Biases in CNC analysis

Notation

q —> SNR

Ideal —> Refers to noise variance being known i.e. $\langle \sigma^2 \rangle$

AK —> (**A**ll **K**nown) Size & location is assumed perfectly known

SK —> (**S**ize **K**nown) Size of the cluster is assumed perfectly known

LK —> (**L**ocation **K**nown) Location of the cluster is assumed perfectly known

NK —> (**N**o **K**nown)

Semi-ideal —> Noise realization known

Real —> Totally blind matched filtering analysis

Real iterative —> Real + estimated signal subtraction

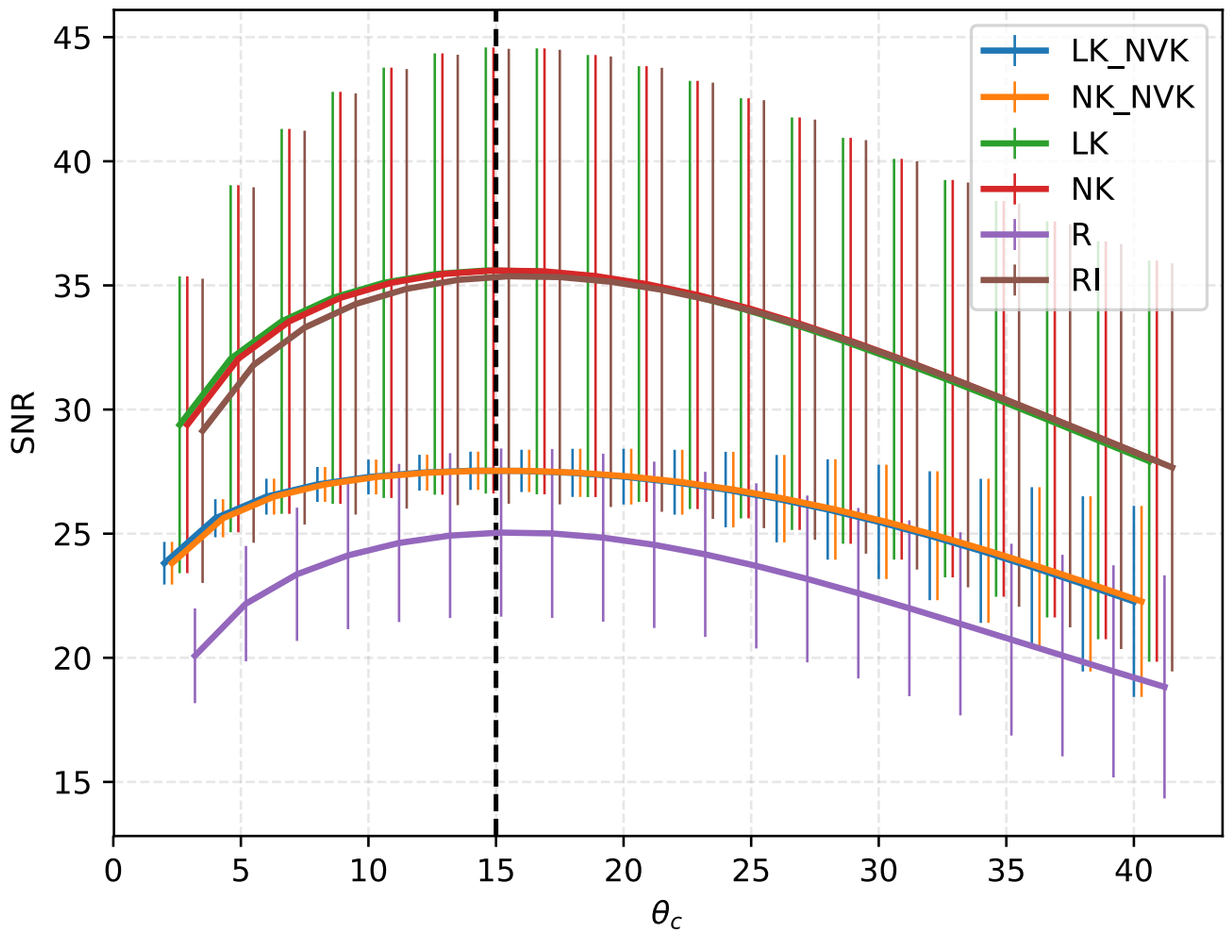
d.o.f —> 0 for AK, 1 for LK, 2 for SK, and 3 in all other cases

