Density modification and Model building

Some slides are courtesy of…

Kevin Cowtan

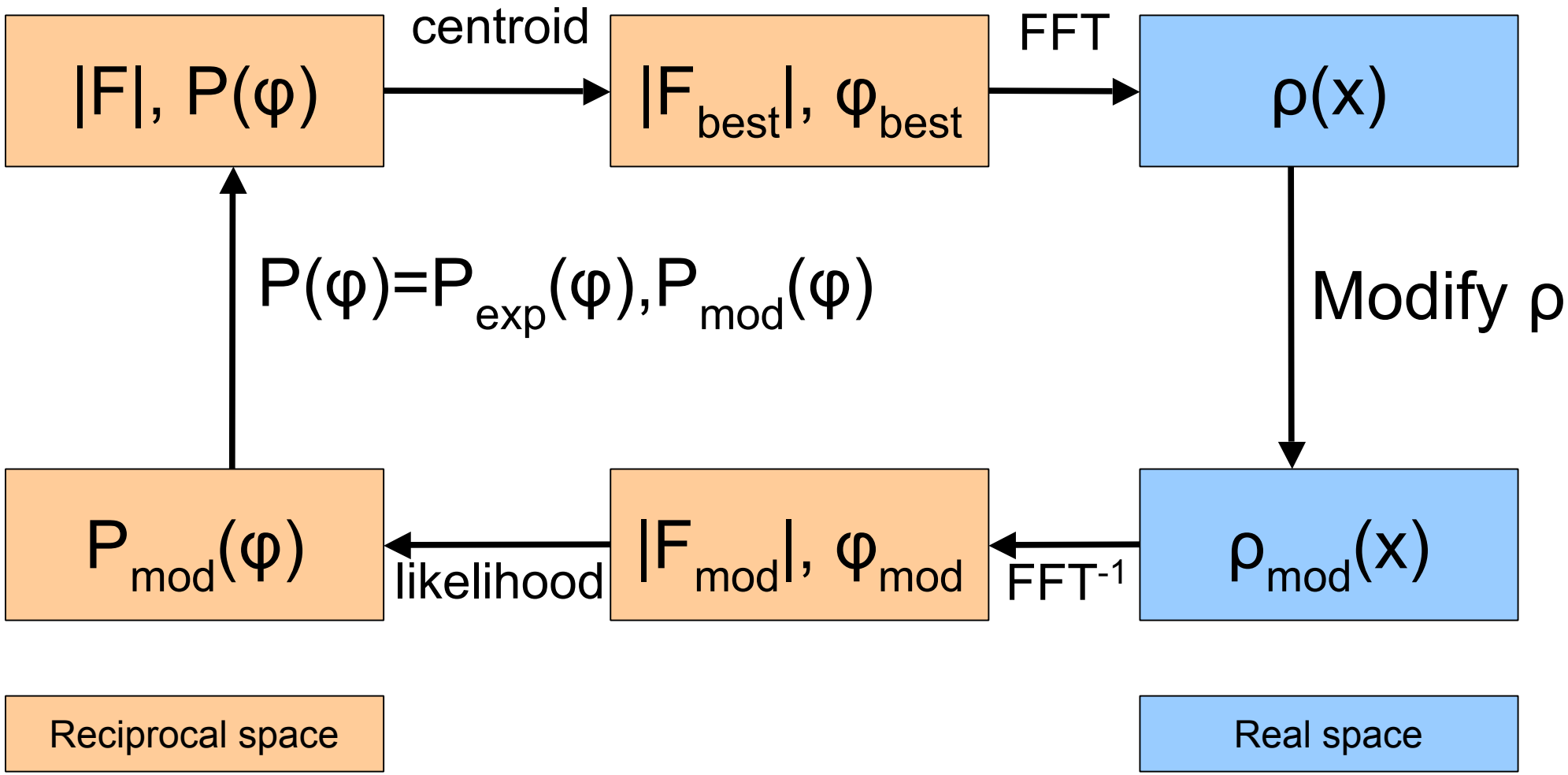
Parrot, Pirate, DM, Buccaneer

Victor Lamzin

ARP/wARP

Recall: Density modification

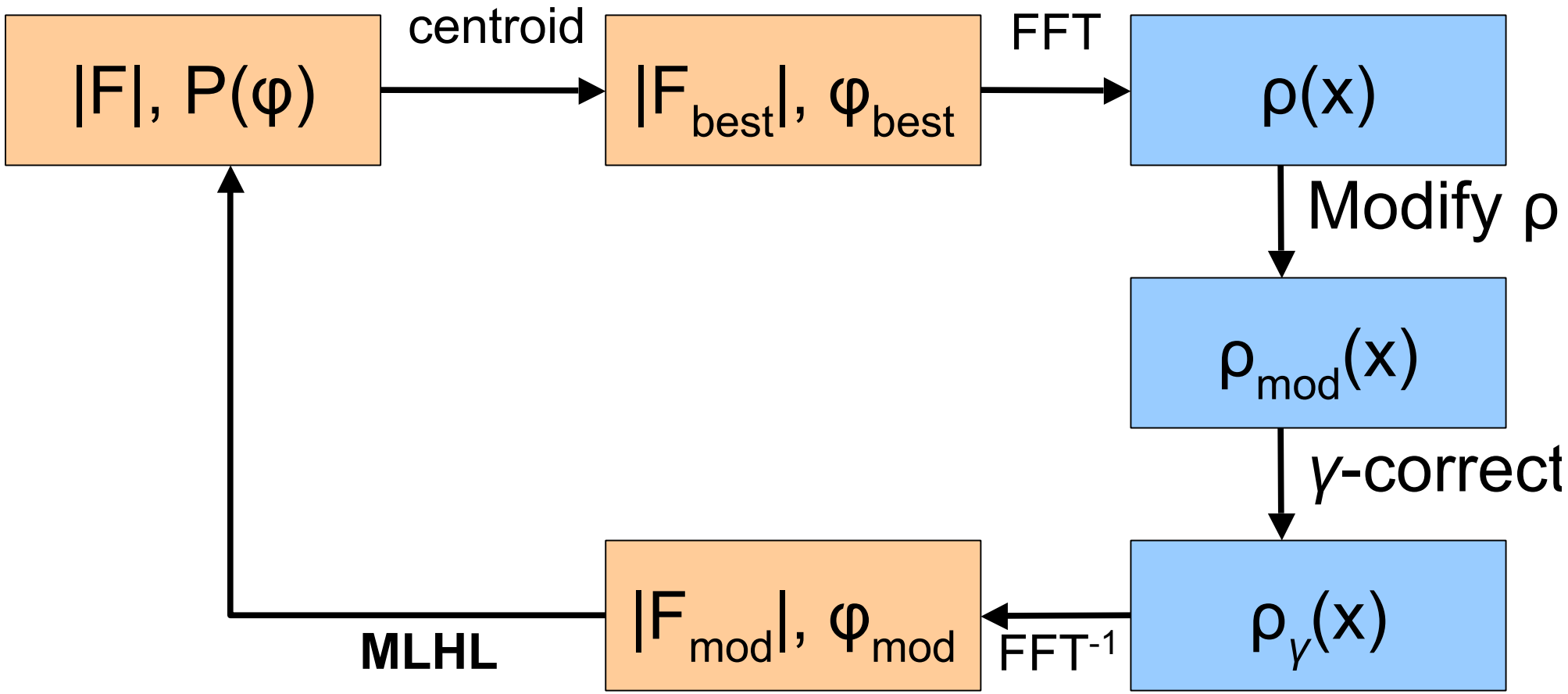
3. Phase probability distributions:



Density modification

PARROT

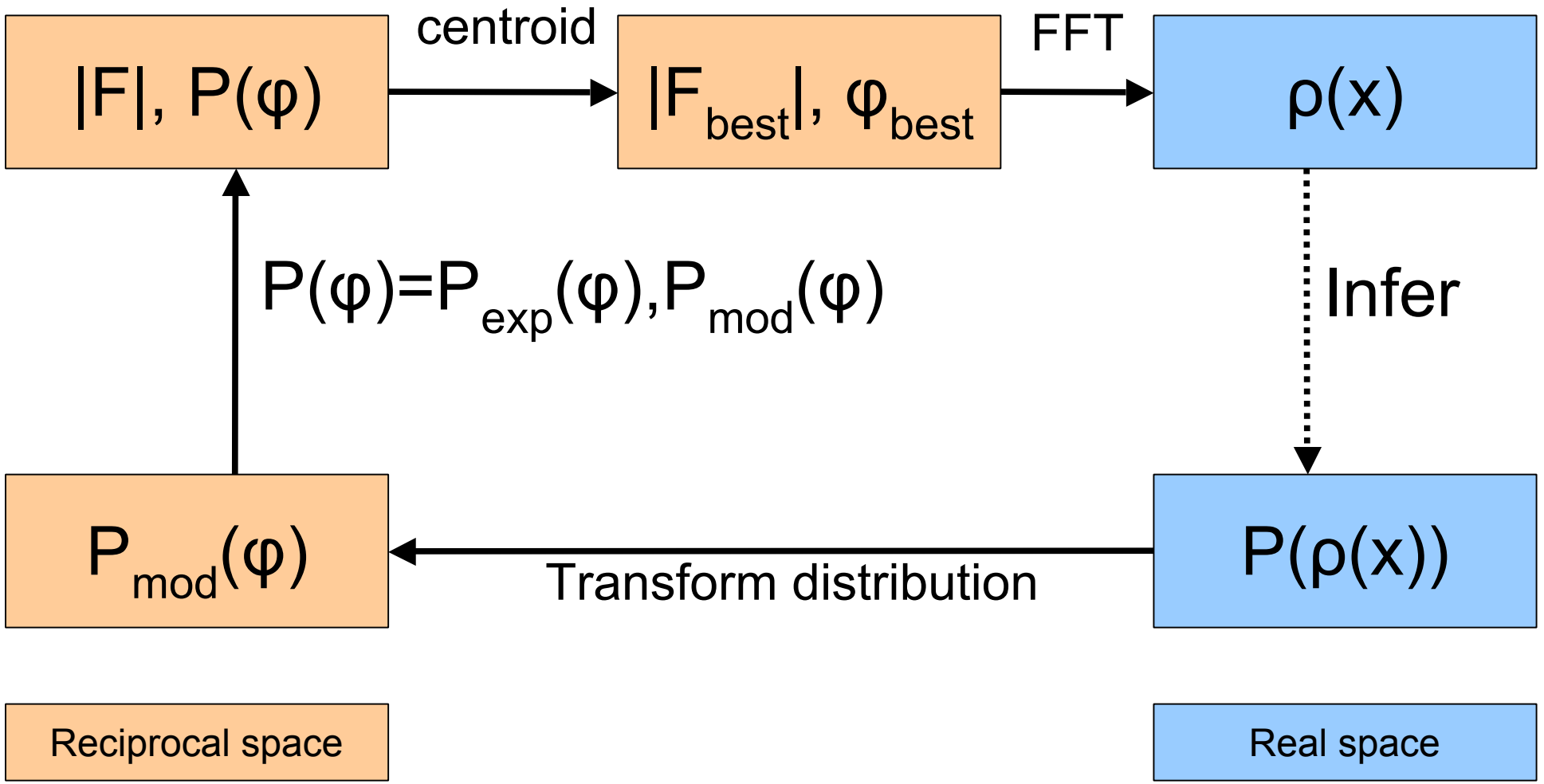
5. Maximum Likelihood H-L:



Density modification

RESOLVE, PIRATE

6. Statistical density modification:



Density modification in Parrot

Builds on existing ideas:

- DM:
  - Solvent flattening
  - Histogram matching
  - NCS averaging
  - Perturbation gamma
- Solomon:
  - Gamma correction
  - Local variance solvent mask
  - Weighted averaging mask

Density modification in Parrot

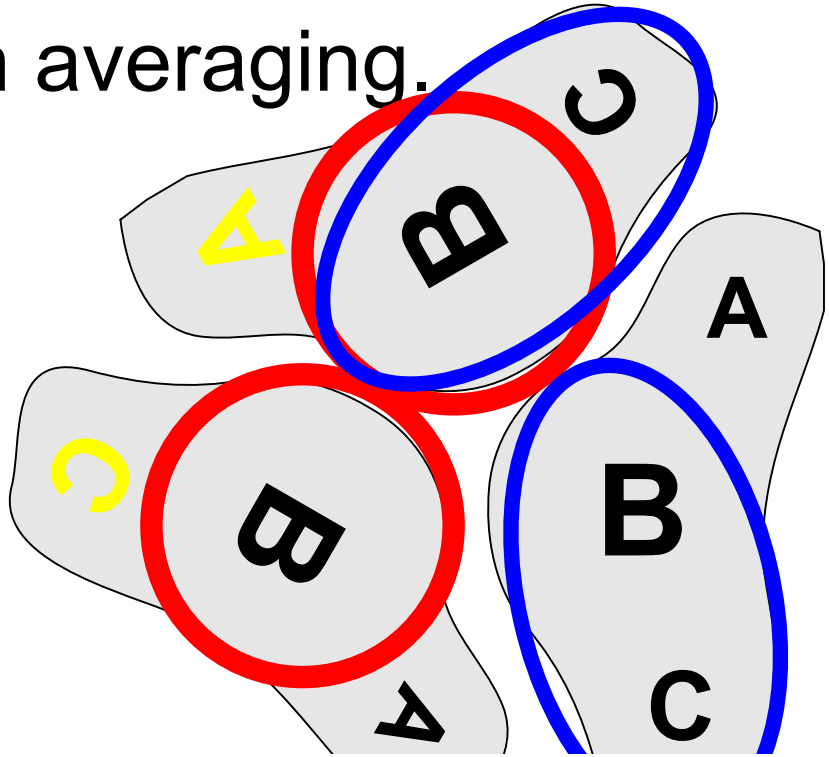
New developments:

- MLHL phase combination
  - (as used in refinement: *refmac, cns*)
- Anisotropy correction
- Problem-specific density histograms
  - (rather than a standard library)
- Pairwise-weighted NCS averaging...

