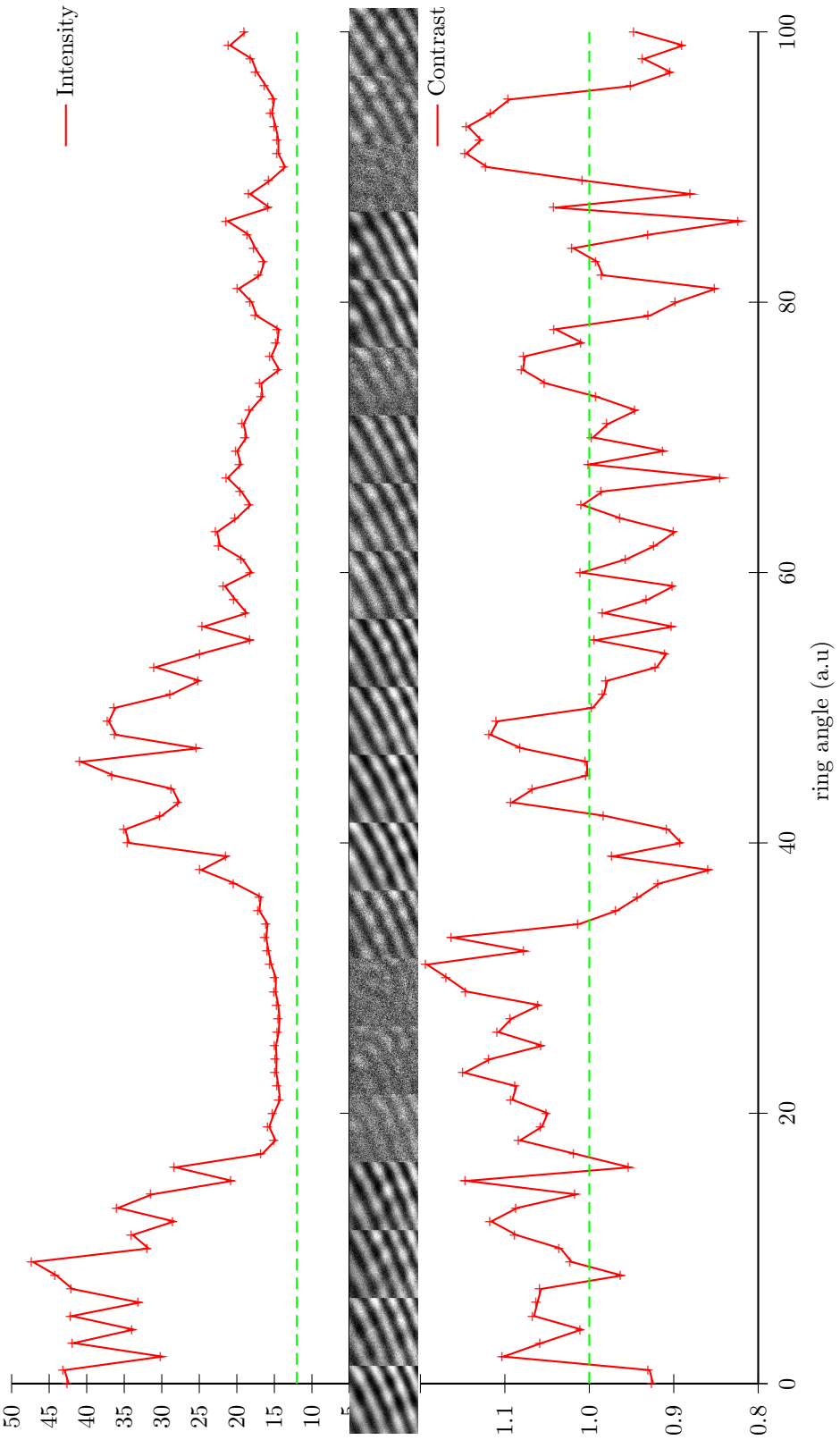


Figure 21: Intensity and contrast in a "vortex" area. Contrast is defined here as the ratio of the range (average of the top 1 % minus the average of the bottom 1 % of the intensity values) divided by the mean intensity. Note that in certain dark areas of the ring you can still still have as much contrast as in some of the the brightest places. The green dashed line in the intensity plot is the noise threshold.