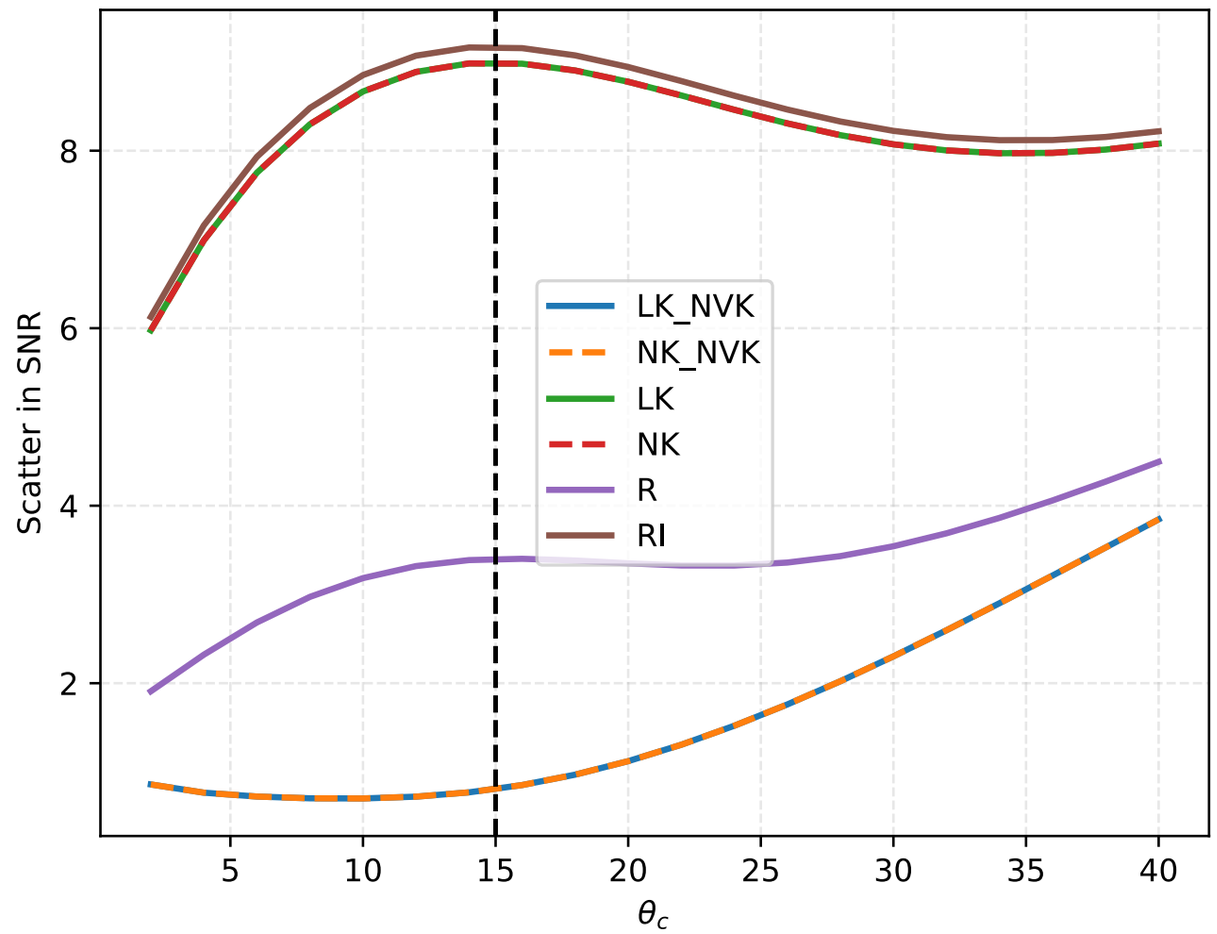
Some details to bear in mind:

- In most likelihood analyses, the noise covariance needs to be measured from data.
- In MF analysis, the case is no different. Note that the noise covariance depends on measurement noise as well as subtle foreground details, which in principle may only be measured once you have the data.