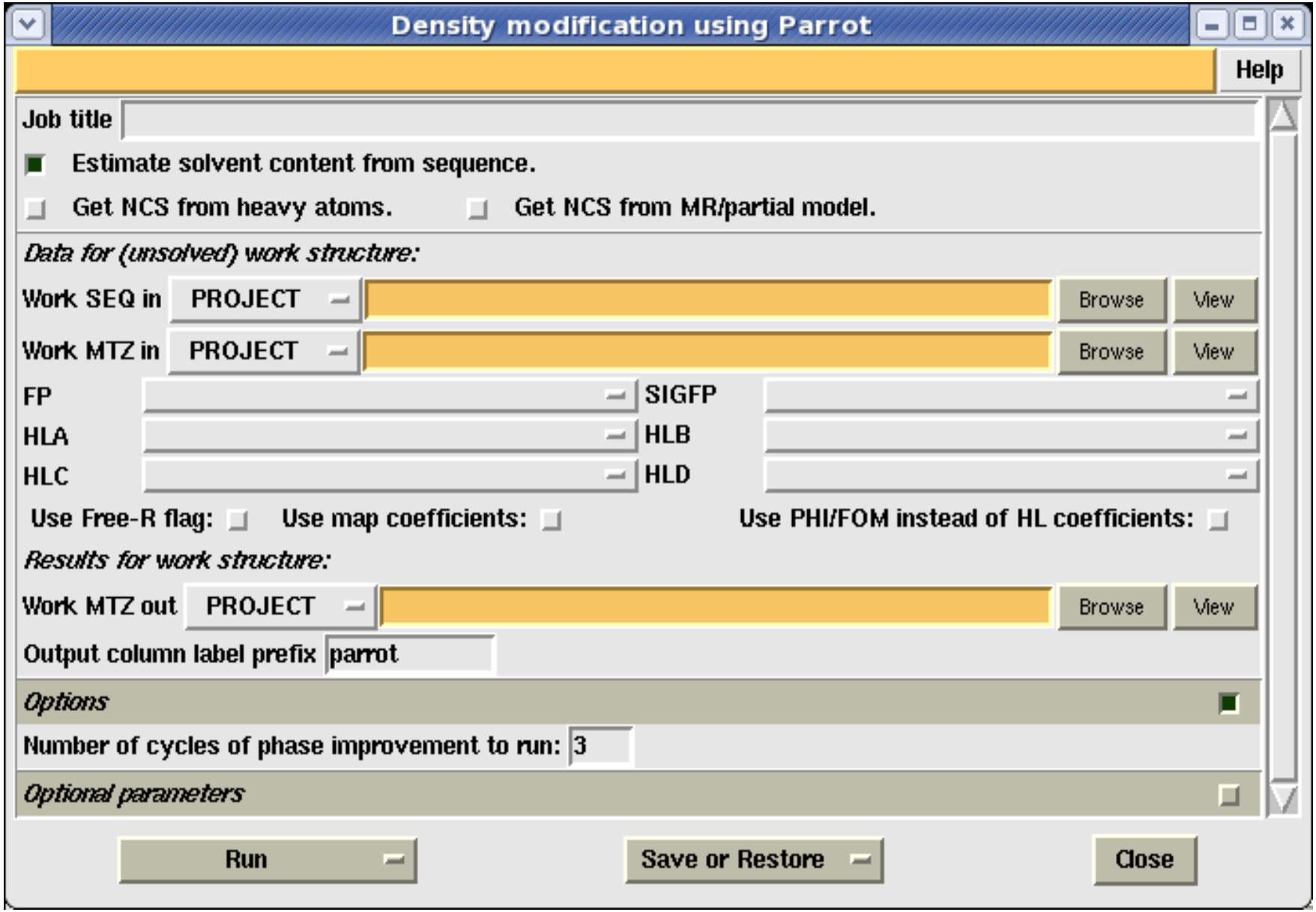
Recent Developments:

Pairwise-weighted NCS averaging:

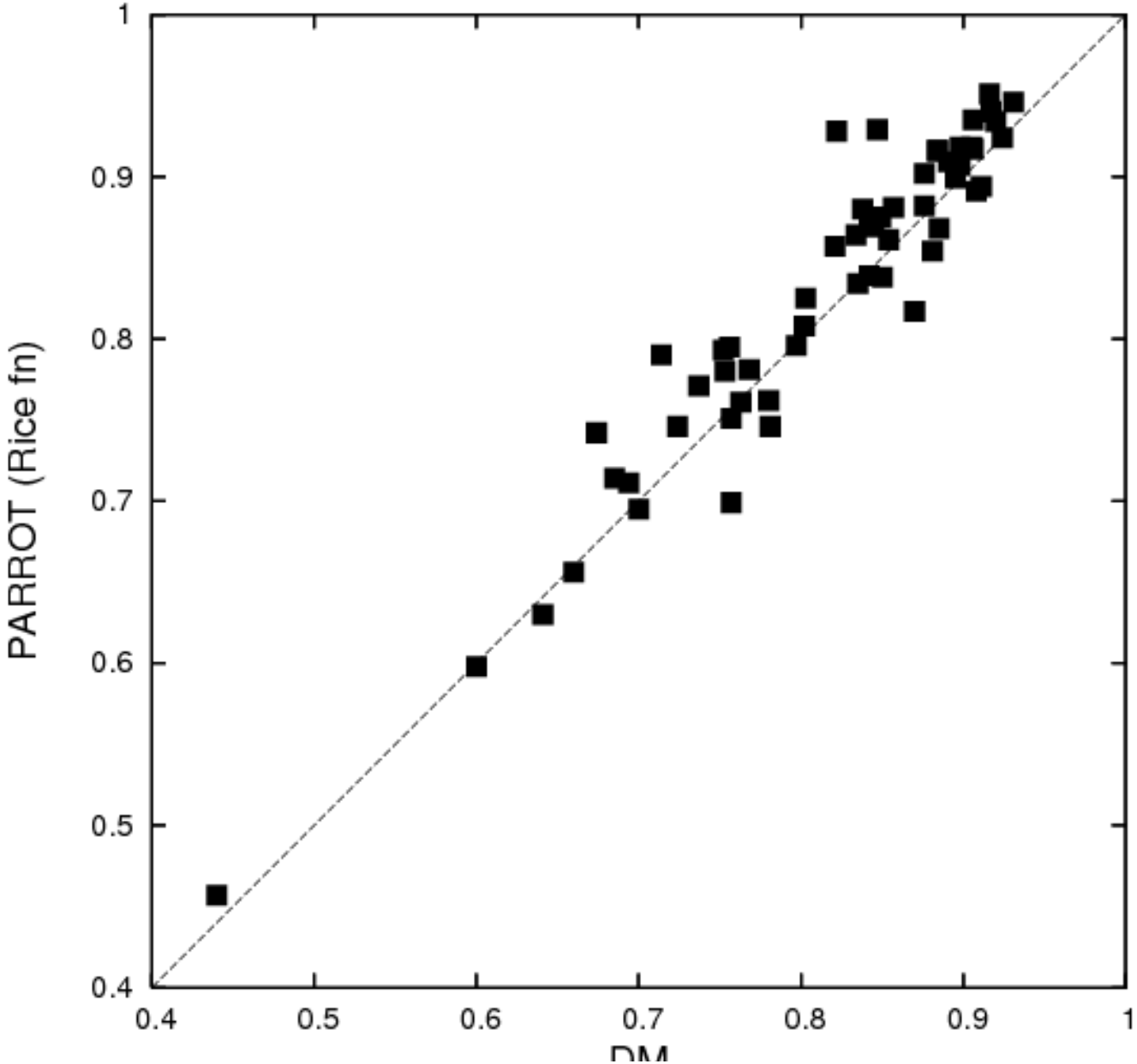
- Average each pair of NCS related molecules separately with its own mask.
- Generalisation and automation of multi-domain averaging.



Parrot


Density modification using Parrot
Help
Job title
Estimate solvent content from sequence.
Get NCS from heavy atoms.
Get NCS from MR/partial model.
Data for (unsolved) work structure:
Work SEQ in
PROJECT
Browse
View
Work MTZ in
PROJECT
Browse
View
FP
SIGFP
HLA
HLB
HLC
HLD
Use Free-R flag:
Use map coefficients:
Use PHI/FOM instead of HL coefficients:
Results for work structure:
Work MTZ out
PROJECT
Browse
View
Output column label prefix
parrot
Options
Number of cycles of phase improvement to run:
3
Optional parameters
Run
Save or Restore
Close

