Introduction to density modification and some latest density modification improvements

Content

- Density modification principles and problems
- Density modification by Parrot
- Multivariate phase combination
- Bias reduction by β-correction
- Combining density modification with model building and refinement

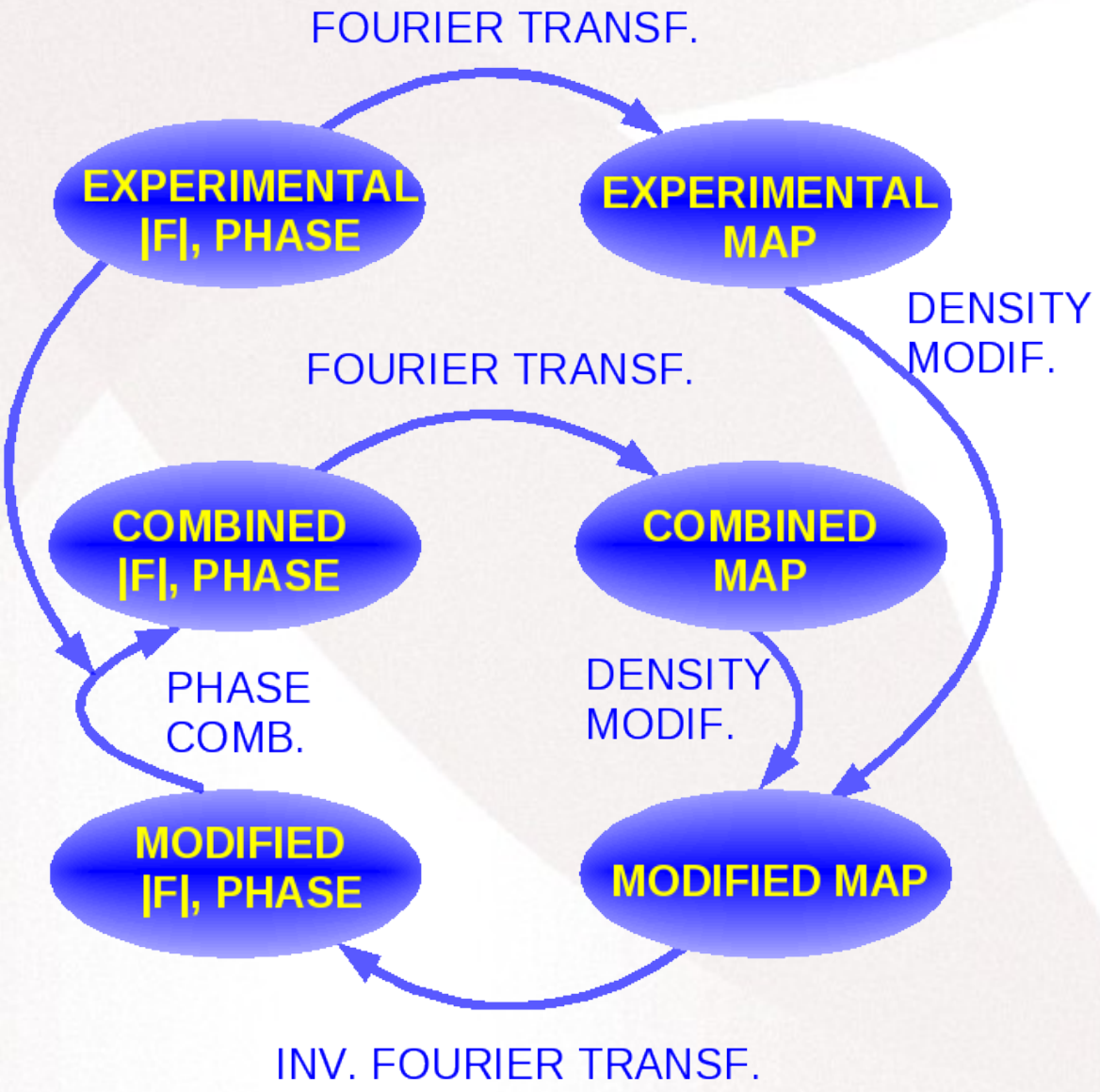
Content

- Density modification principles and problems
- Density modification by Parrot
- Multivariate phase combination
- Bias reduction by β-correction
- Combining density modification with model building and refinement

Density modification principles

- Idea: apply prior information about protein electron density maps to the experimental map
- Solvent flatness prior: the density in solvent regions is close to constant
- Histogram similarity prior: the histograms of protein density maps are similar
- NCS prior: the density in NCS related regions should be very similar

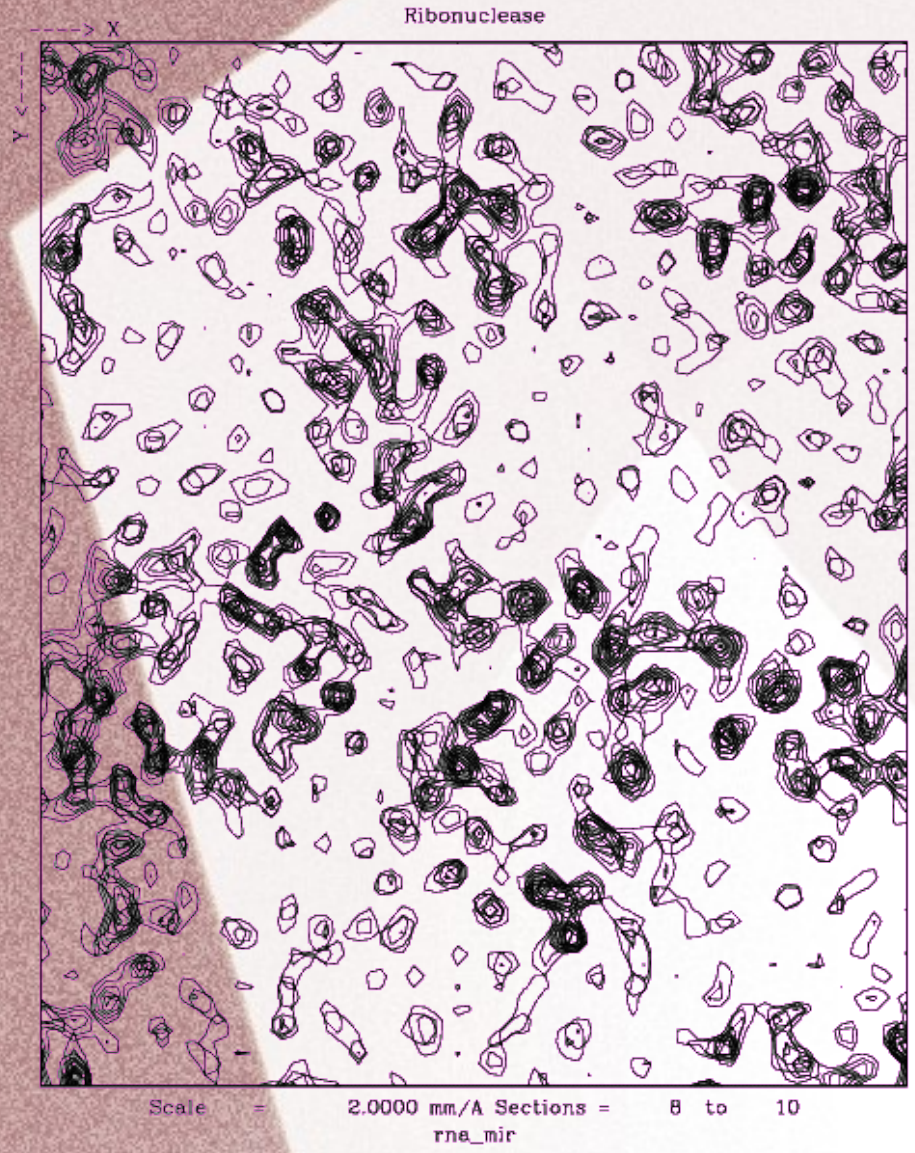
Density modification procedure

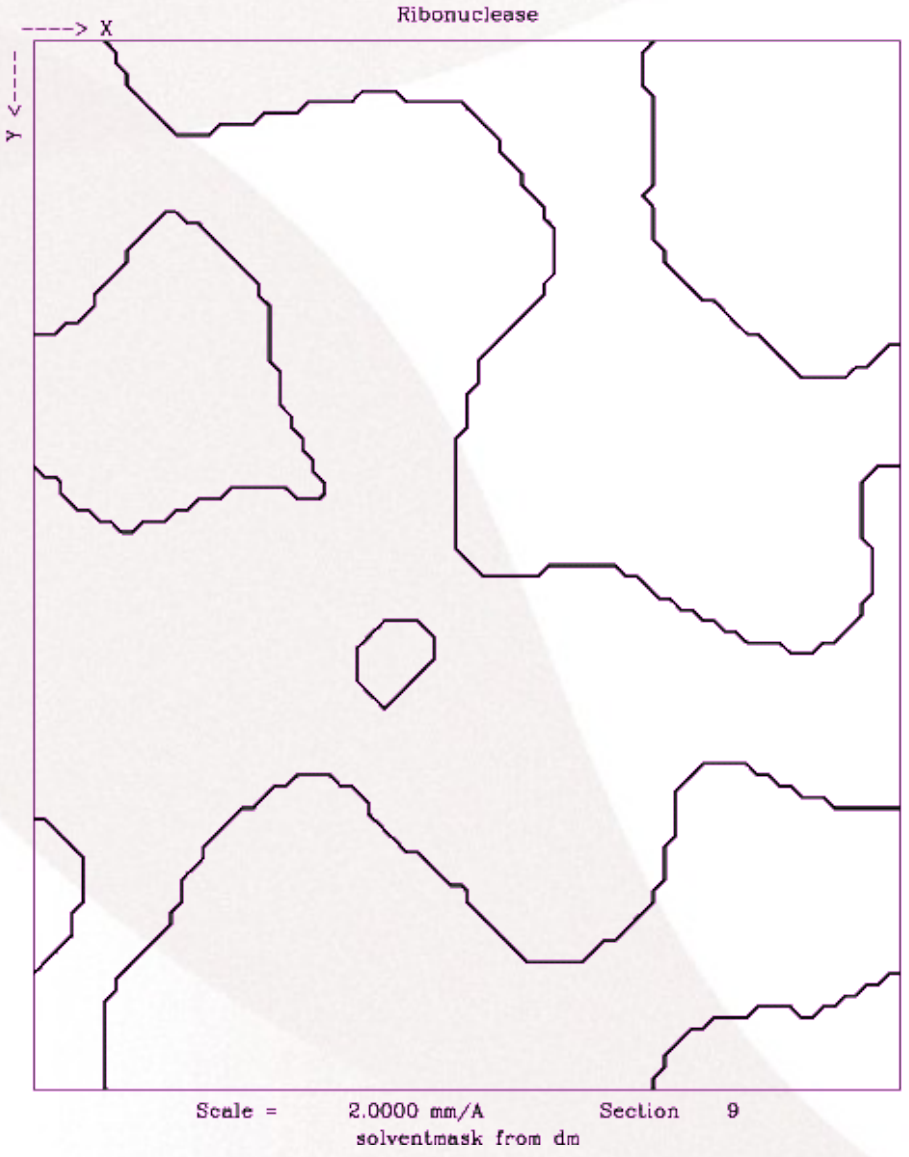


Solvent flattening

- The likely solvent regions in the map are identified by looking at local variance map
- The density in these regions is flattened
- If a solvent region was identified correctly, its flattening should remove noise from the map, improving the phases