MF behaviour

q~25 cluster

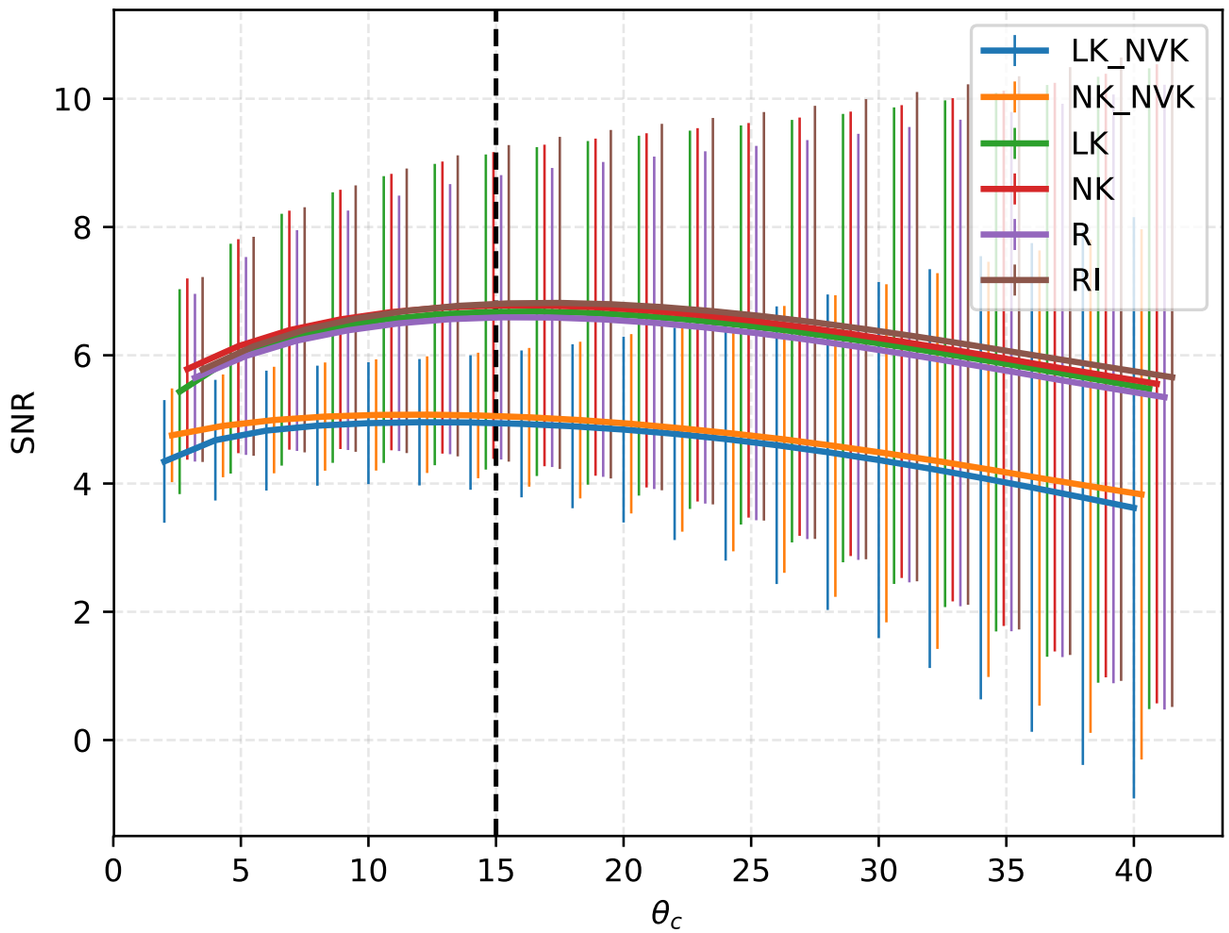


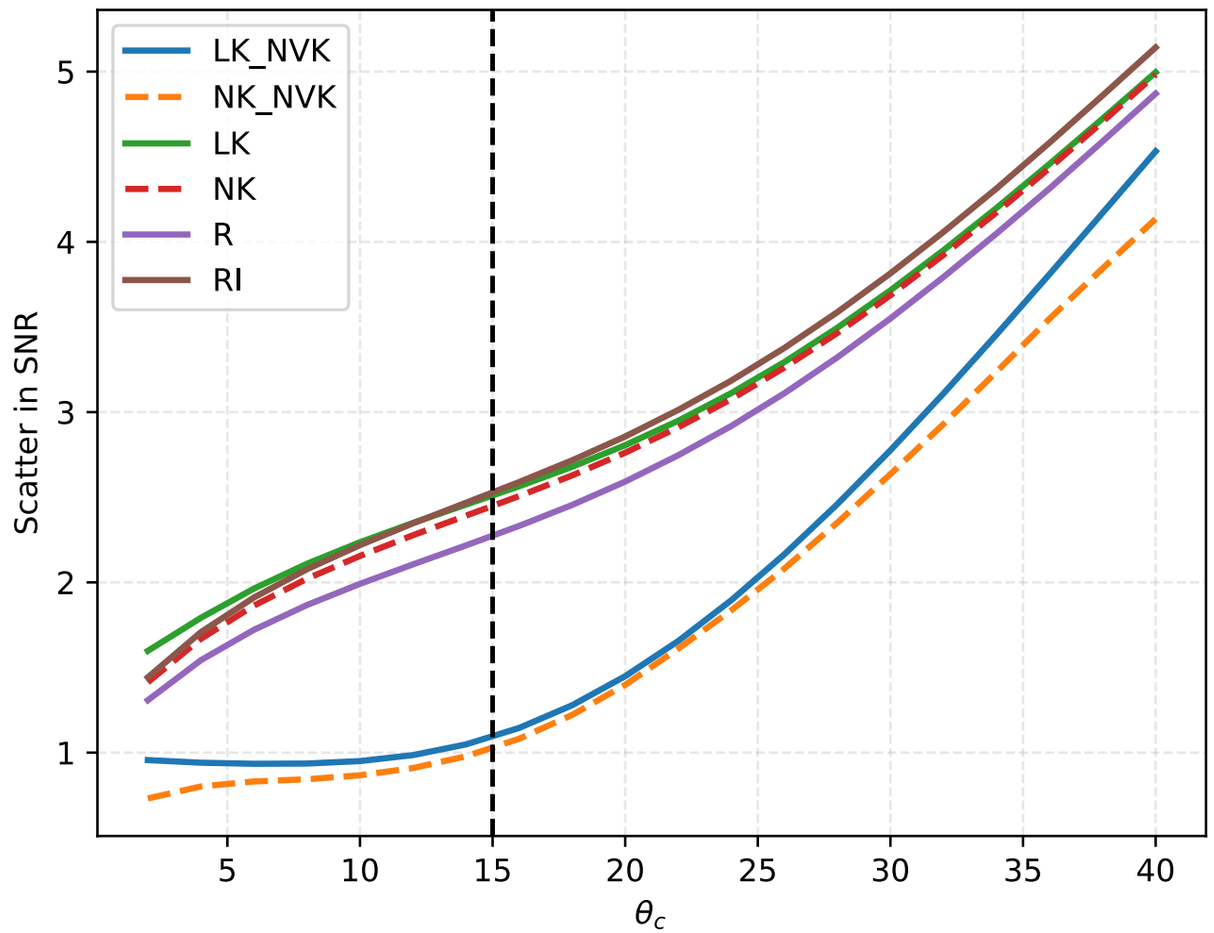


- Assuming noise variance to be known
    1. results in lower scatter in SNR
    2. also results in significantly lower SNR on the mean
- High SNR clusters, can significantly bias the true SNR estimates.
- Iterative procedure yields the true SNR (though current demonstration is for a single cluster in the field. In practice there are likely to be multiple clusters in the field.)