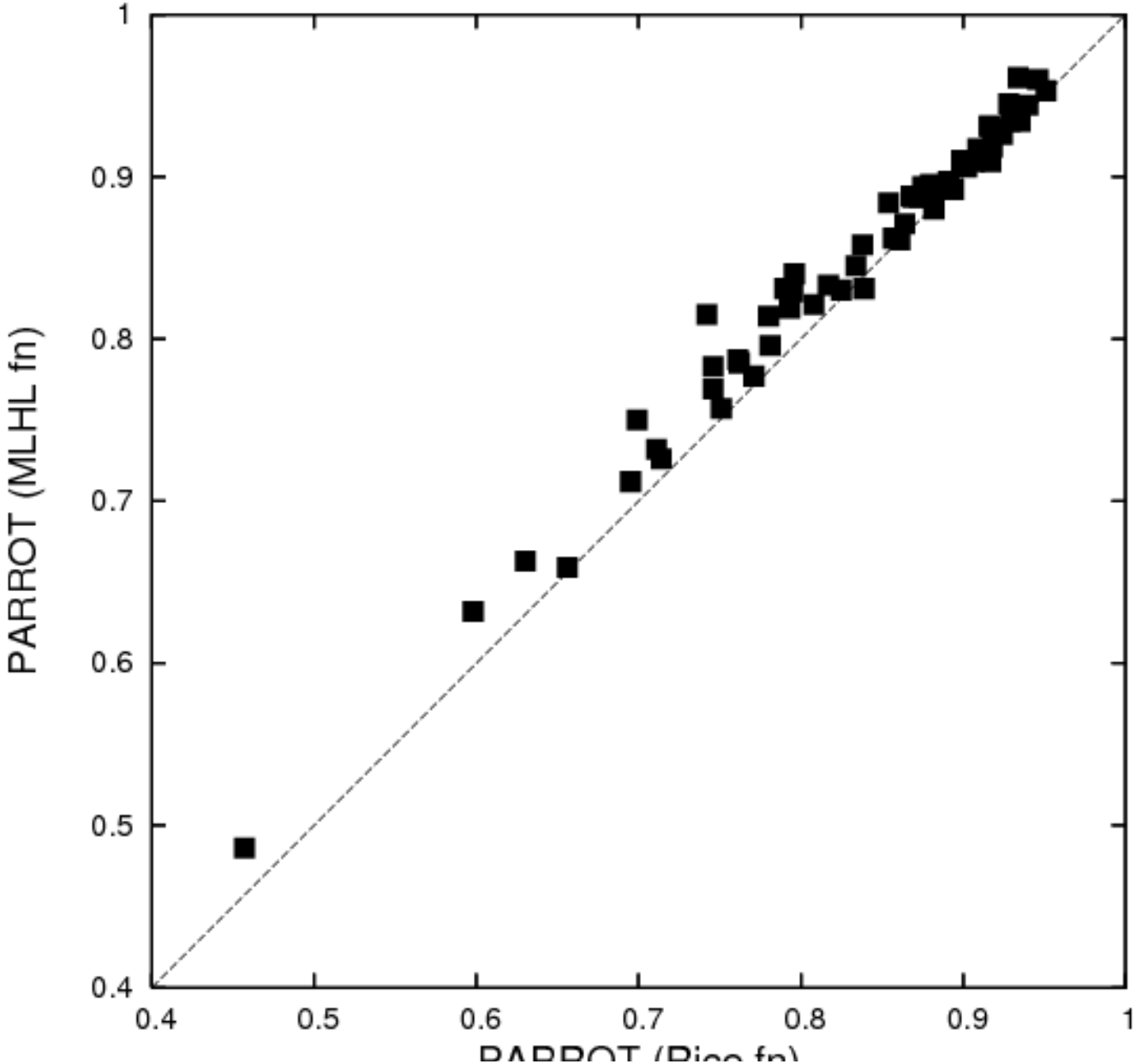
DM vs Parrot



Map correlations

Parrot: No new features enabled.

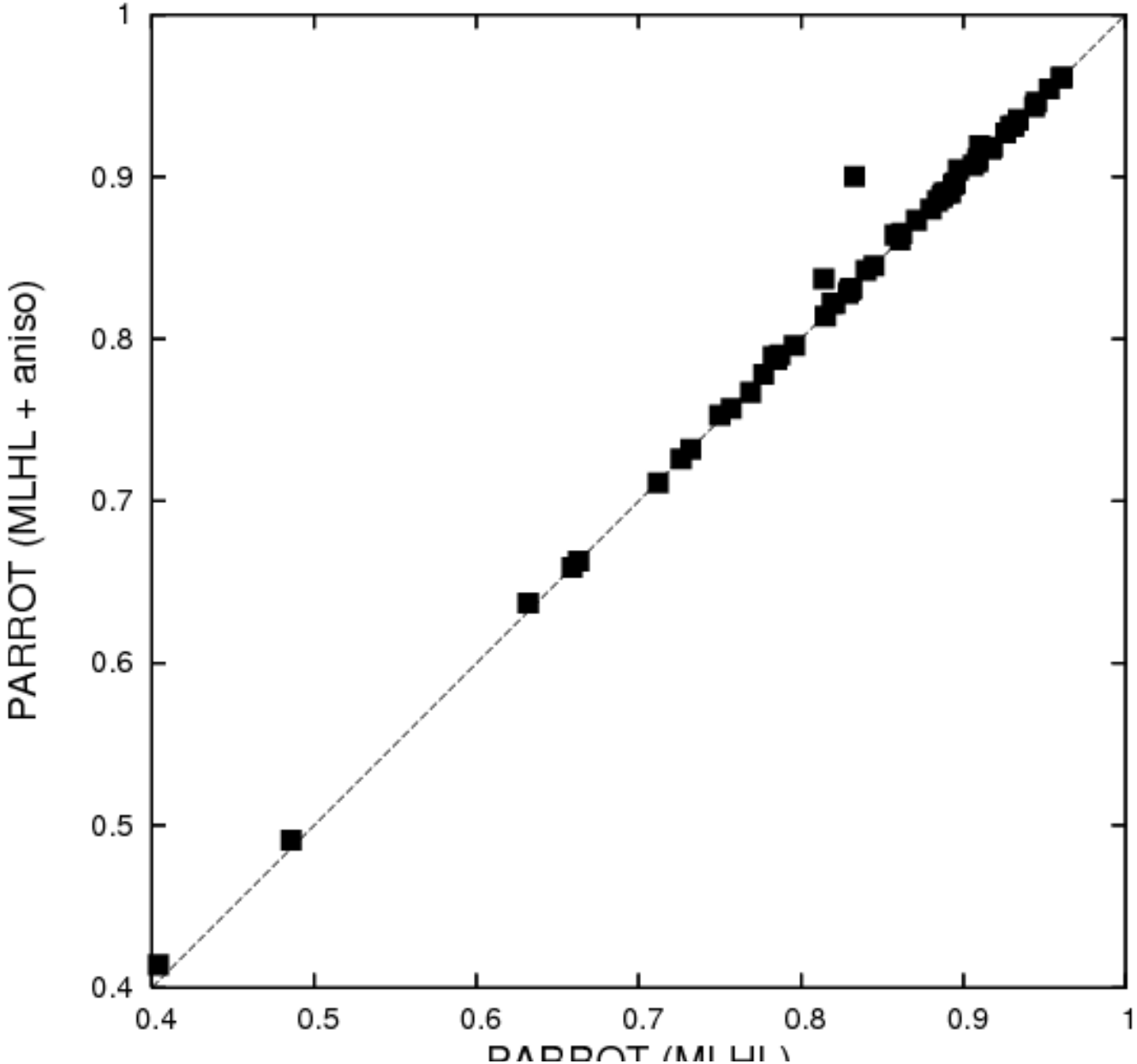
Parrot: Rice vs MLHL



Map correlations

Comparing old and new likelihood functions.