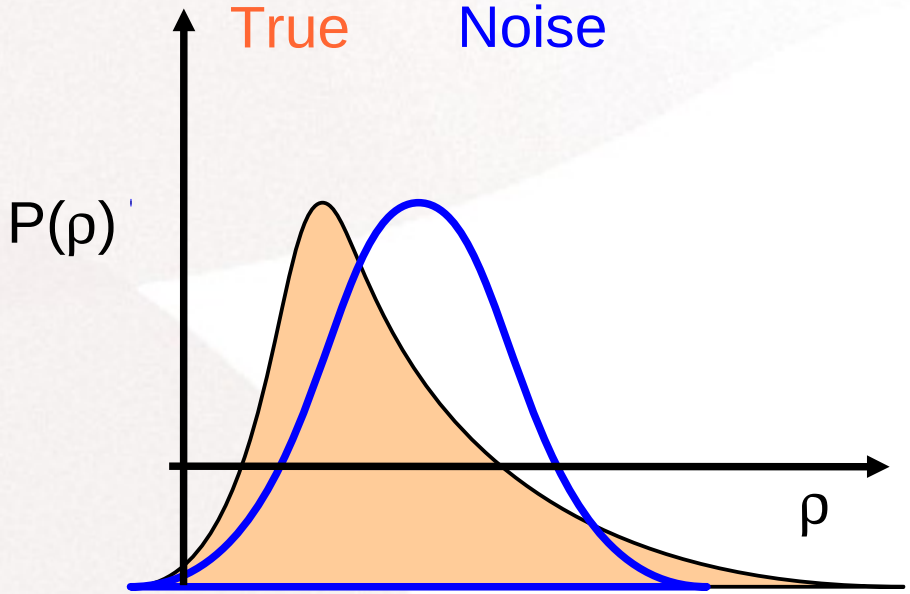
Solvent flattening

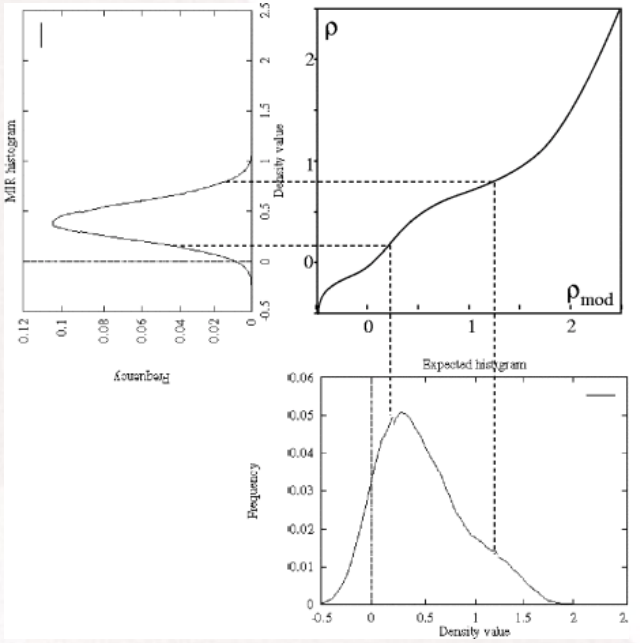




Histogram matching

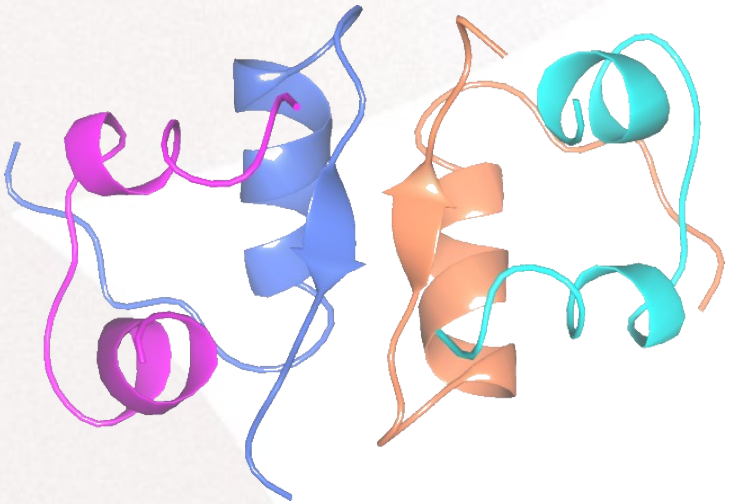
- A technique from image processing for modifying the protein region.
- Noise maps have Gaussian histogram.
- Well phased maps have a skewed distribution: sharper peaks and bigger gaps.
- The protein density is sharpened by a transform in order to match histogram of a well phased map.

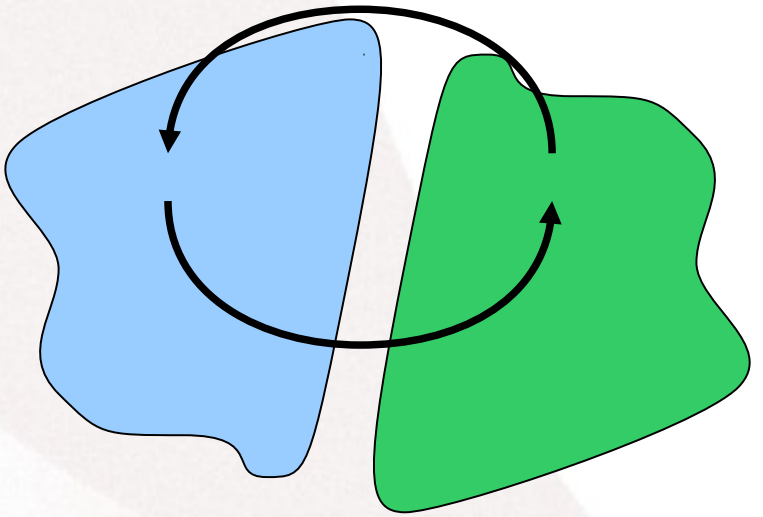




Non-crystallographic symmetry

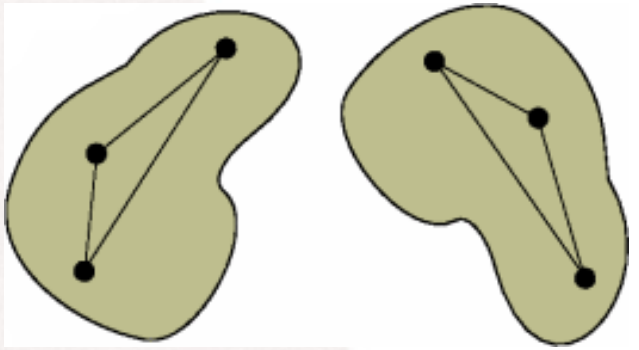
- If the molecule has internal symmetry, we can average together related regions.
- In the averaged map, the signal-noise level is improved.
- If a full density modification calculation is performed, powerful phase relationships are formed.





Non-crystallographic symmetry

- How do you know if you have NCS?
  - Cell content analysis – how many monomers in ASU?
  - Self-rotation function.
  - Difference Patterson (pseudo-translation only).



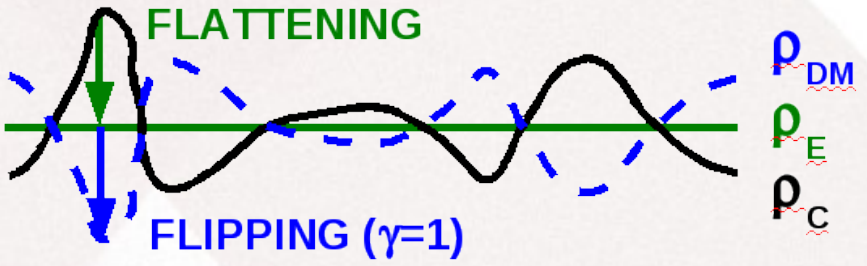
- How do you determine the NCS?
  - **From heavy atoms.**
  - From model building.
  - From molecular replacement.

Bias in density modification

- Usual assumption used in phase combination: Independence between the density modified phases and the experimental phases
- Reality: The density modified phases are derived from the experimental phases -> they are significantly correlated
- Result: Measures of phase quality (FOM) are overestimated, HL coefficients are biased and phases derived may be suboptimal

Usual bias reduction methods

- Map flipping (gamma correction): The (likely) solvent regions are not flattened but flipped



- Structure factor flipping: $\overrightarrow{F_{DM}} = 2\overrightarrow{F_b} - \overrightarrow{F_c}$

- Statistical density modification (Pirate)

Content

- Density modification principles and problems
- Density modification by Parrot
- Multivariate phase combination
- Bias reduction by β-correction
- Combining density modification with model building and refinement

Density modification by Parrot

Kevin Cowtan: author of Parrot and several slides of this talk

Parrot (2010) builds on existing ideas:

- Solvent flattening
- Histogram matching
- NCS averaging
- Gamma correction

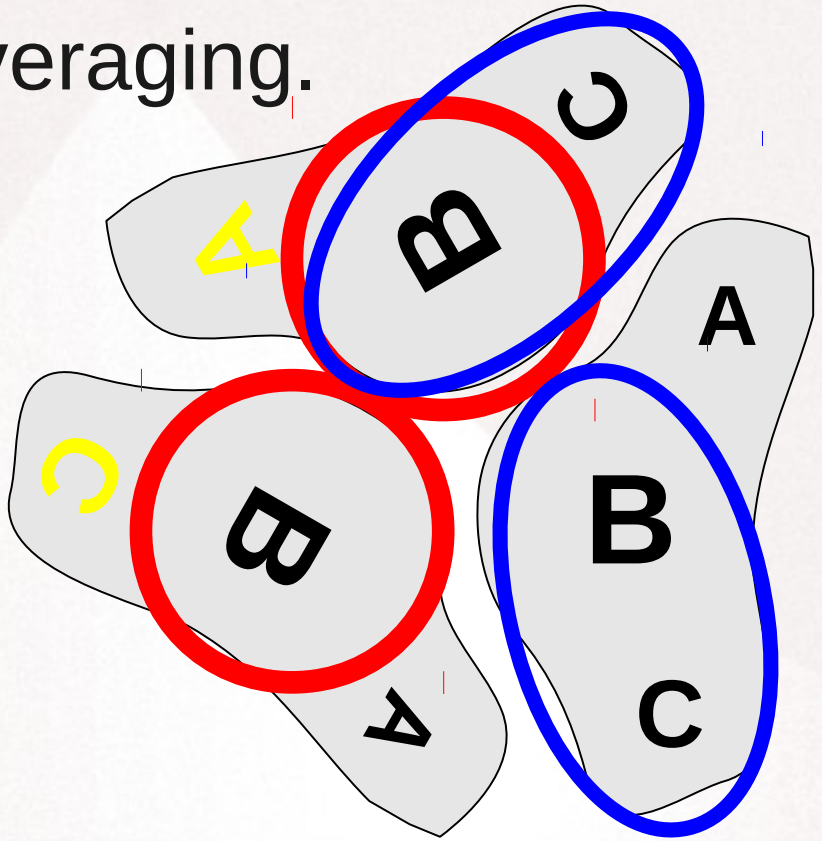
Density modification by Parrot

New developments:

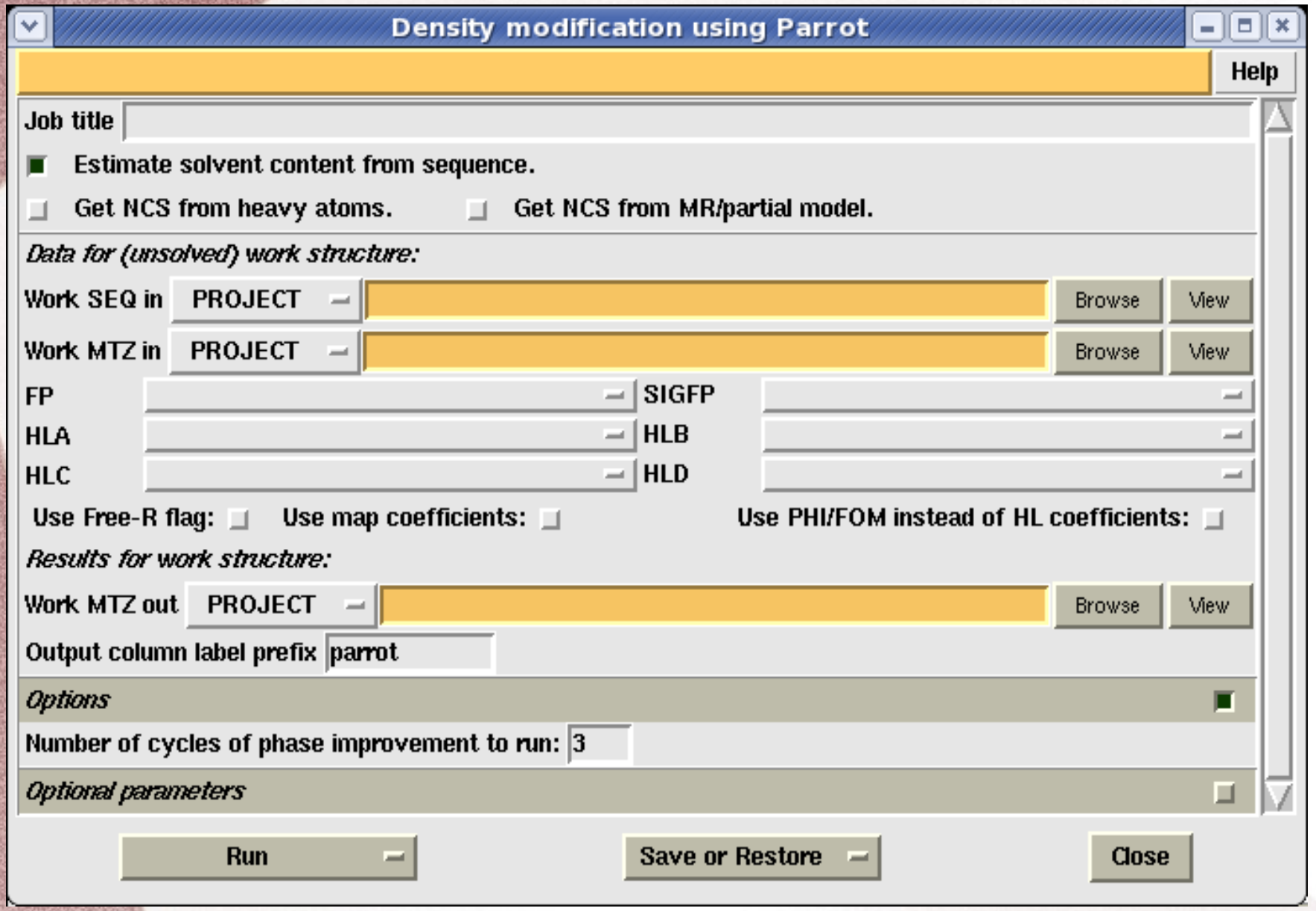
- MLHL phase combination
  (similar to MLHL used in refinement)
- Anisotropy correction
- Problem-specific density histograms
  (rather than a standard library)
- Automatic NCS determination
- Pairwise-weighted NCS averaging

Pairwise-weighted averaging

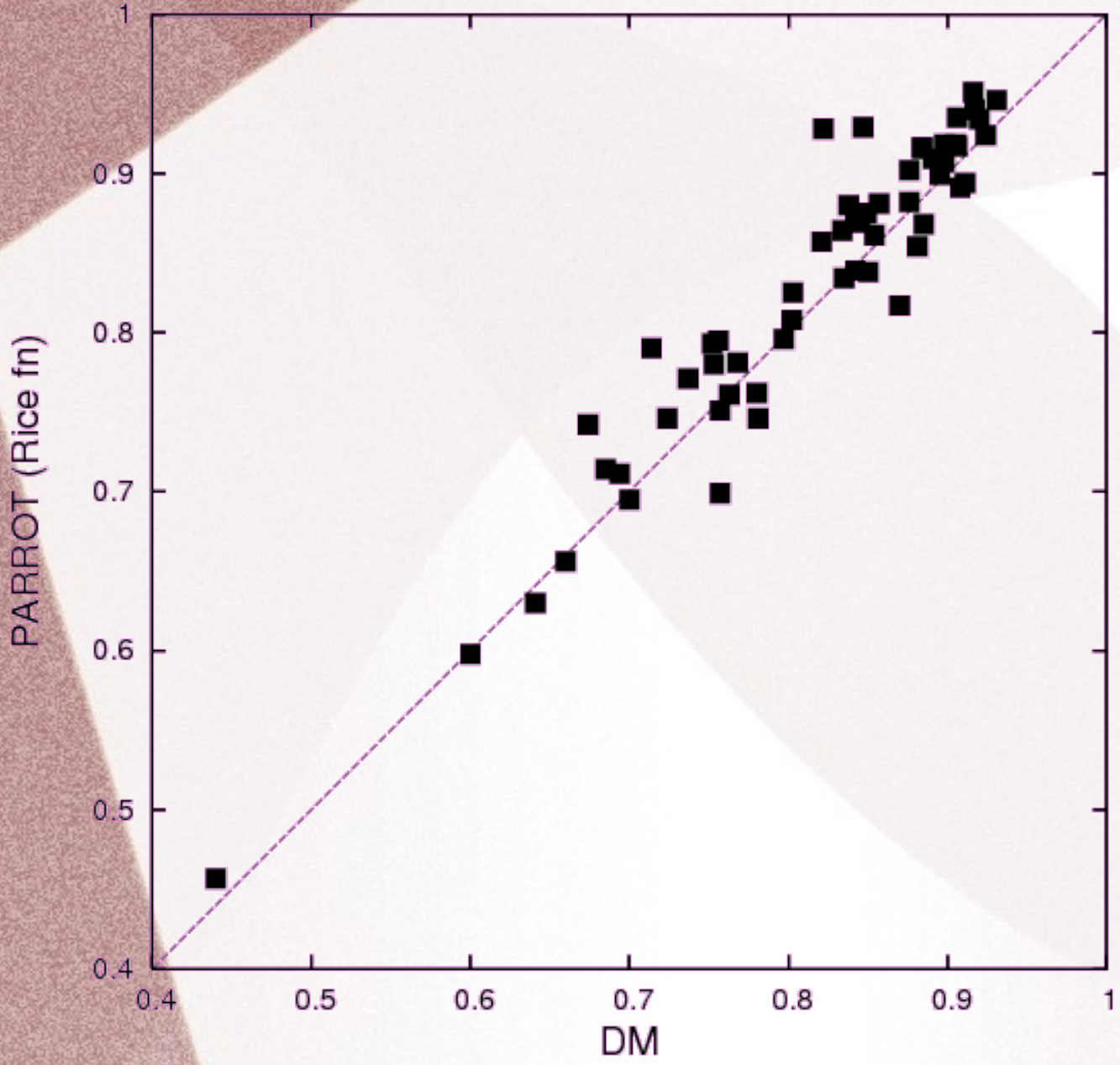
- Average each pair of NCS related molecules separately with its own mask.
- Generalisation and automation of multi-domain averaging.



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

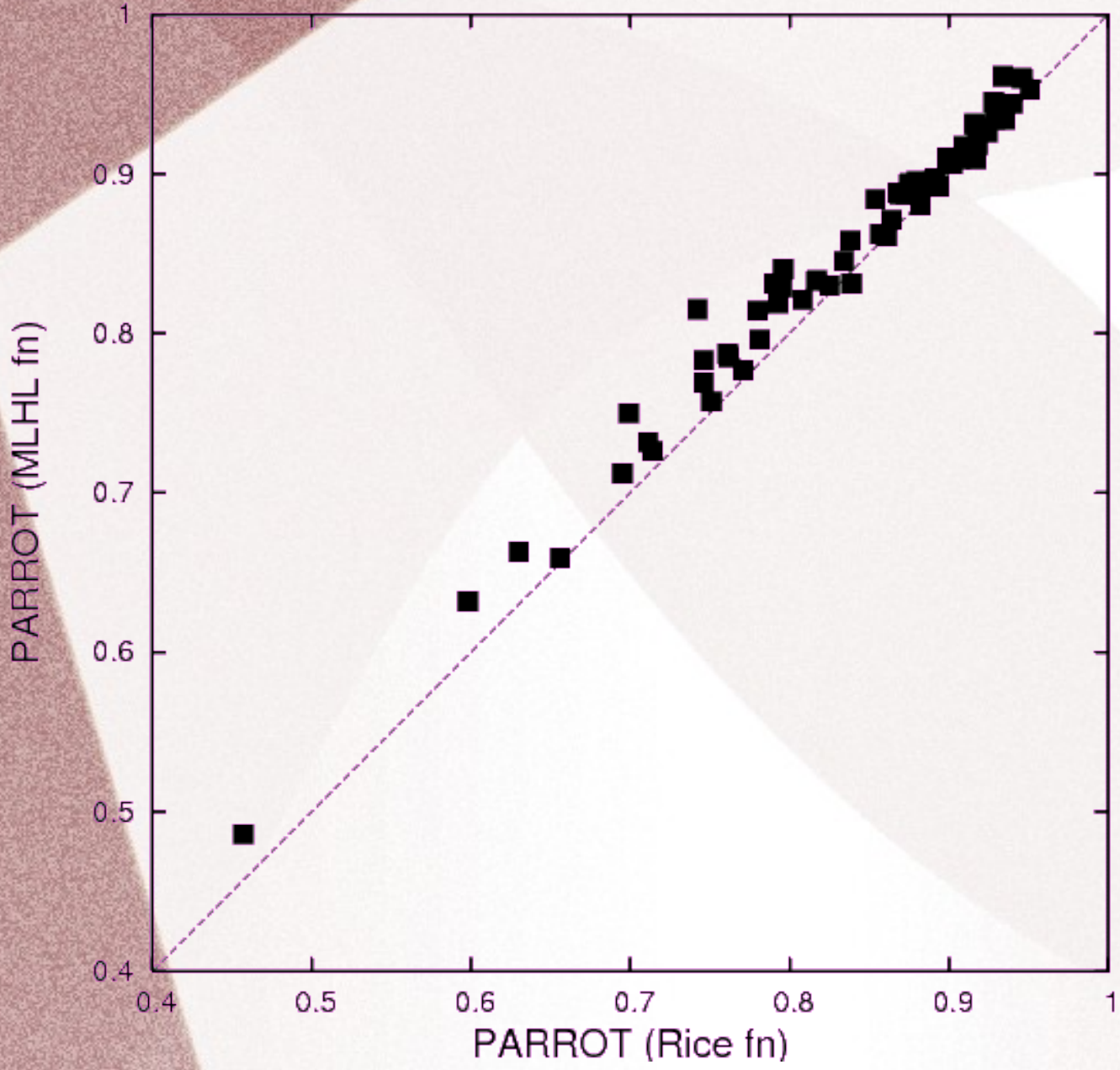
DM vs Parrot



Map correlations

Parrot: No new features enabled.

