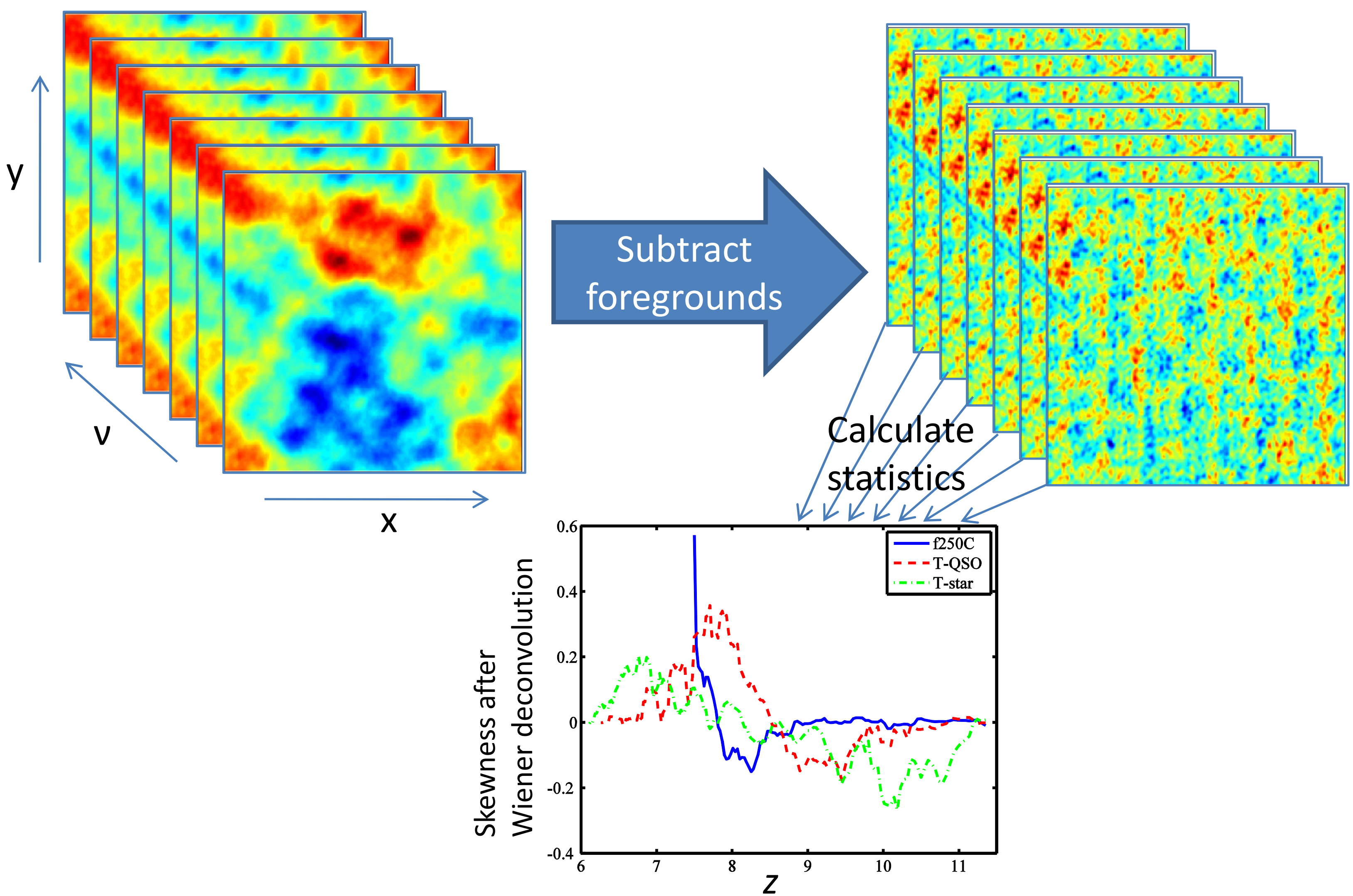
Non-parametric foreground fitting

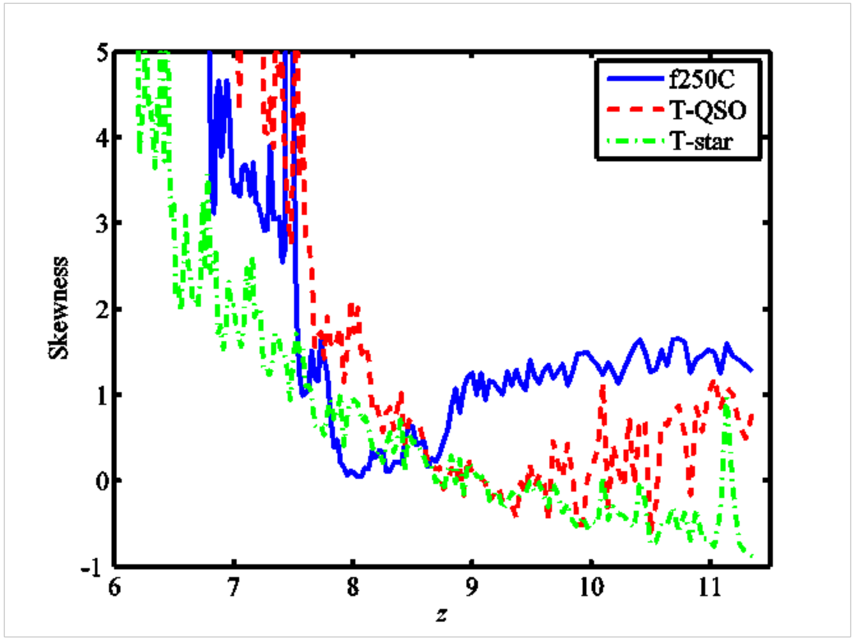
Geraint Harker