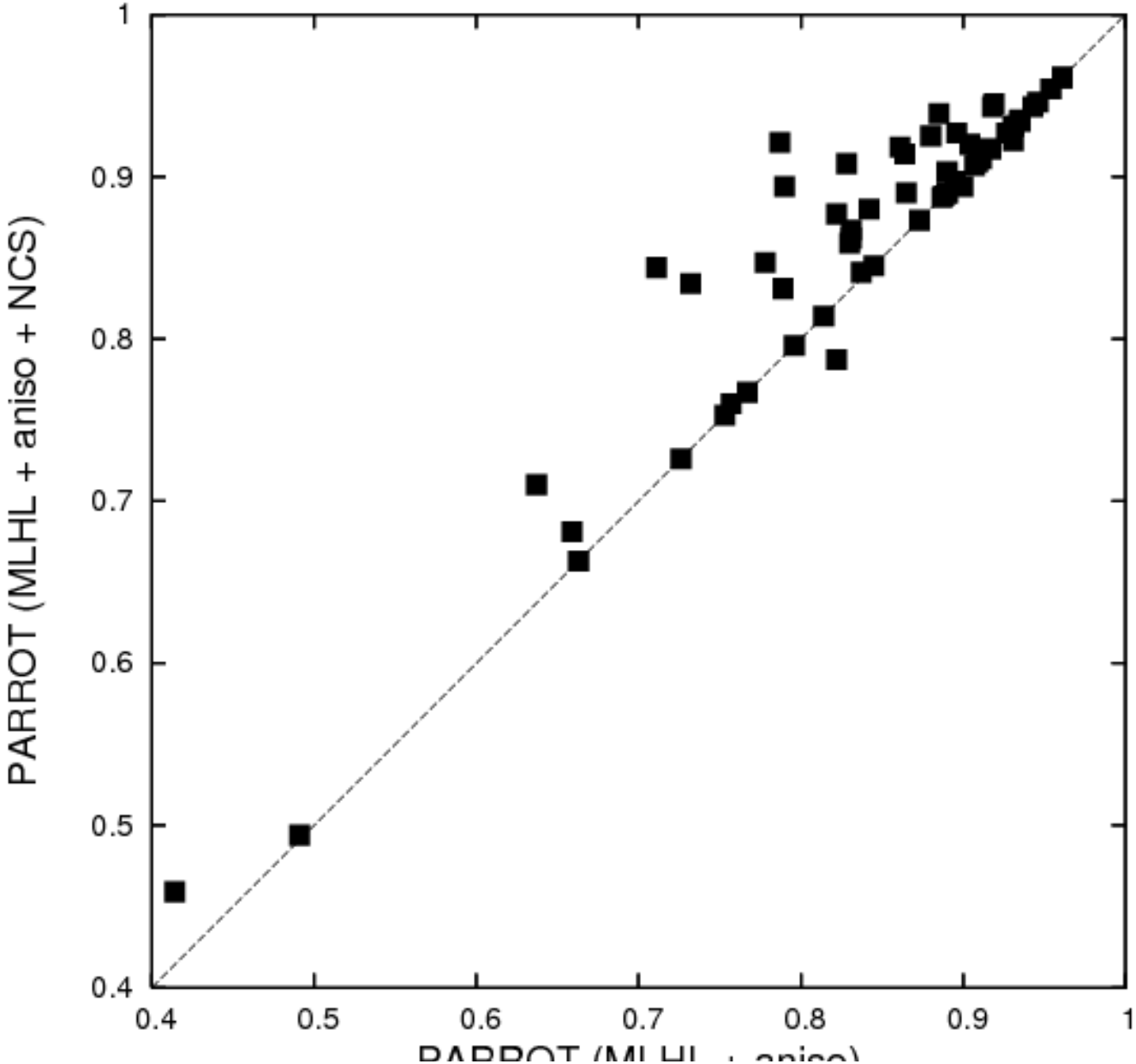
Parrot: Isotropic vs Anisotropic



Map correlations

Comparing with and without anisotropy correction.

Parrot: simple vs NCS averaged



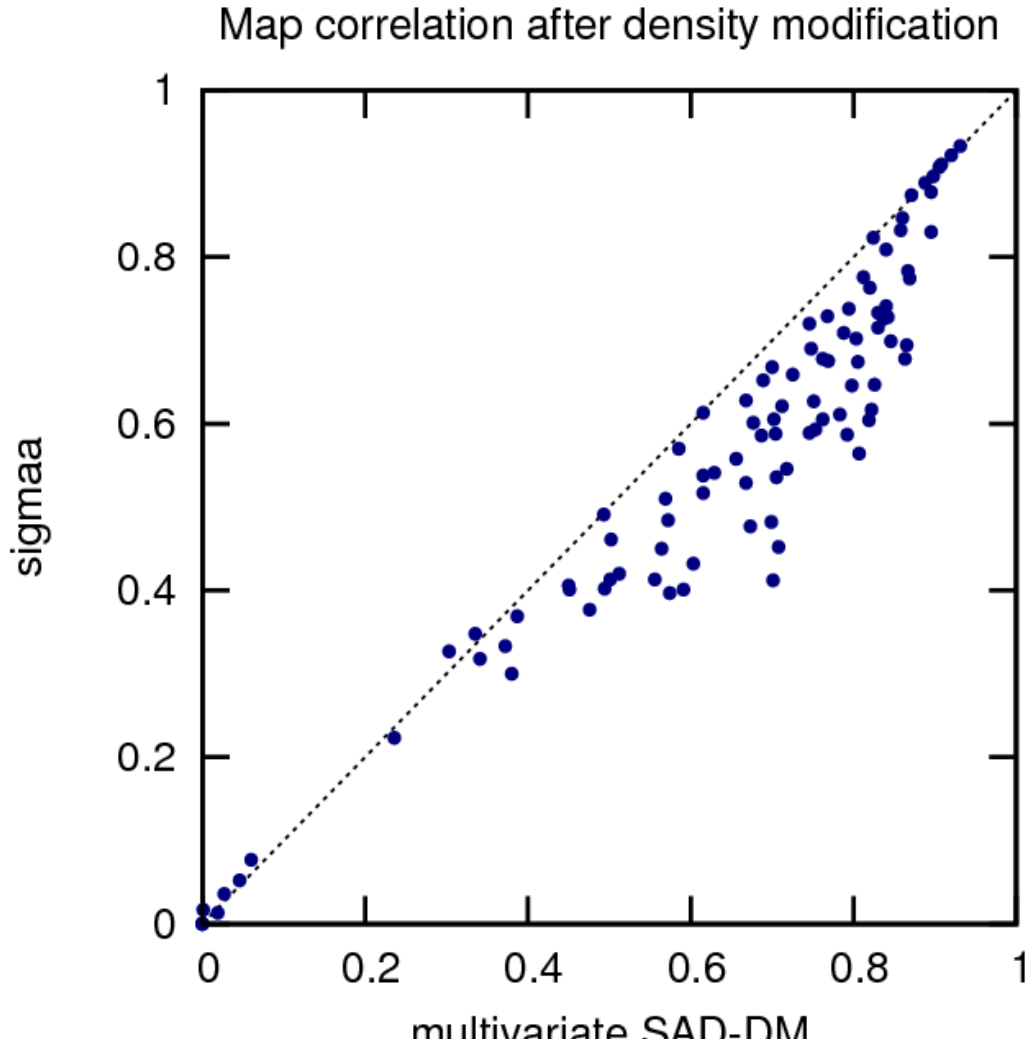
Map correlations

Comparing with and without NCS averaging.

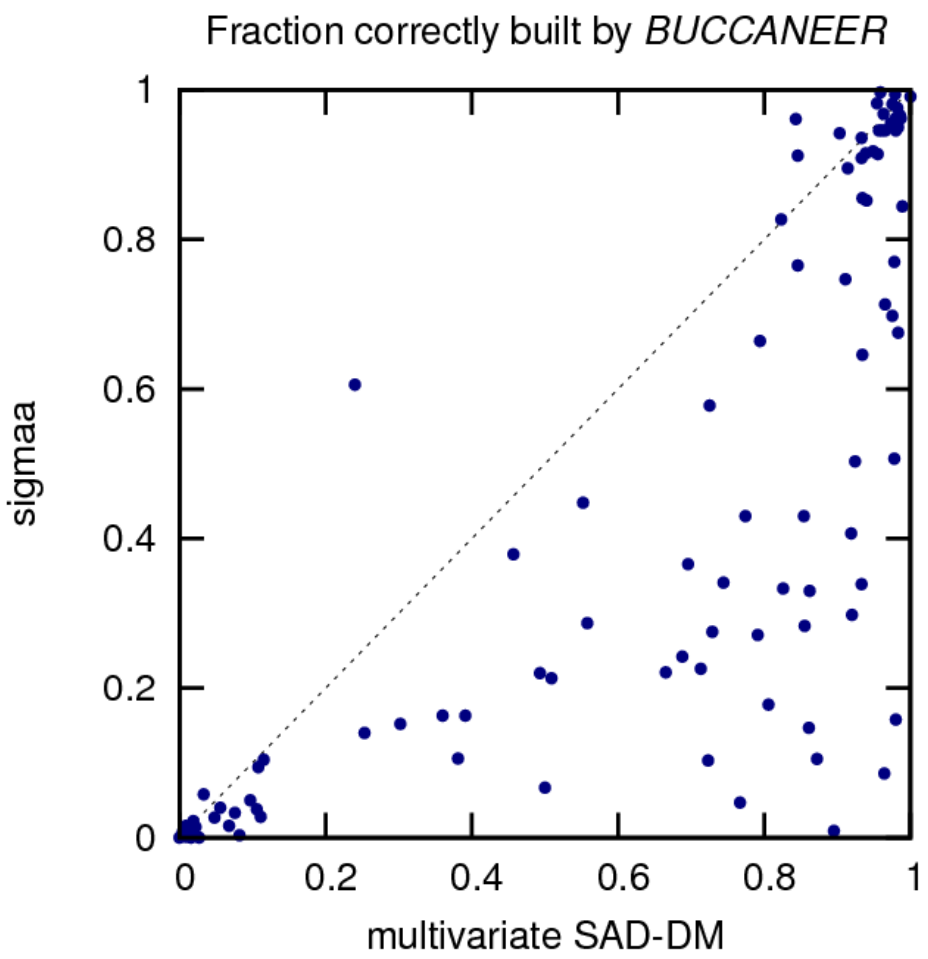
MULTICOMB: Multivariate phase combination for density modification