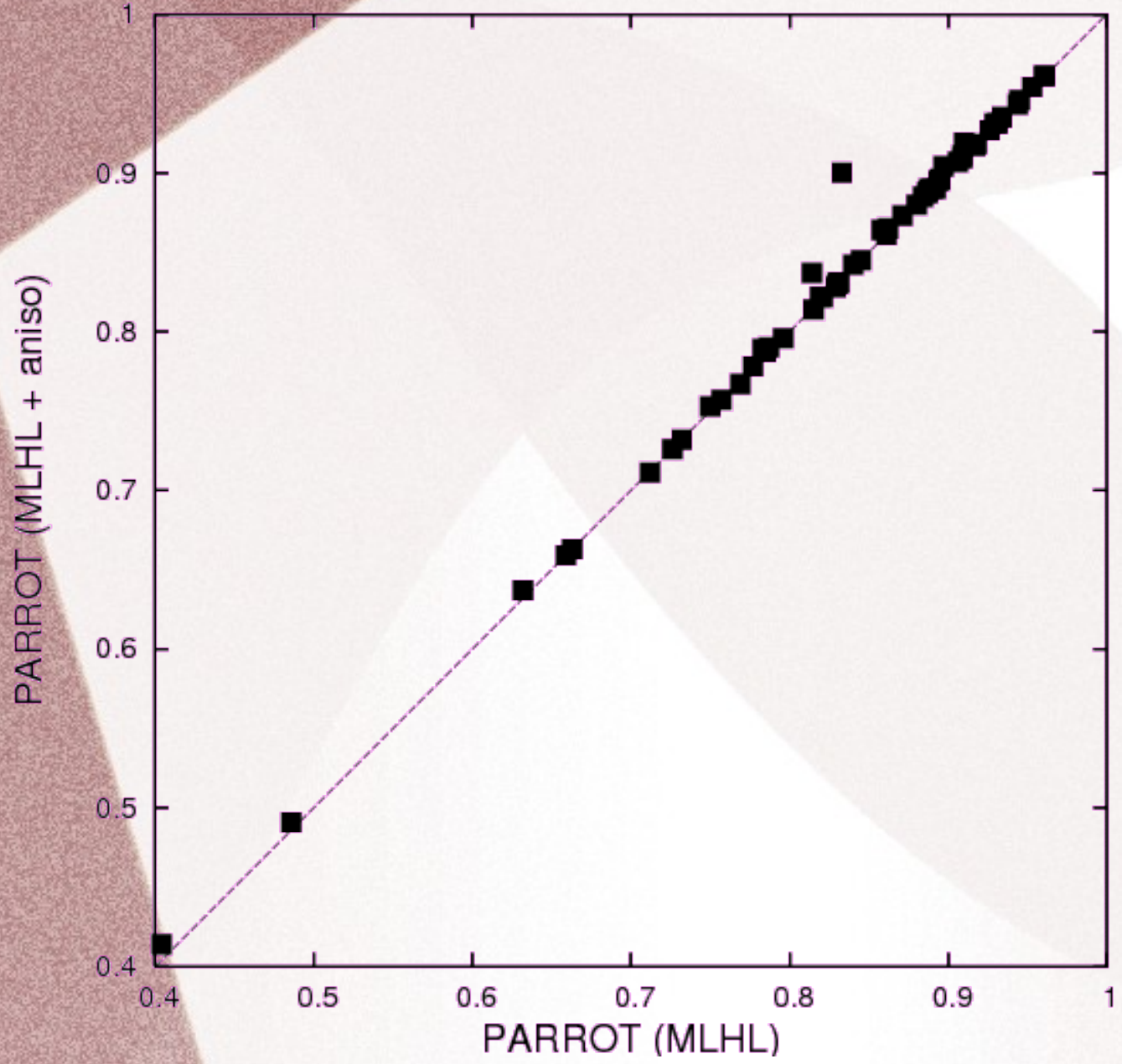
Parrot: Rice vs MLHL



Map correlations

Comparing old and new likelihood functions.

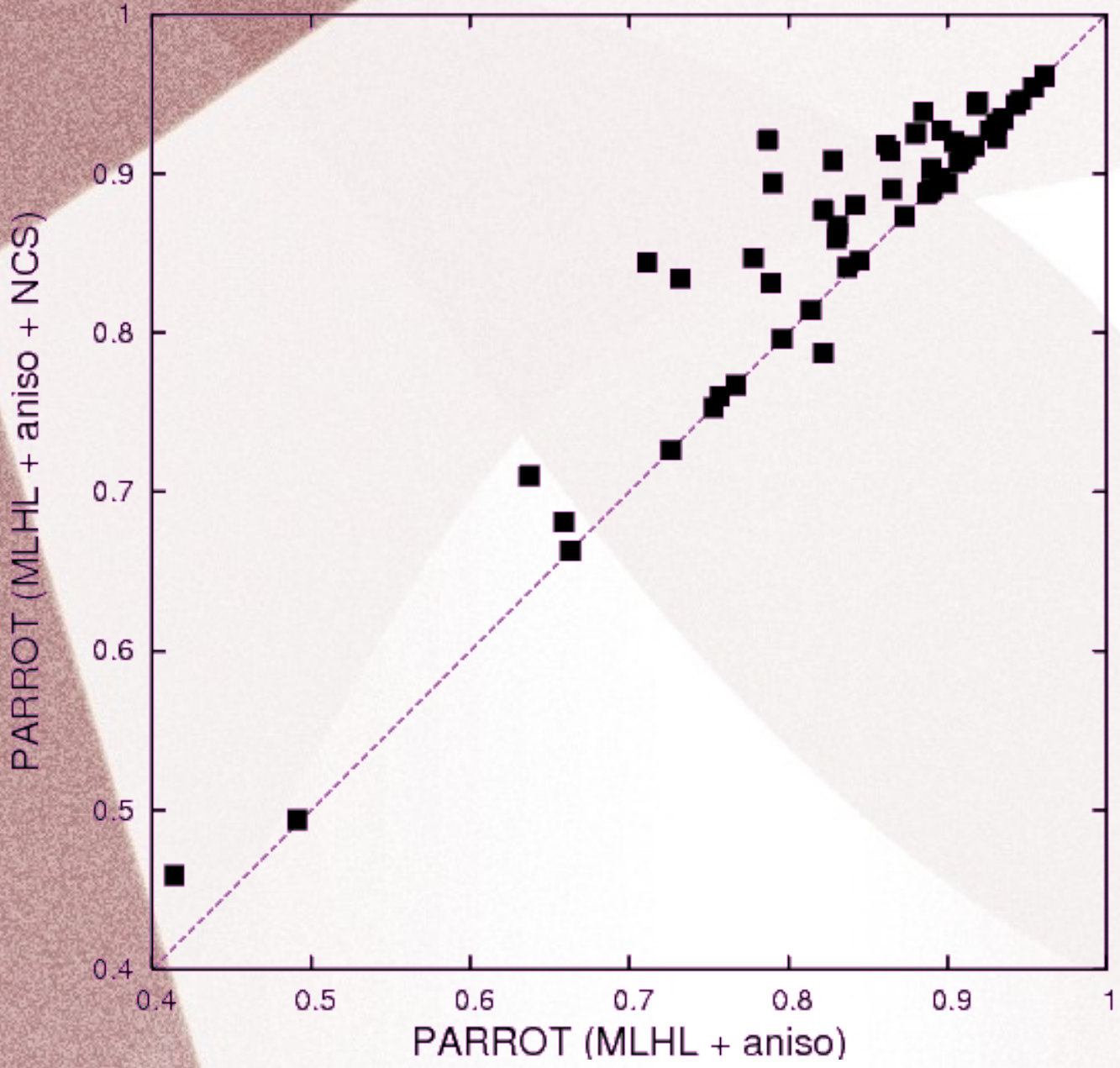
Parrot: Isotropic vs Anisotropic



Map correlations

Comparing with and without anisotropy correction.

Parrot: simple vs NCS averaged



Map correlations

Comparing with and without NCS averaging.

Content

- Density modification principles and problems
- Density modification by Parrot
- Multivariate phase combination
- Bias reduction by β-correction
- Combining density modification with model building and refinement

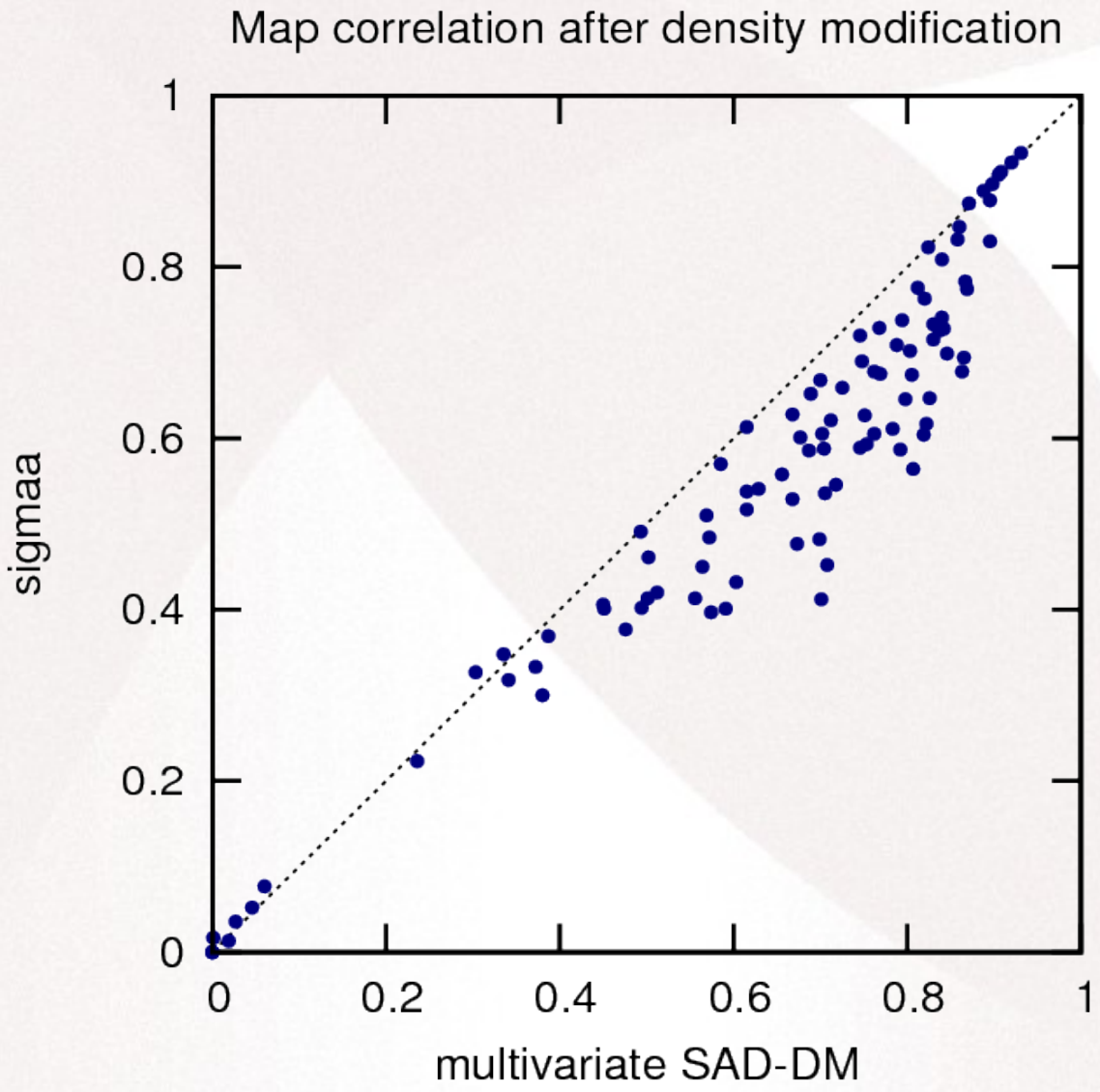
Multivariate phase combination for density modification