MF behaviour

q~5 cluster





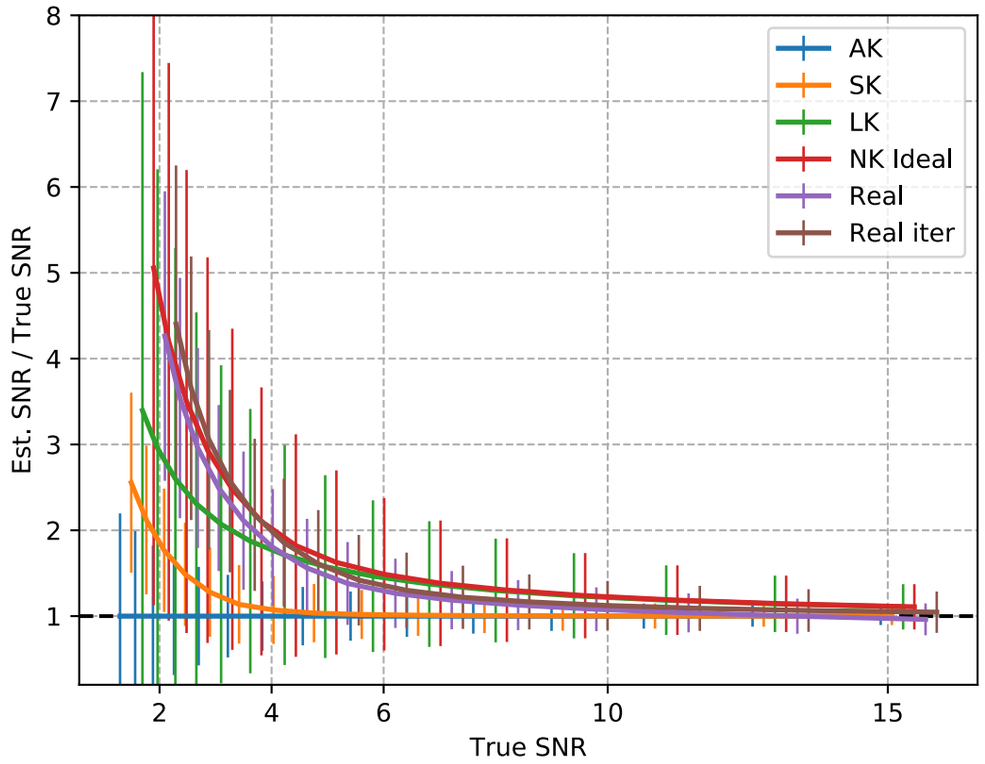
True also for low q clusters

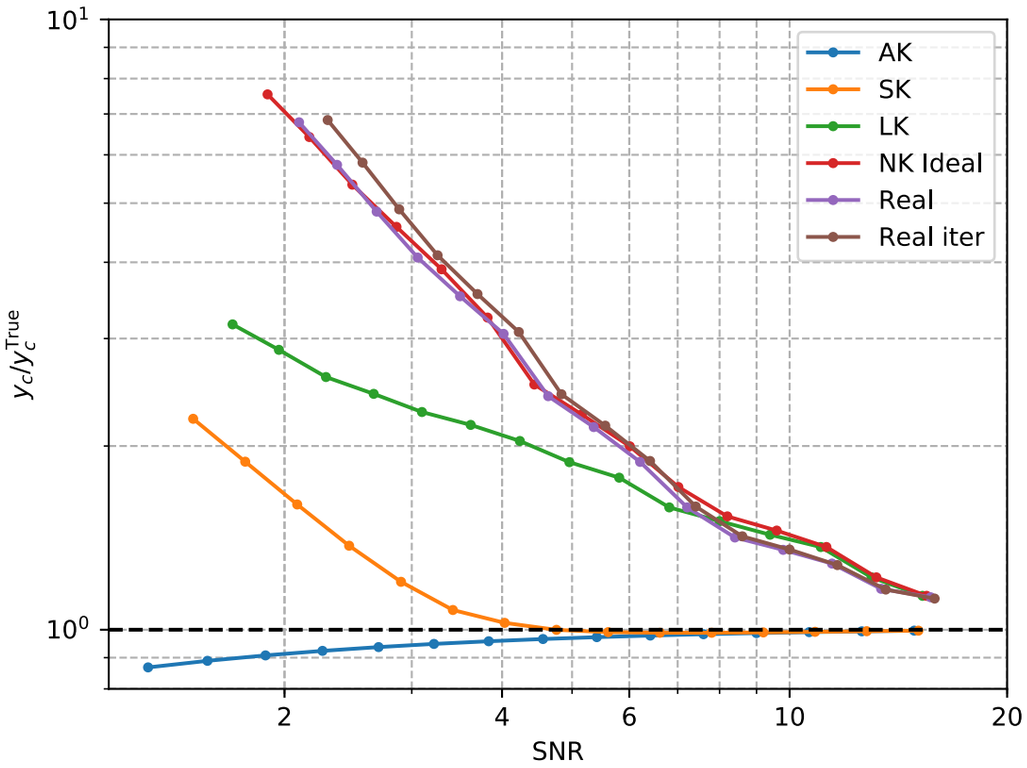