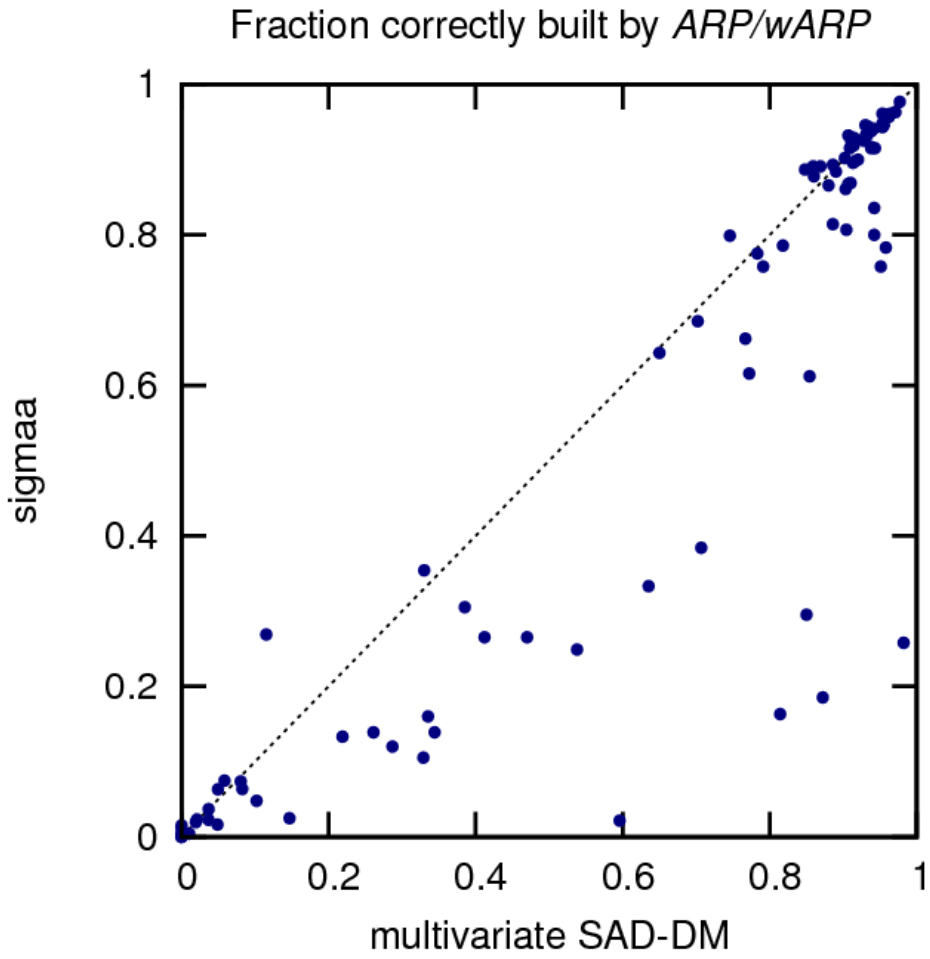
- Current density modification procedures
  - neglect the correlation between the original map and the density modified map.
  - approximate the original phase information with a 1 dimensional Hendrickson-Lattman distribution
- To overcome these shortcomings, we implemented a multivariate function which explicitly takes into account the correlation between the original, density modified and heavy atom structure factors.