- Current density modification procedures neglect the correlation between the original map and the density modified map
- Our approach: multivariate SAD-DM probability distribution

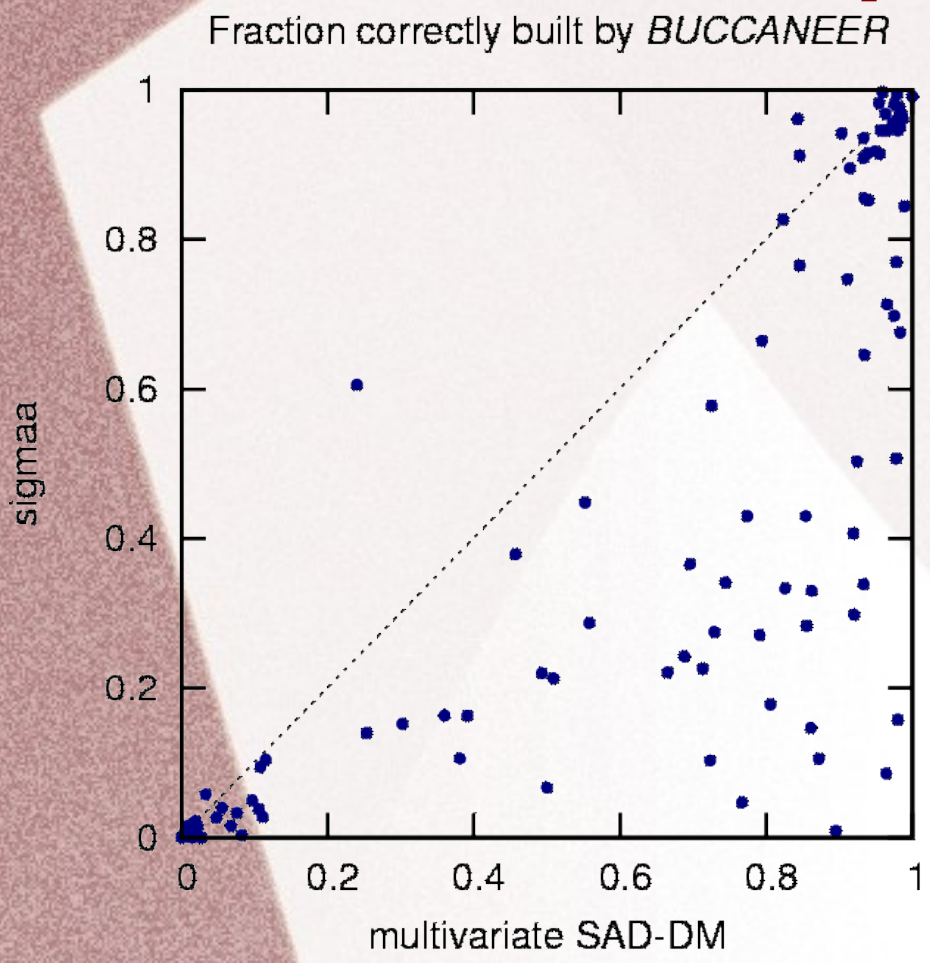
$$P_{DM}\left(F_o^+, F_o^- \middle| F_H^+, \varphi_H^+, F_H^-, \varphi_H^-, F_{DM}, \varphi_{DM}\right)$$

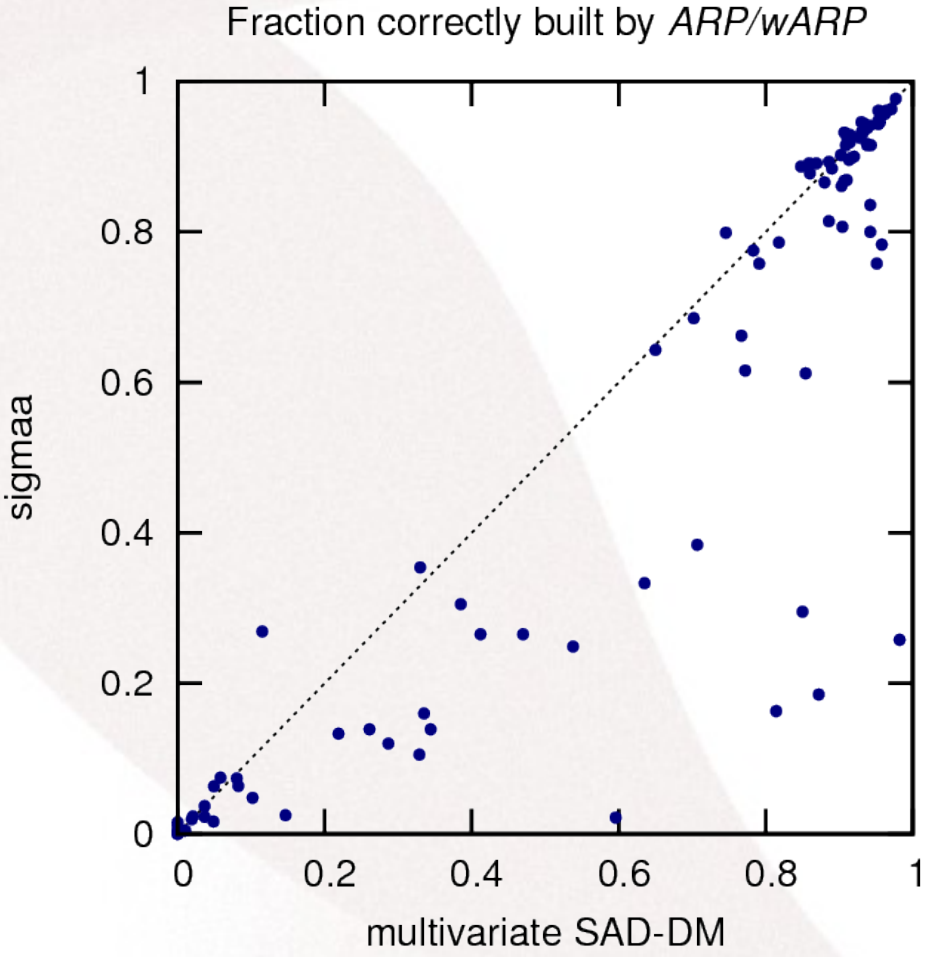
- Advantages: no independence assumption, dynamic construction of phase probabilities rather than static HL
- Implementation: MULTICOMB, REFMAC

Comparison of sigmaa vs. multivariate SAD-DM function



Results of model building after DM with sigmaa and SAD-DM phase combination





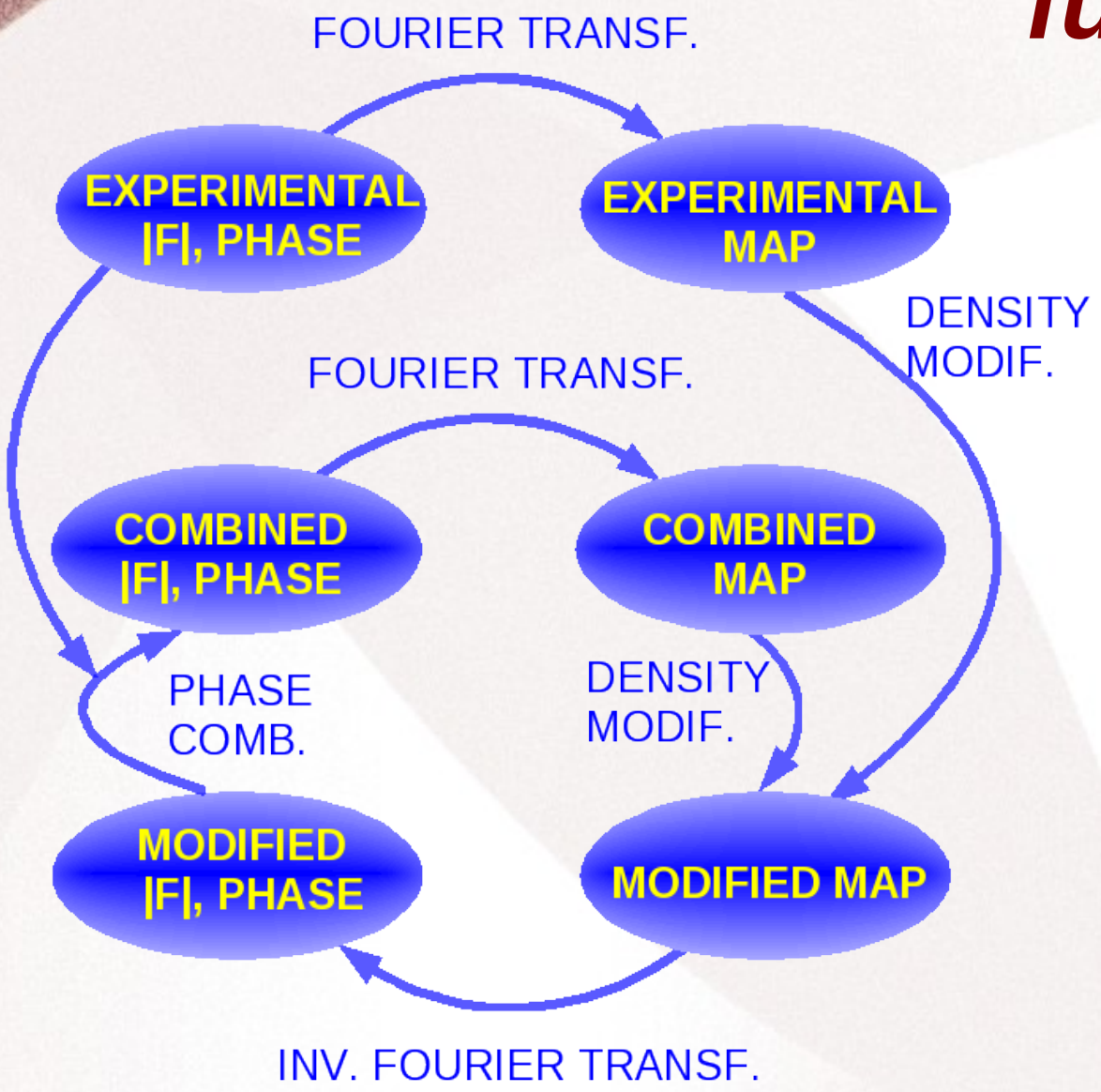
Bias of SAD-DM function

- Average bias (as difference between FOM and mean cosine of phase error):

Phase comb.	Sigmaa	SAD-DM
Bias	0.27	0.24

- Bias only slightly reduced
- Reason: the phase quality is estimated from fit between the observations and the model - which was constructed from the observations