Biases

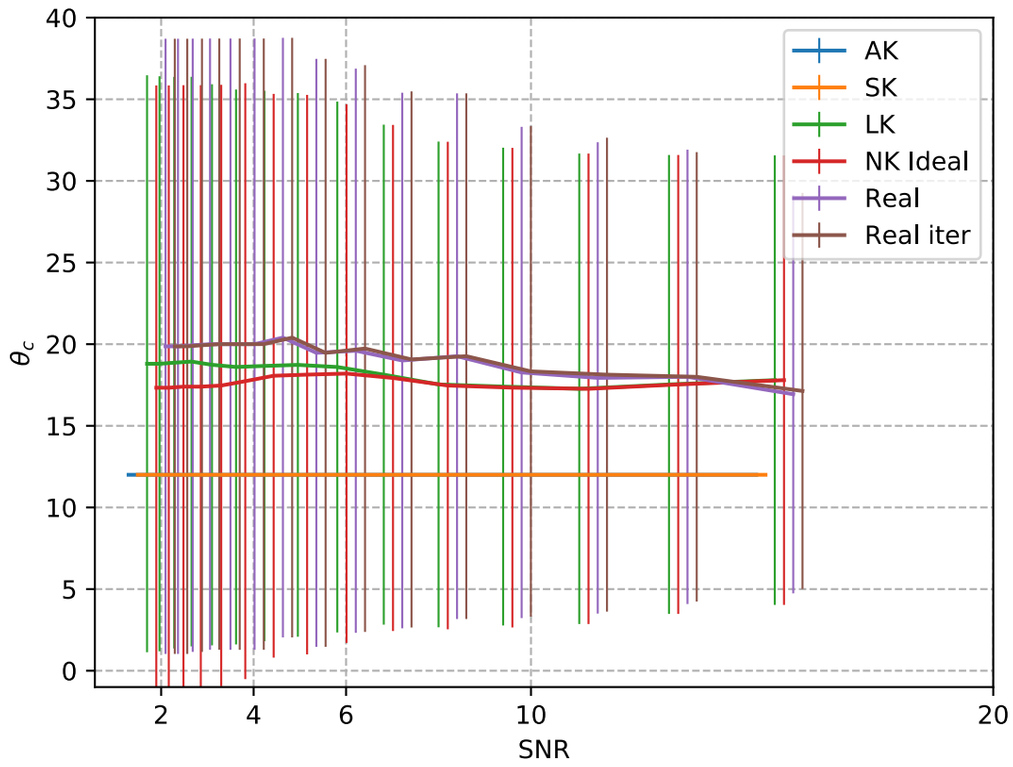
$\mathbf{q}, y_c, \theta_c$

Averaged noise variance known

(relevant only to ideal analysis)