Comparison of sigmaa vs. multivariate SAD function



Results of model building: sigmaa vs. multivariate SAD

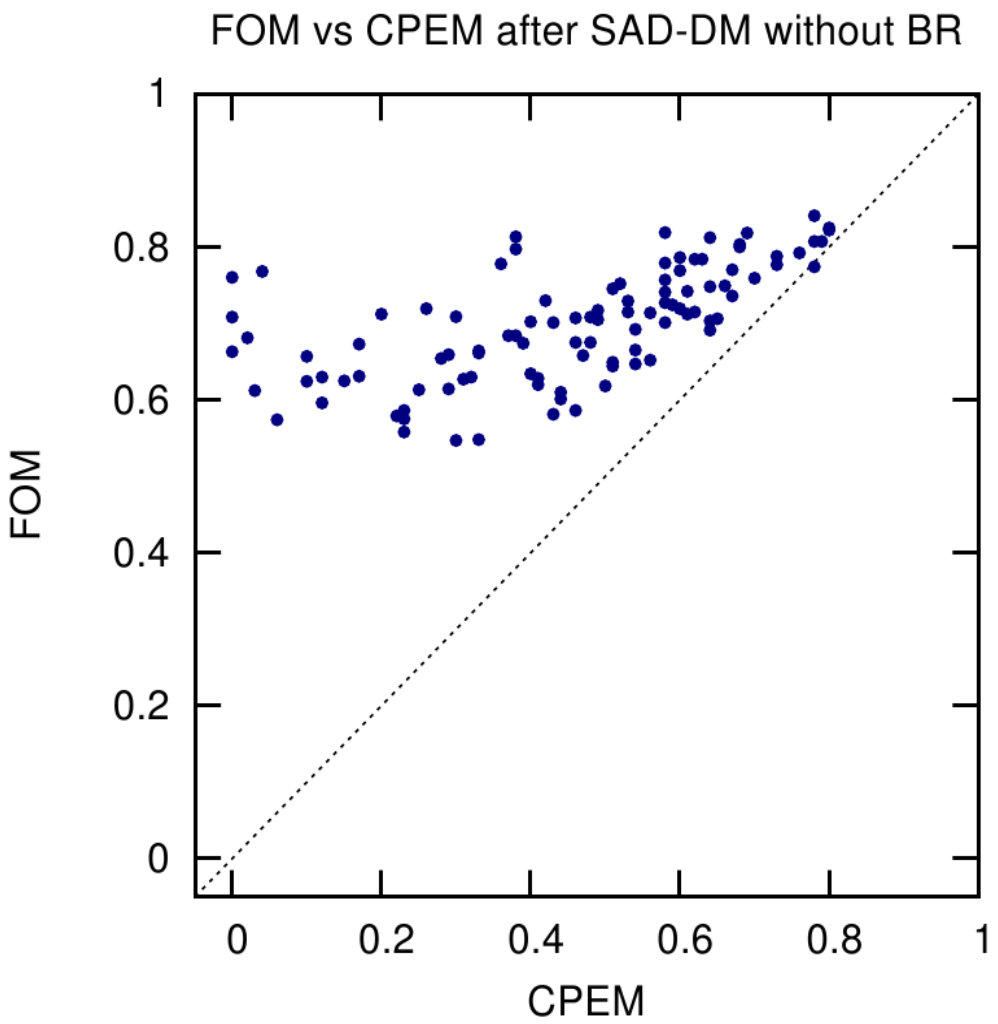


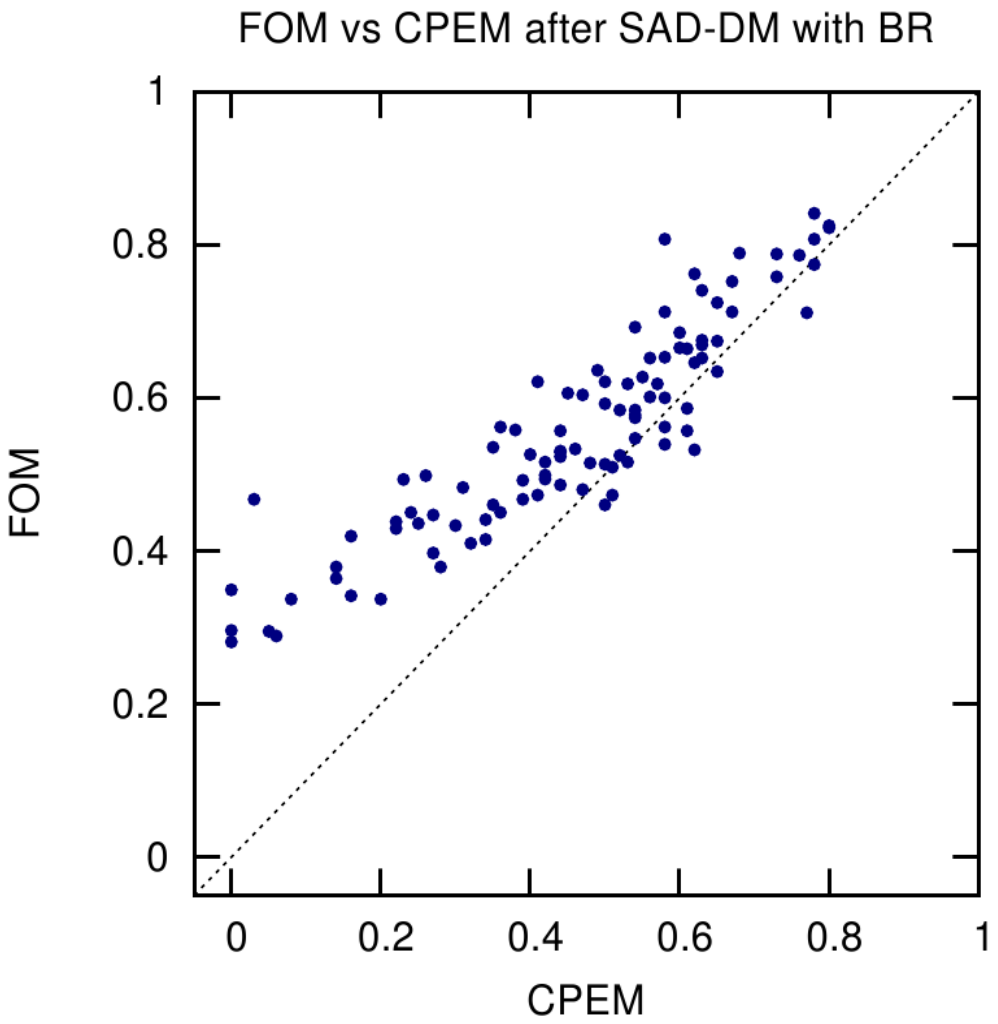


β-correction method: bias reduction in density modification

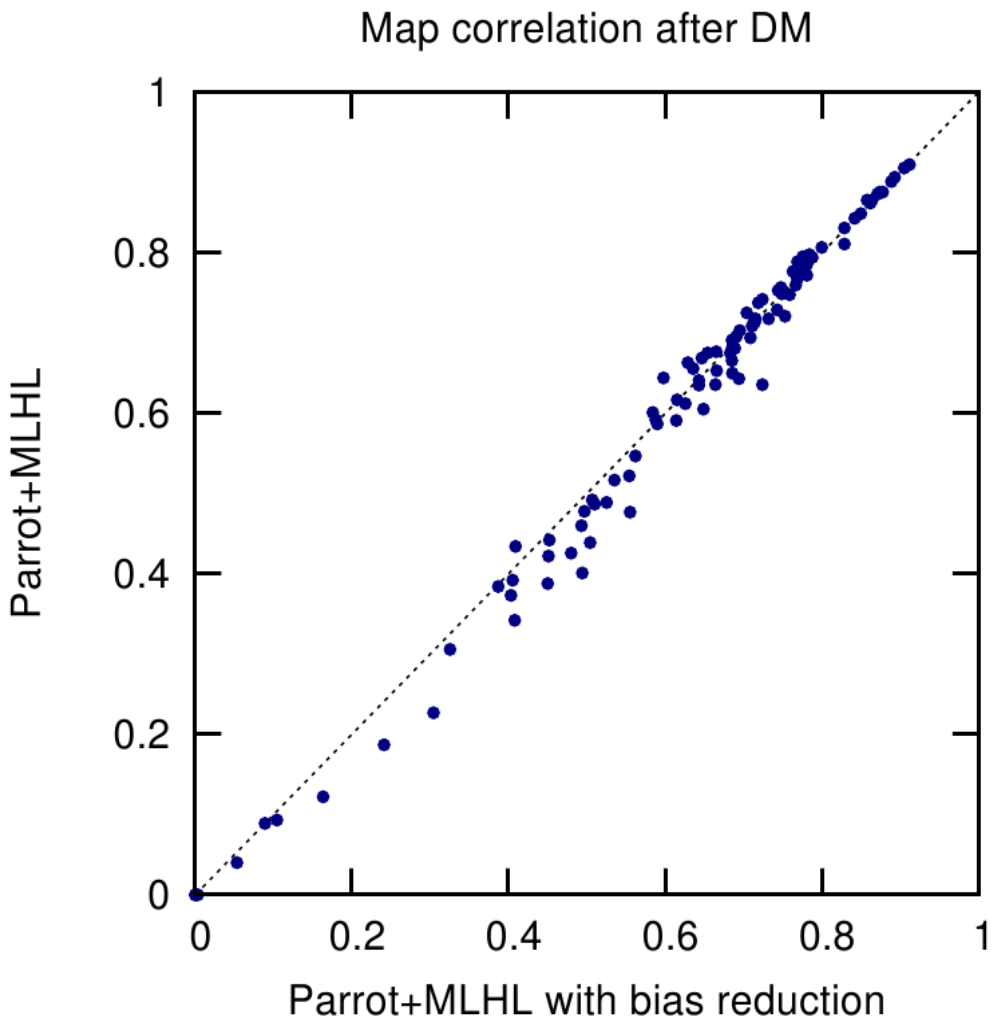
- Density modified map is obtained from experimental map leading to artificially high correlations between the observed and modified amplitudes.
- β correction is applied to the Luzzati error parameter to reduce bias of modified data.