Bias of the SAD-DM function



$$P_{DM}\left(F_o^+, F_o^- \middle| F_H^+, \varphi_H^+, F_H^-, \varphi_H^-, F_{DM}, \varphi_{DM}\right)$$

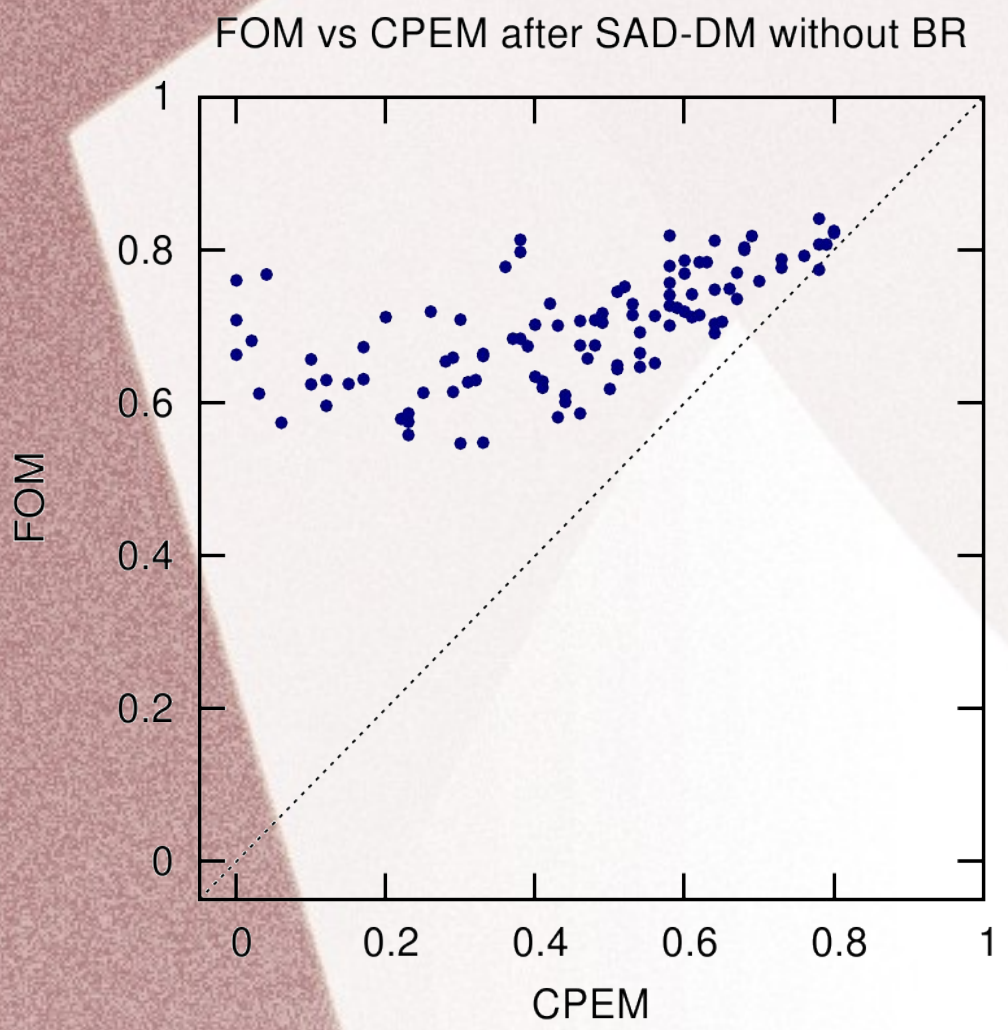
Content

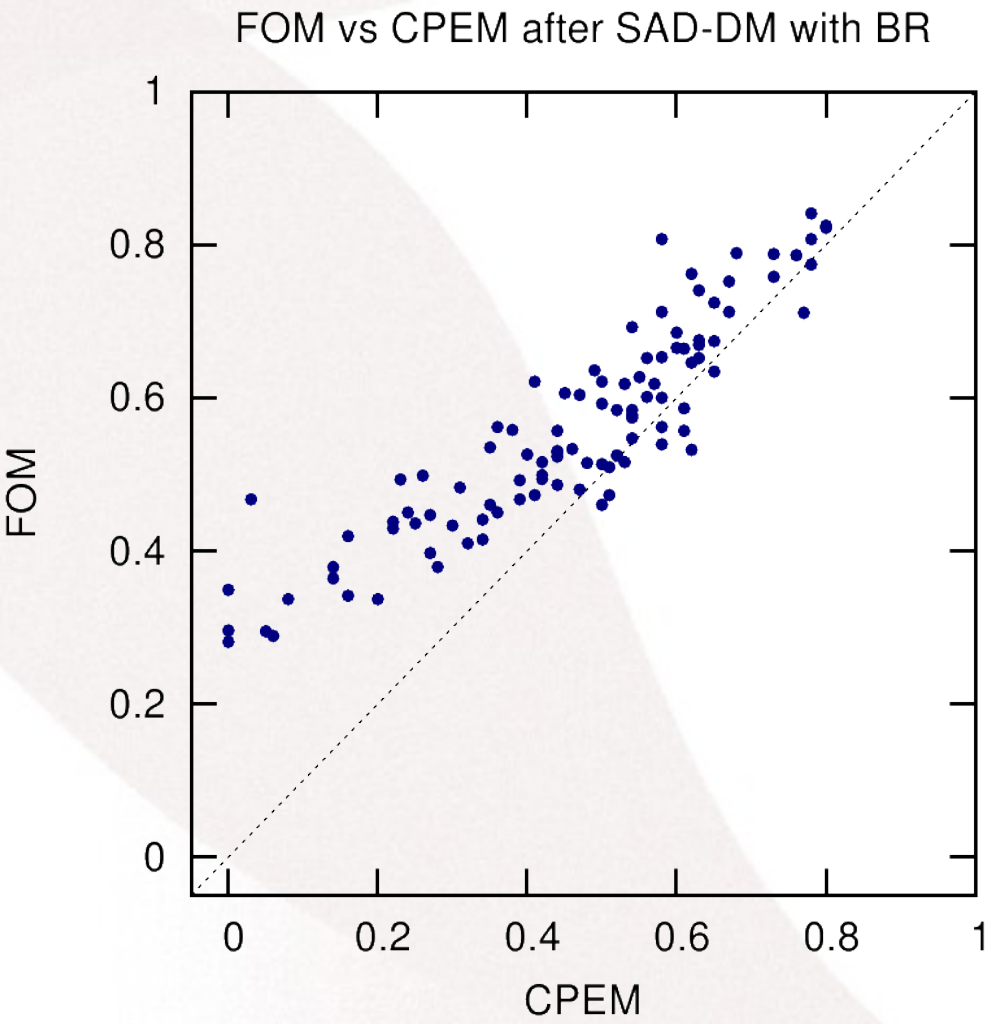
- Density modification principles and problems
- Density modification by Parrot
- Multivariate phase combination
- Bias reduction by β-correction
- Combining density modification with model building and refinement

Solution: β-correction

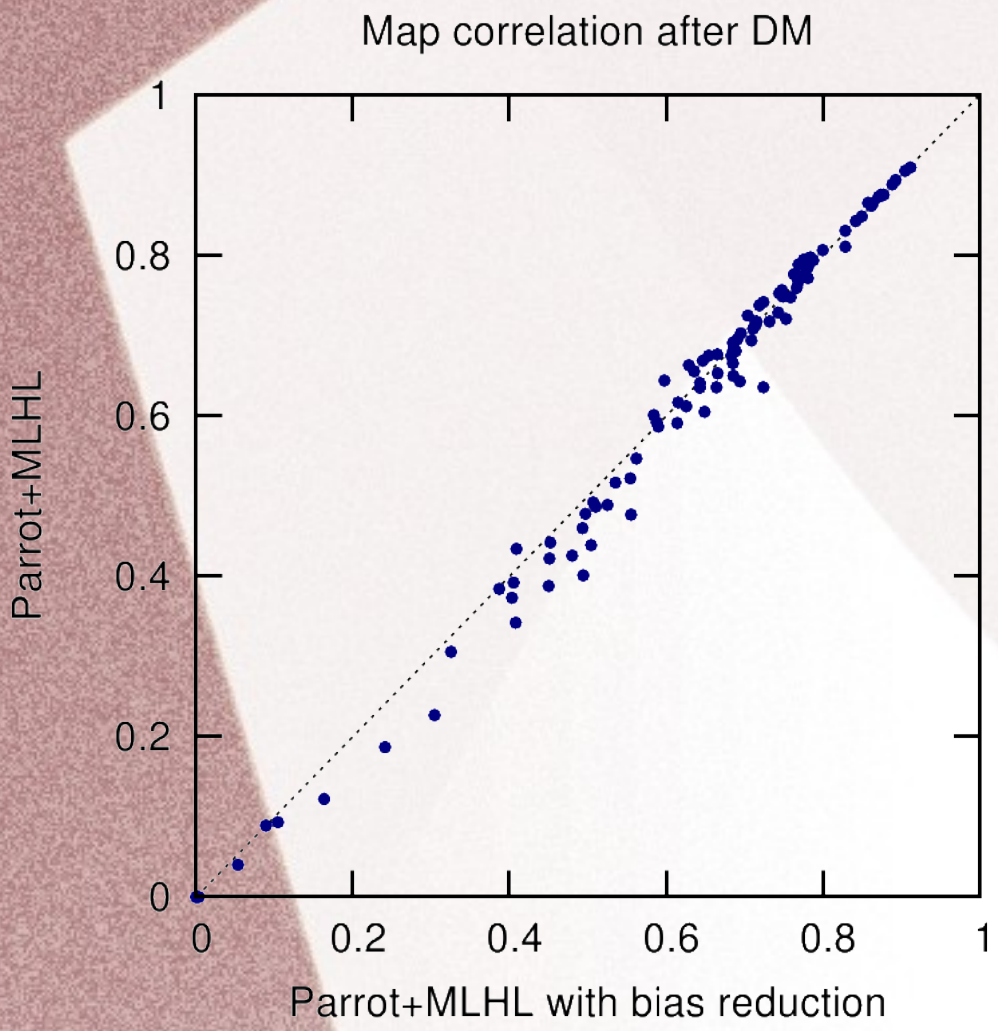
- Estimate overall bias parameter β:
  - run 5 DM cycles with 'free' set of reflections set aside and comparing the correlation between F_o and F_c in free and working set
- Run standard DM using β (and all reflections):
  - β is applied to a correlation term betwen F_o and F_c in the covariance matrix
- β is 1 in case of no bias and can reach 0.2-0.3 in case of very strong bias