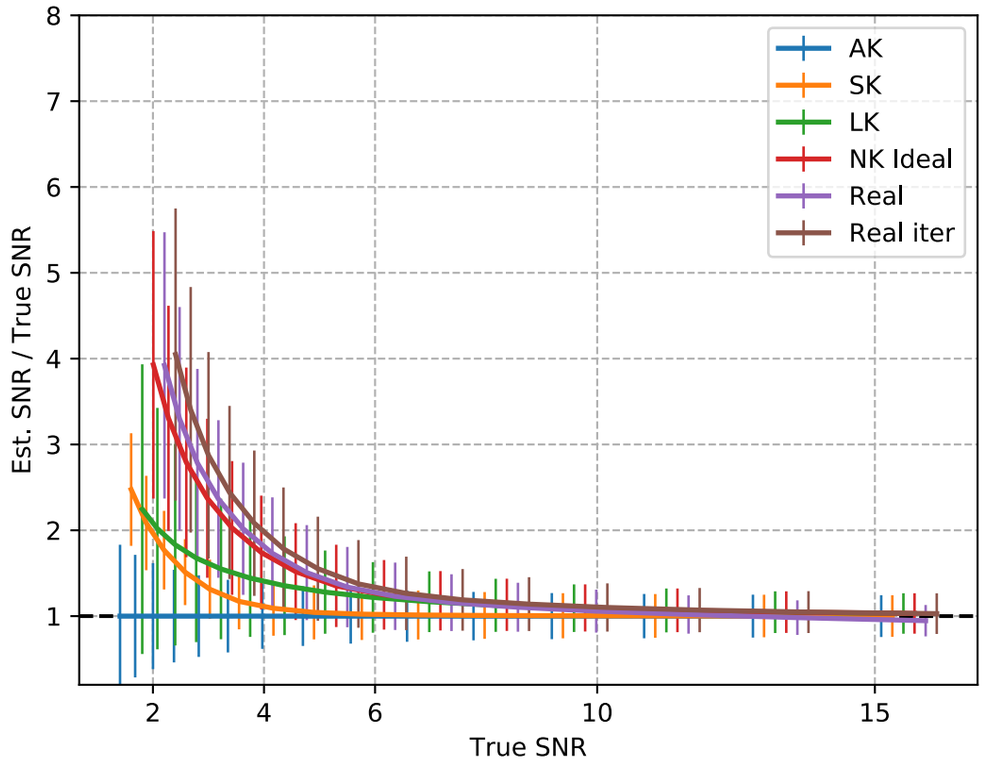


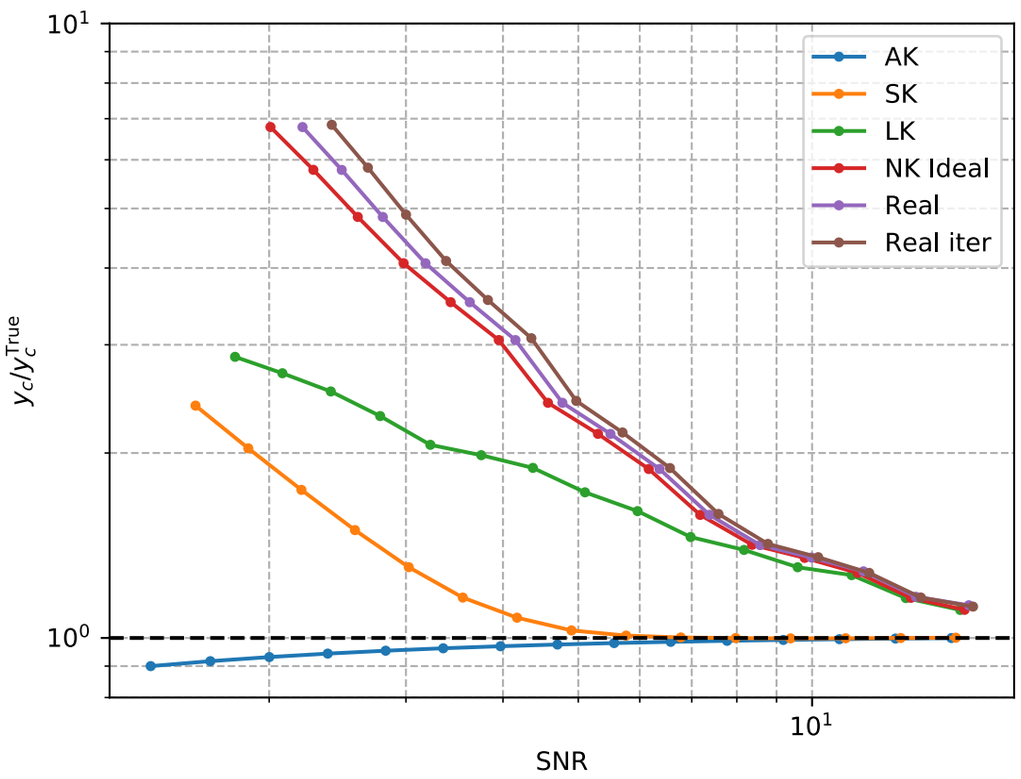


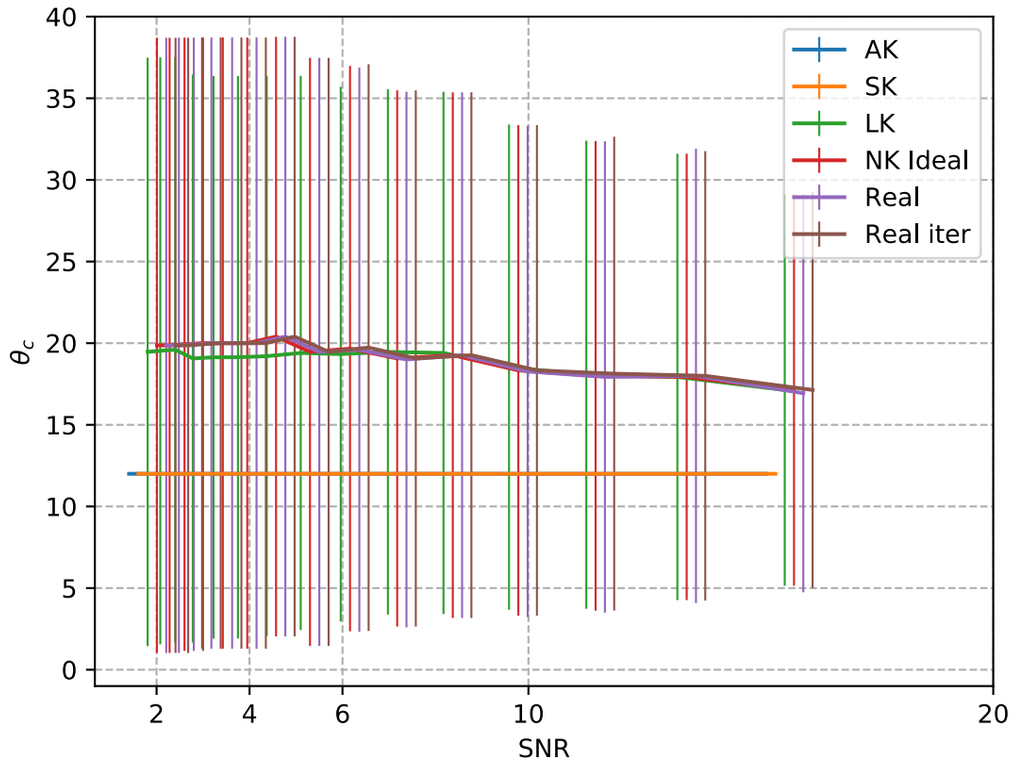
$\mathbf{q}, y_c, \theta_c$

Noise realization known

(relevant only to ideal analysis)







Some comments and queries:

- With LK we expect d.o.f to be smaller I.e 1, and we therefore expect the inconsistency with true SNR to be smaller as compared to SK, where the d.o.f is 2. But I seem to get the opposite trend.
- Knowing size seems to be more important than knowing the location of the cluster. (Careful, I have run this exercise only for a single cluster size)
- The bias on y_c seems to drop much faster with increase in SNR, than the size bias. For this reason y_c might be a better measurable to work with. May that is why ACT analysis work in this variable. (Since θ_c estimates can be significantly skewed, integrated quantities like Y_{500} are more likely to be error prone.)