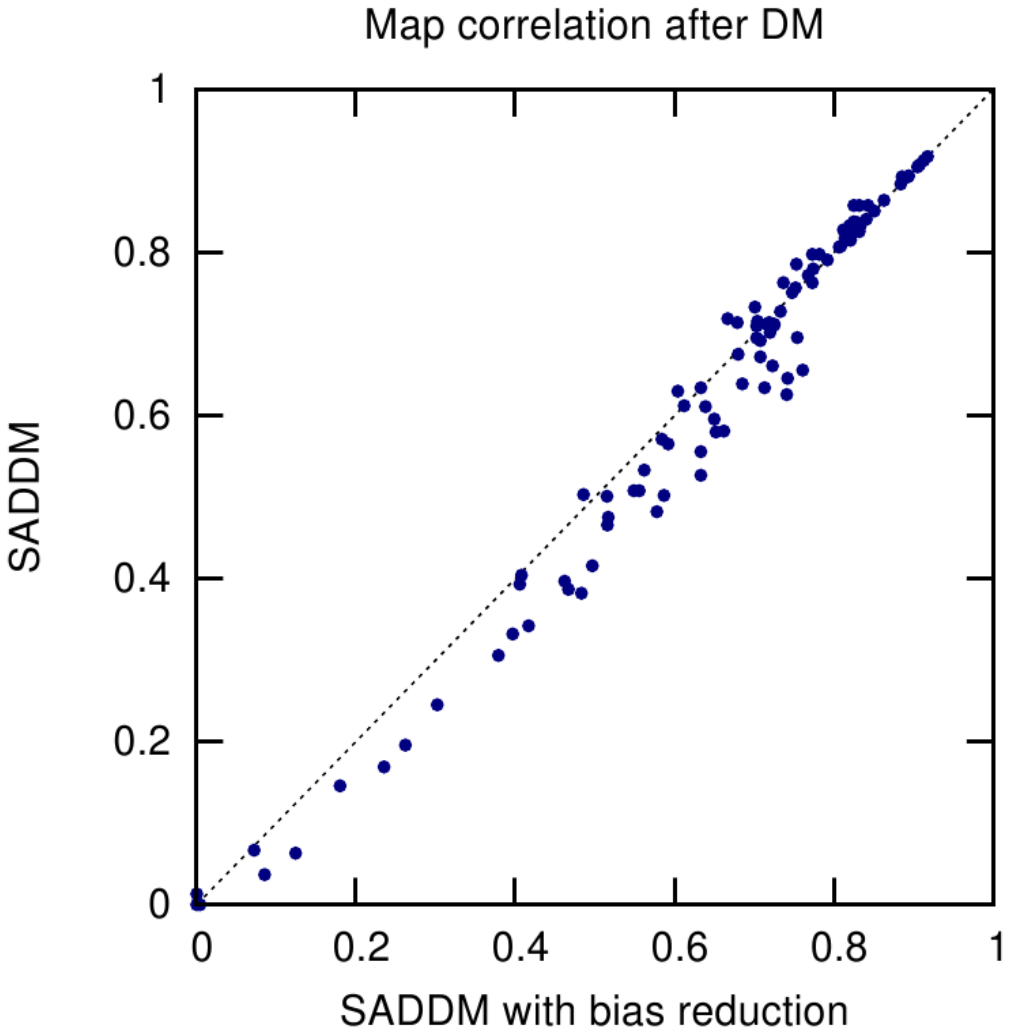
FOM and phase error after DM with/without bias reduction





Map correlations after DM with/without bias reduction





SAD and SIRAS functions in model refinement

- Previous functions in REFMAC:
  - No prior phase information (Rice function) (Murshudov *et al.*,1997), (Bricogne and Irwin, 1996), (Pannu and Read, 1996)
  - Prior phase information used indirectly in the form of Hendrickson-Lattman coefficients (MLHL function) (Pannu *et al.*, 1998)

Features of MLHL function

- Dependent on where you obtained your Hendrickson-Lattman coefficients.
- Assumes that your prior phase information is independent from your model phases!
- ***Benefit***: General approach for all experiments (MAD, SAD, MIRAS).

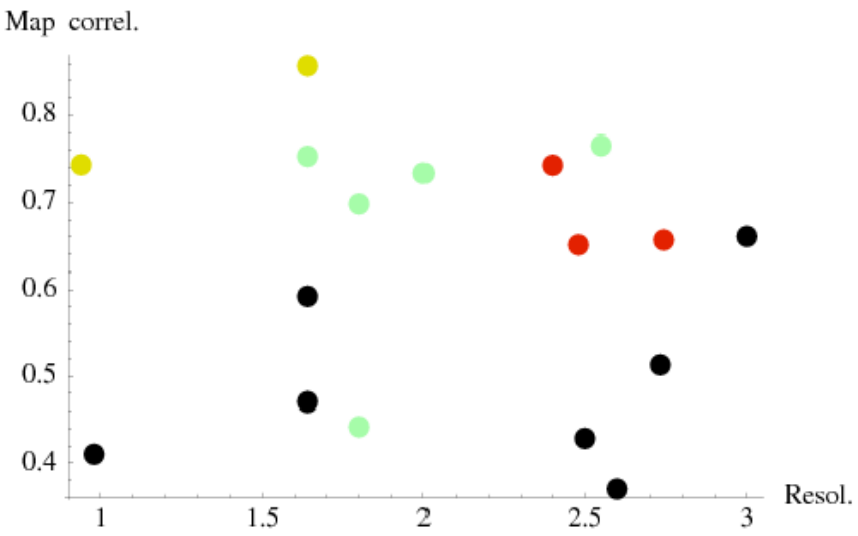
Multivariate SIRAS function for phasing and model refinement

- Currently in BP3 and SHARP, anomalous information is added for SIRAS and MAD by multiplying by a Gaussian term of Bijvoet differences (Thus, assuming independence with isomorphism term.)
- This isomorphic term also assumes uncorrelated errors.
- Better results are obtained by deriving a multivariate function for SIRAS modeling the correlation amongst data sets (Skubak et al. (2009) Acta D).

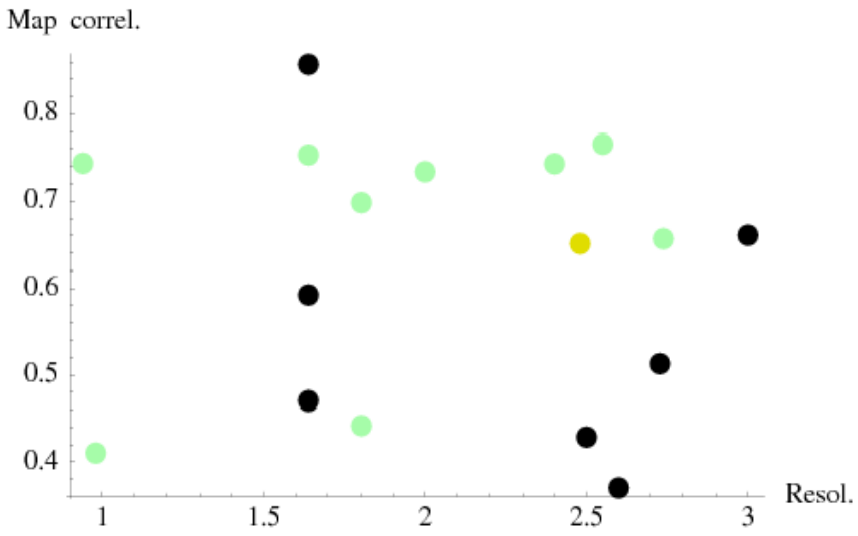
Tests of SAD and SIRAS functions in refinement

- The functions were tested on many real data sets (various phasing signals and resolution ranges) in ARP/wARP + REFMAC.
- Input created by CRANK using CRUNCH2 or SHELXD, BP3 and DM or SOLOMON.
- Skubak *et al.* (2004,2005,2009) Acta D.

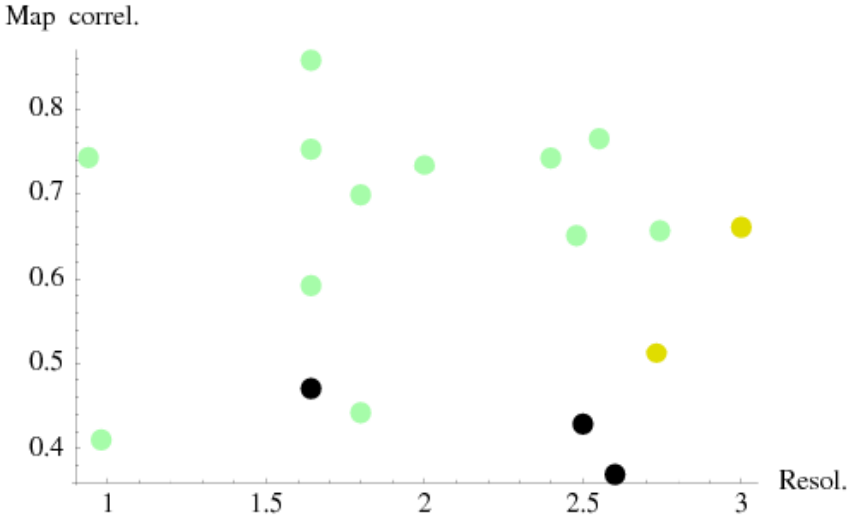
Results from SAD function



Rice function



MLHL function



SAD function

Green: 80 – 100% built

Yellow: 50 – 80% built

Red: 20 – 50% built

Black: 0 - 20% built