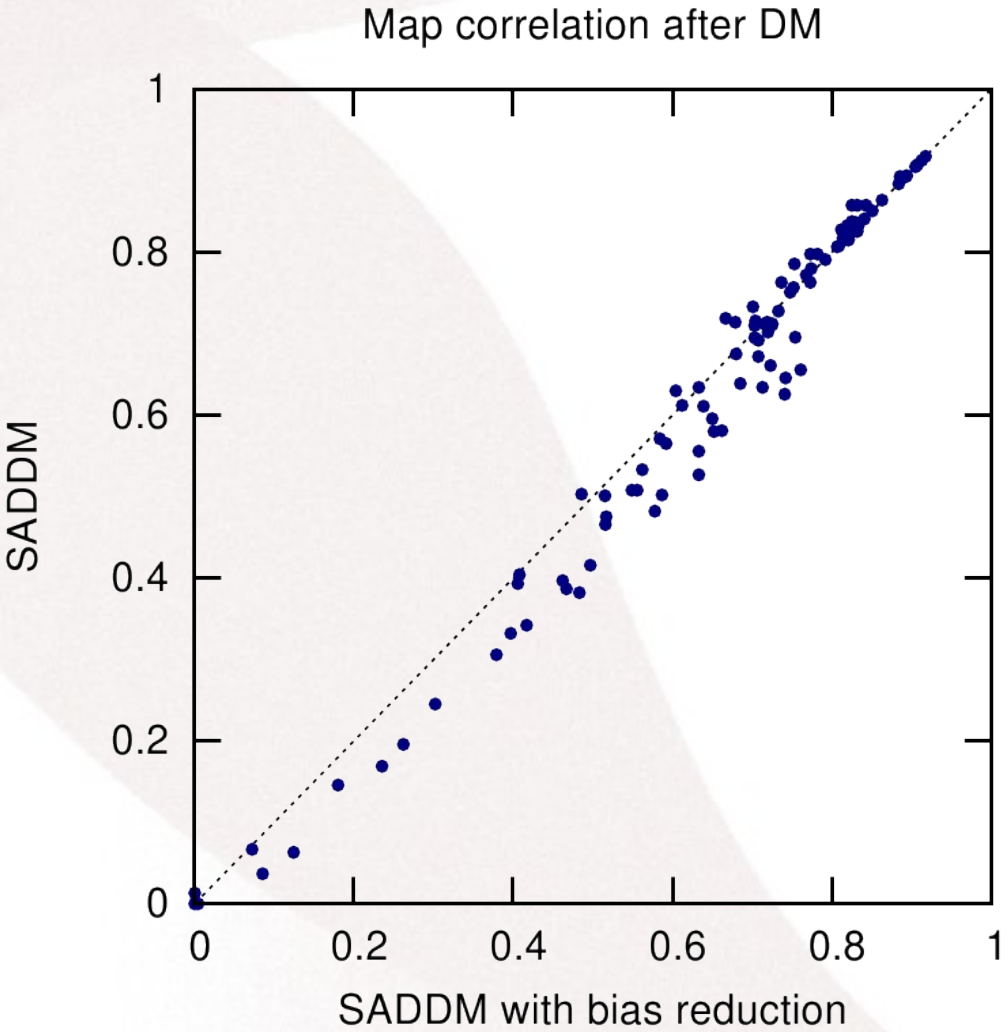
FOM and phase error after DM with/without bias reduction





Map correlation after DM with/without bias reduction





Content

- Density modification principles and problems
- Density modification by Parrot
- Multivariate phase combination
- Bias reduction by β-correction
- Combining density modification with model building and refinement

Previously mentioned SAD distributions

- function for SAD substructure refinement

$$P_{ph}\left(F_o^+, F_o^- \middle| F_H^+, \varphi_H^+, F_H^-, \varphi_H^-\right)$$

  - implemented in BP3, Refmac5

- function for SAD phase combination in DM

$$P_{DM}\left(F_o^+, F_o^- \middle| F_H^+, \varphi_H^+, F_H^-, \varphi_H^-, F_{DM}, \varphi_{DM}\right)$$

  - implemented in Multicomb, Refmac5

- function for SAD phased protein model refinement

$$P_{ref}\left(F_o^+, F_o^- \middle| F_P^+, \varphi_P^+, F_P^-, \varphi_P^-\right)$$

  - implemented in Refmac5

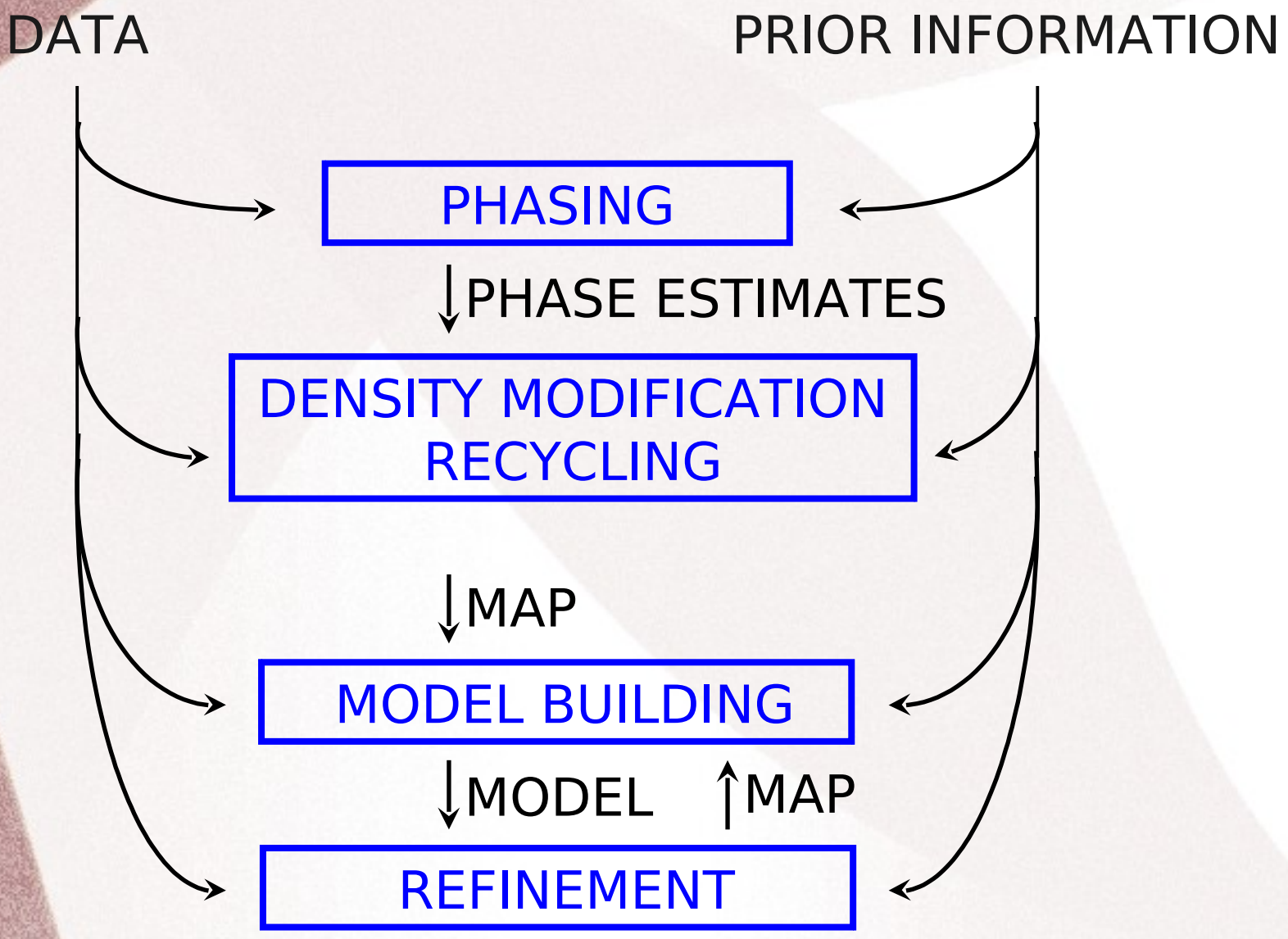
SAD-DM-ref function

- function for SAD phased and DM restrained refinement:

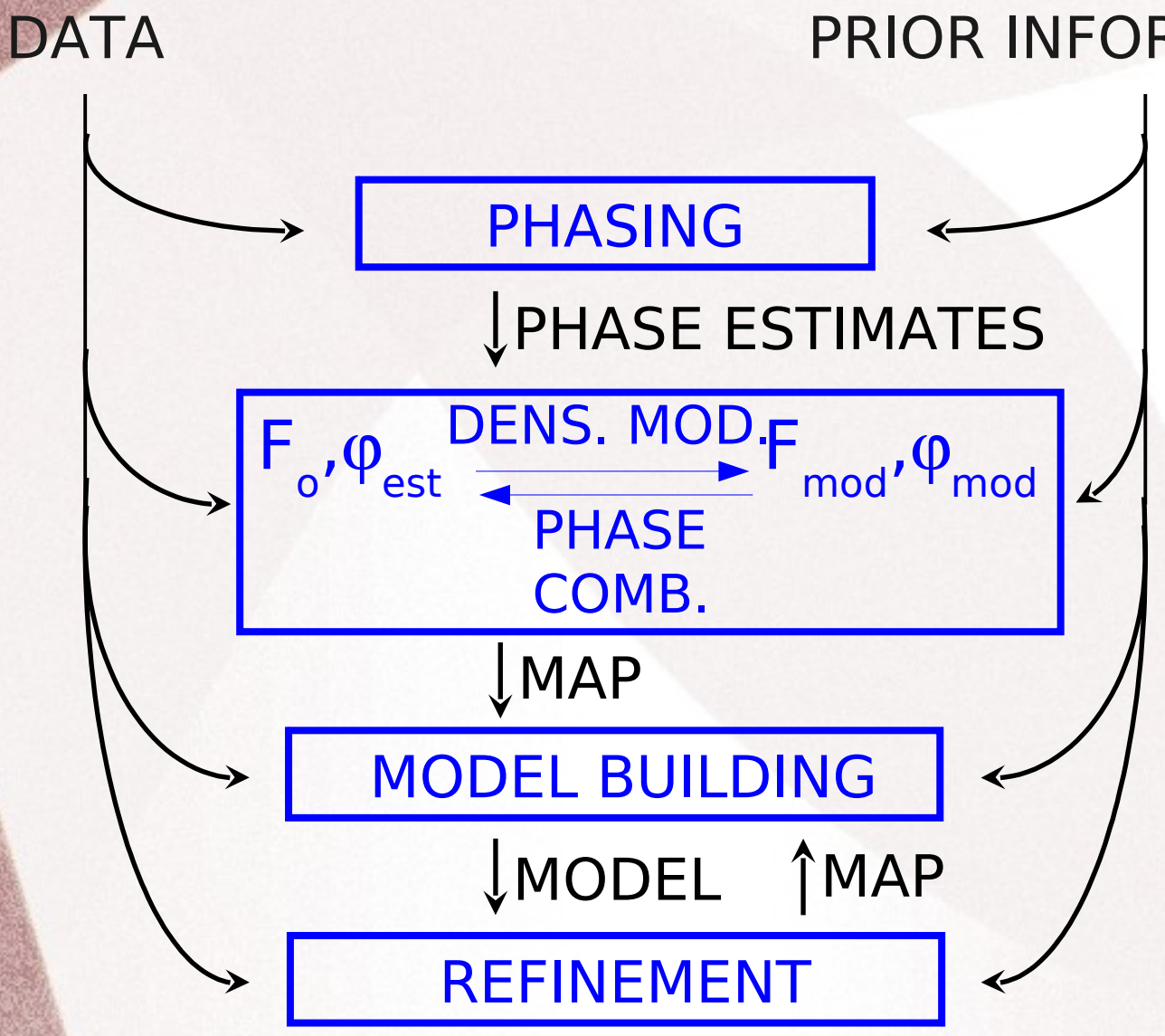
$$P_{DM+ref}\left(F_o^+, F_o^- \middle| F_P^+, \varphi_P^+, F_P^-, \varphi_P^-, F_{DM}, \varphi_{DM}\right)$$

- combines SAD phase information with information from density modified map and uses it for refinement of the current model
- all previous function are theoretically a special case of this function
- implemented in Refmac5 and Crank 1.5 (development version)

