The "Intrinsic" Obscured Quasar Bias

Motivation

My recent work shows that the obscured quasar clustering amplitude (and therefore the bias) is higher than that of unobscured quasars. This implies that the simple orientation model is incorrect. But, surely some subset of the obscured sample is actually just unobscured objects seen at the right viewing angle so the "torus" makes them seem obscured. Therefore, the bias of the quasars obscured for some other reason is actually larger than what I'm measuring. The question is - can we use modeling to figure out what that bias is?

Method

We need to know three things:

1. The (intrinsic) clustering/bias of unobscured quasars.

2. The *observed* clustering/bias of obscured quasars (combined torus and "other" obscured).

3. The level of "contamination" in our obscured sample, i.e. how many objects are really just the same as unobscured objects but are seen through the torus?

Numbers 1 and 2 come straight from my paper. Number 3 can be determined from a given torus model. In the simplest case, you just say that from 0 to ϕ degrees (0 being pole-on), you see an unobscured quasar, and from ϕ to 90 degrees it is obscured. For a given sample size (the obscured sample in my paper for example), you just use the probability of seeing a quasar at a given ϕ (just a sine function in this simple case) to determine the number of torus obscured things (N_{torus}) and "other" obscured things (N_{other}). You could of course make this more complicated, such as with a luminosity dependent opening angle, but we'll start simple. Note that there is an implicit assumption here, that the obscured sample is unbiased in orientation in any way. Maybe not true, but not obviously untrue either.

To back out the "intrinsic" obscured quasar bias from what is observed, you follow these steps:

1. Make an initial guess of the intrinsic clustering properties of obscured quasars, and build a mock sample of size N_{other} with these properties.

2. Make a mock sample of size N_{torus} with clustering properties of unobscured quasars.

3. Combine these samples, and measure their bias like you would with real data.

4. Return to (1), and loop over some parameter space for the intrinsic clustering.

5. Compare these results to the observed bias of the total obscured sample, and find the best match to see what the most likely "intrinsic" bias is.

Mock samples

How do we build a mock sample with given clustering properites/bias? We start with the definition of the angular autocorrelation function $\omega(\theta)$, where θ is the projected separation on the sky:

$$dP = n(1 + \omega(\theta))d\Omega. \tag{1}$$

dP is the probability that a given pair of objects with number density n (easily converted to number N in our case, because we'll be using samples over the same areas) with separation θ lie within solid angle $d\Omega$. We assume that the autocorrelation function is a power-law:

$$\omega(\theta) = A\theta^{\gamma}. \tag{2}$$

The factors A and γ for the unobscured and total obscured samples come from my paper. Integrating over larger and larger annuli (whose radii are determined by the binning of the real data we are comparing to (e.g. 4/dex) essentially gives us the number of objects to place at within $\theta_1 < \theta < \theta_2$ from each point:

$$P = \int n(1 + A\theta^\gamma)d\Omega \tag{3}$$

$$= n\left[\int d\Omega + A\int \theta^\gamma d\Omega\right]$$

$$= n\left[\int_0^{2\pi}\int_{\theta_1}^{\theta_2} \sin\theta d\theta d\phi + A\int_0^{2\pi}\int_{\theta_1}^{\theta_2} \theta^\gamma \sin\theta d\theta d\phi\right]$$

$$= 2\pi n\left[\int_{\theta_1}^{\theta_2} \sin\theta d\theta + \int_{\theta_1}^{\theta_2} \theta^\gamma \sin\theta d\theta\right]$$

$$= 2\pi n\left[\cos\theta_1 - \cos\theta_2 + \int_{\theta_1}^{\theta_2} \theta^\gamma \sin\theta d\theta\right]$$

I think this last integral will have to be done numerically, but shouldn't be a problem. I still haven't quite worked out in my head how to take this probability and apply it properly to generate a list of RA and Dec values of the mock sample, but it should be possible. Once you can generate the mock samples, then the rest is just repeating with different inputs, and measuring the clustering like you would on real data.

Maybe their is a purely analytical way to get the answer using the equations above, but it isn't immediately obvious to me.