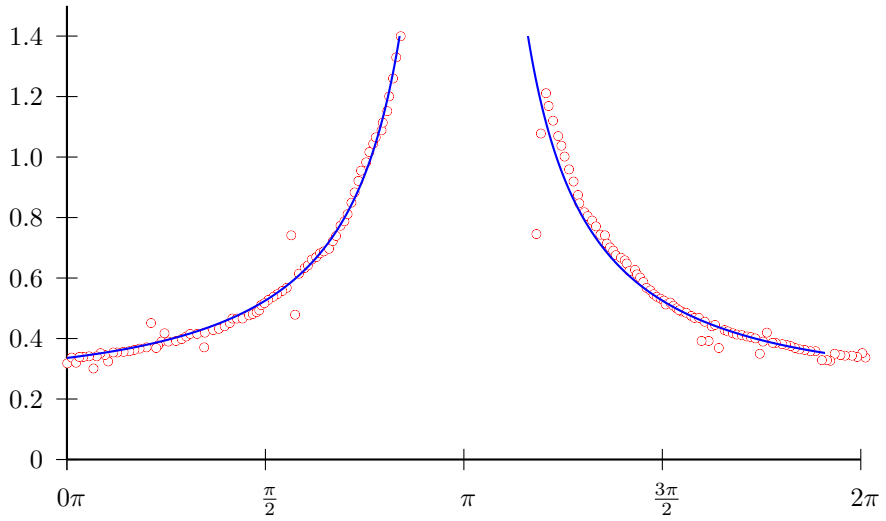


Figure 22: Primary stripe angle as a function position around the ring: experiment and theory. This data was taken from Bert's thesis.

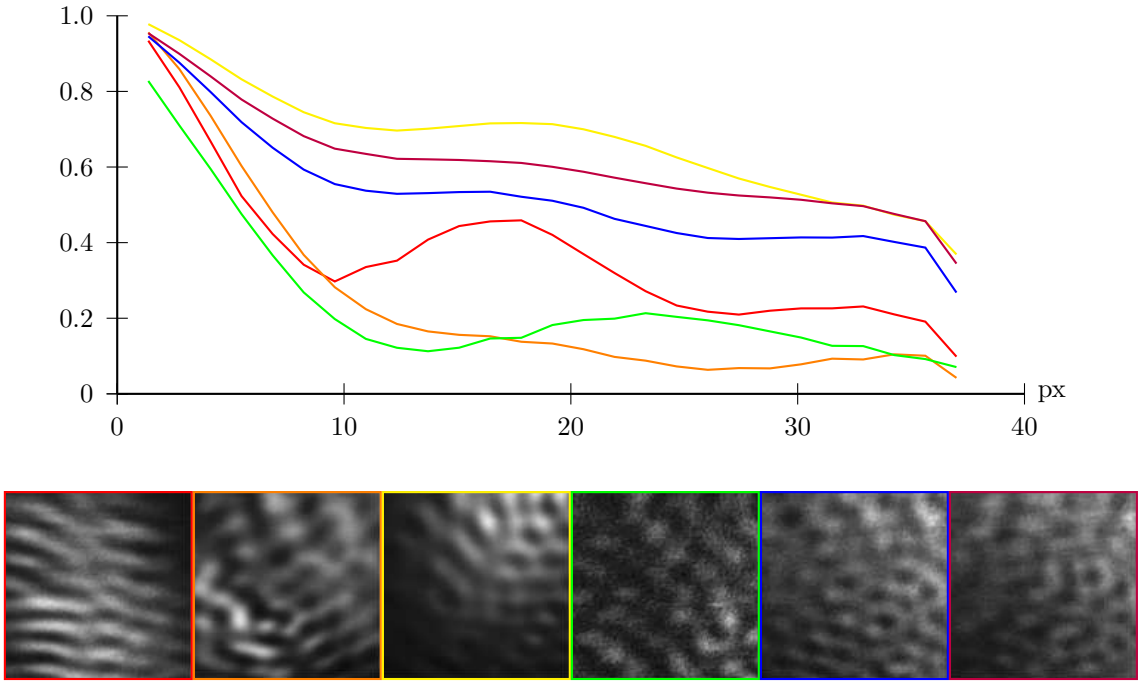


Figure 23: Radially averaged autocorrelations taken in some of the dark areas around the ring for the experimental weirdospace images. Units are in pixels.