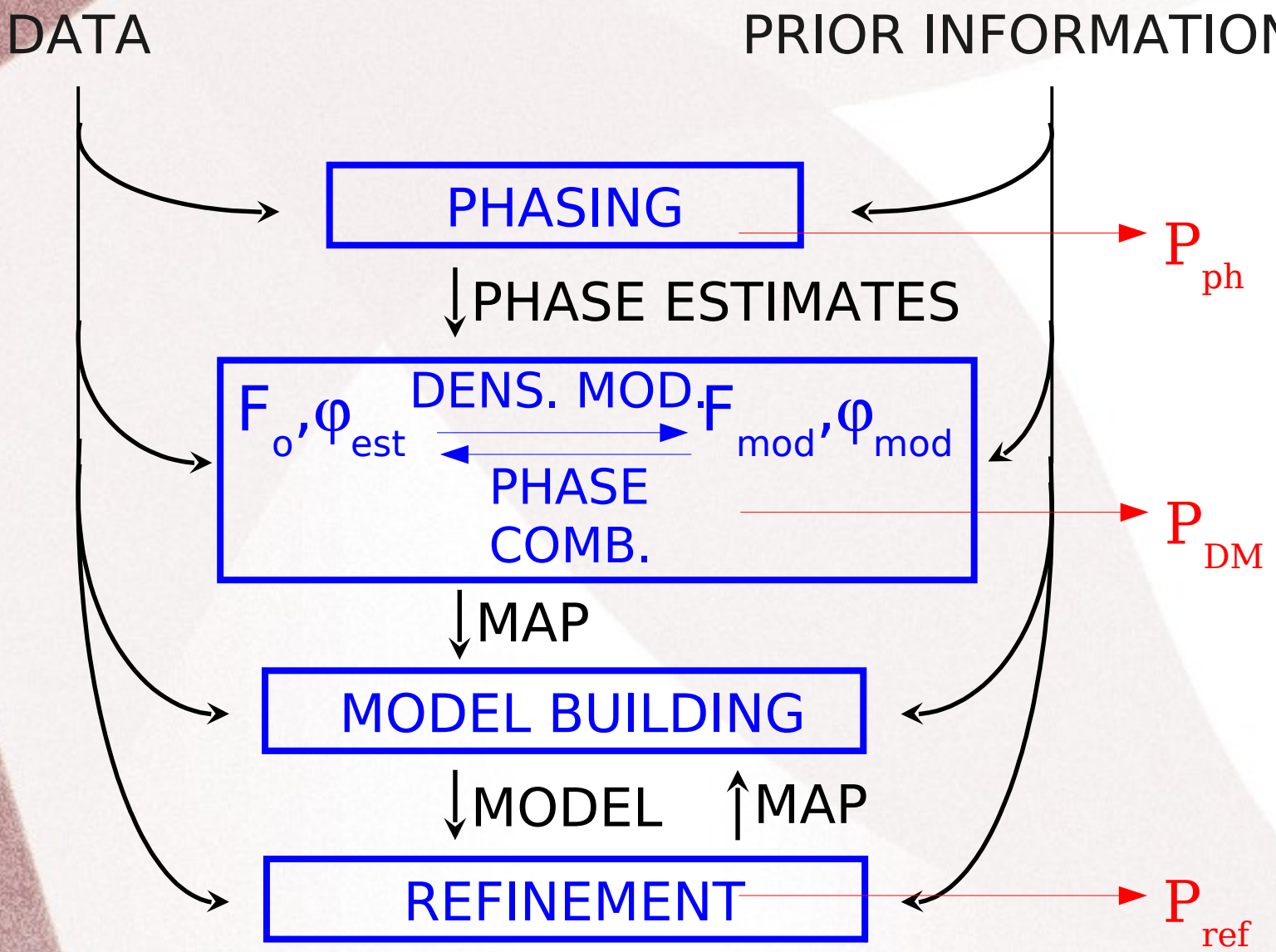
Usual structure solution from SAD phases



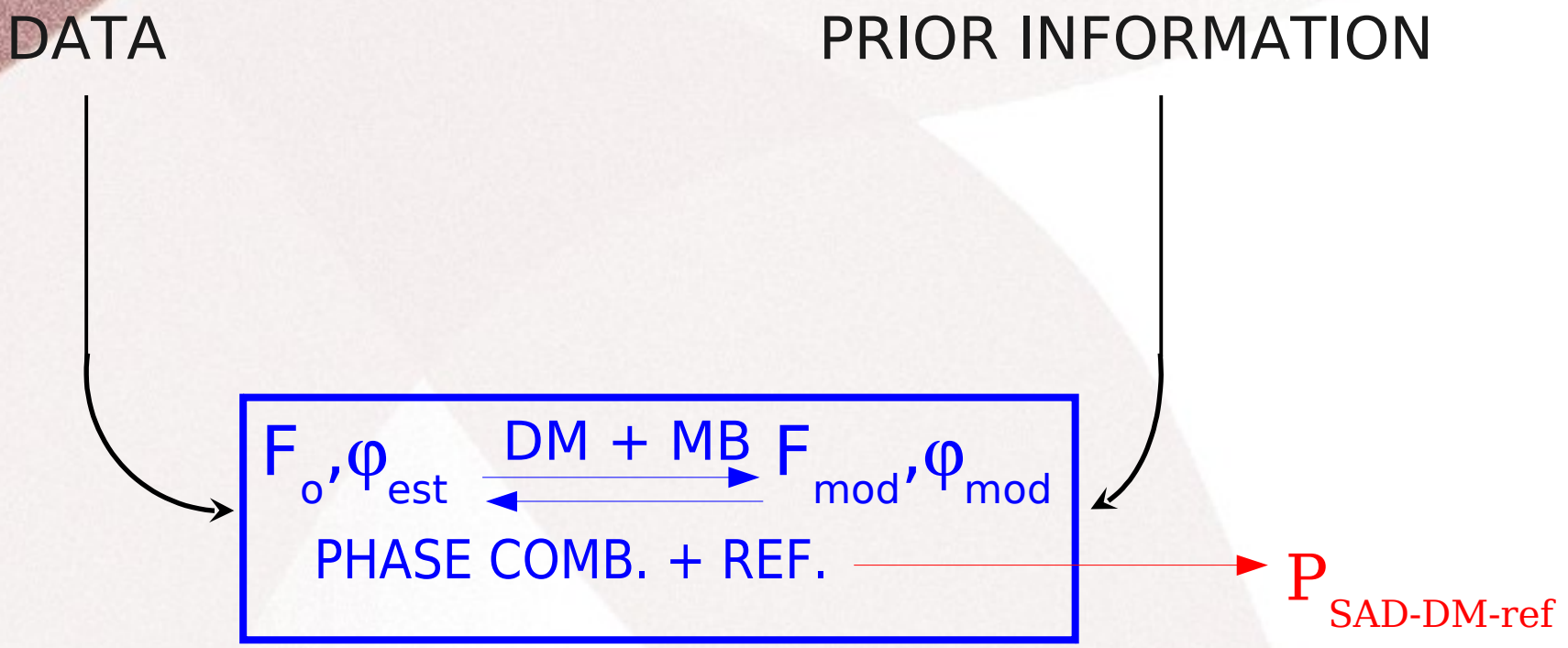
Usual structure solution from SAD phases



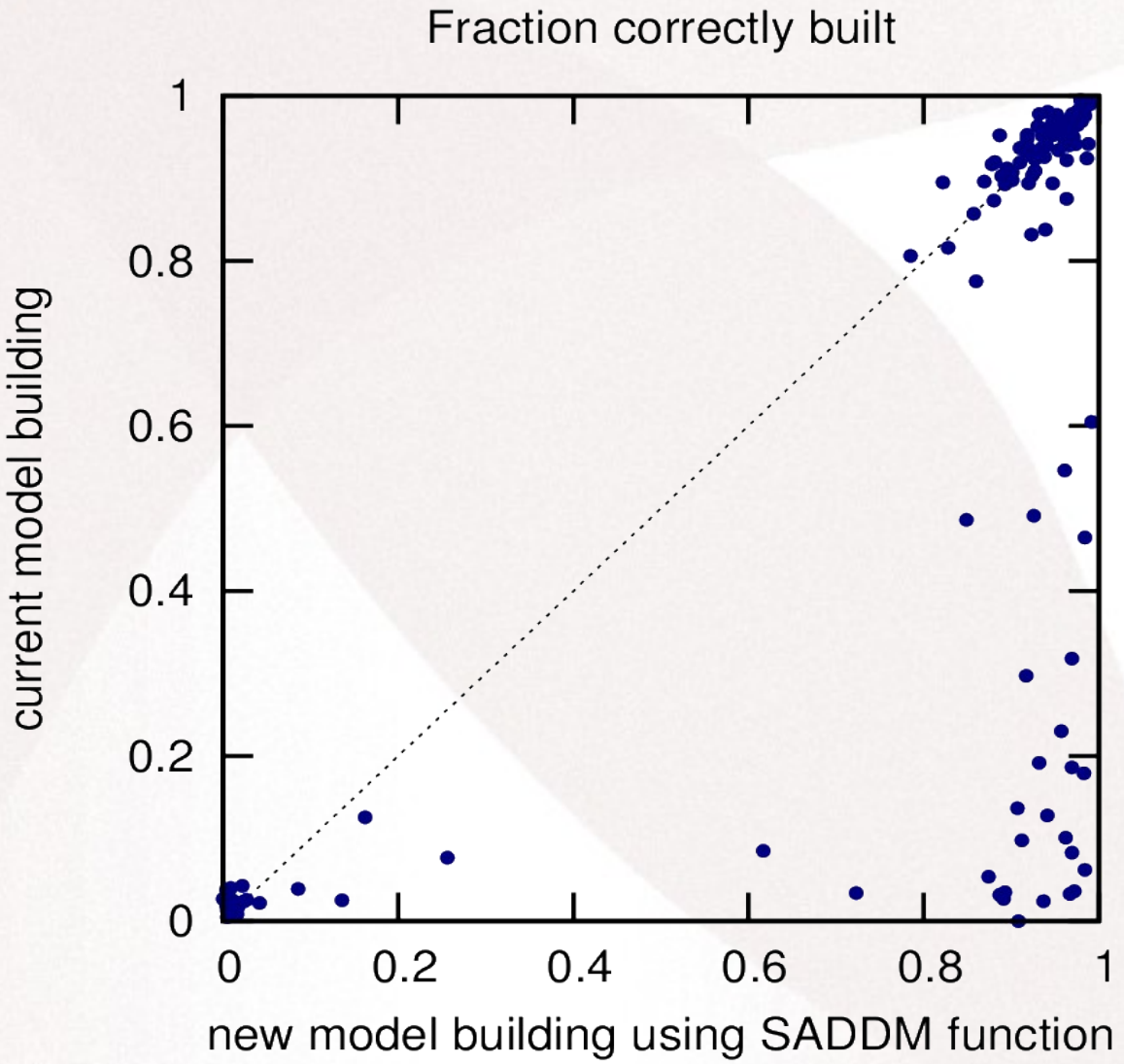
Usual structure solution from SAD phases



Combined structure solution from SAD phases



Model building results



Conclusions

- Density modification primarily incorporates prior information about map in the real space
- In order to reduce errors and bias, powerful phase combination algorithms are used in reciprocal space and the process is iterated
- Density modification can be incorporated in model building and refinement, leading to many more models automatically built