Signal extraction in two stages

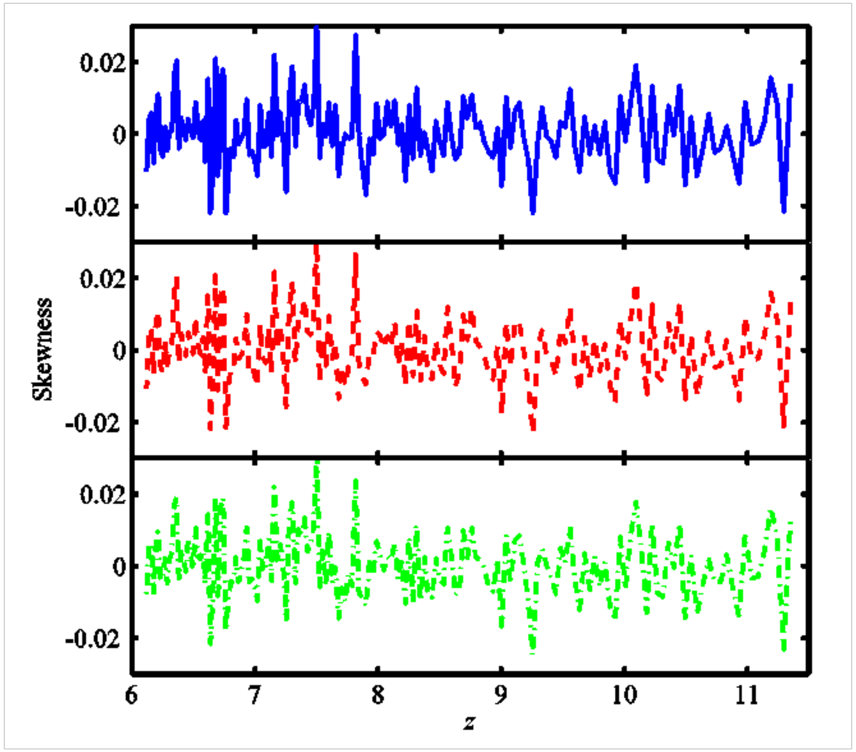


The importance of good foreground fitting

Original simulations

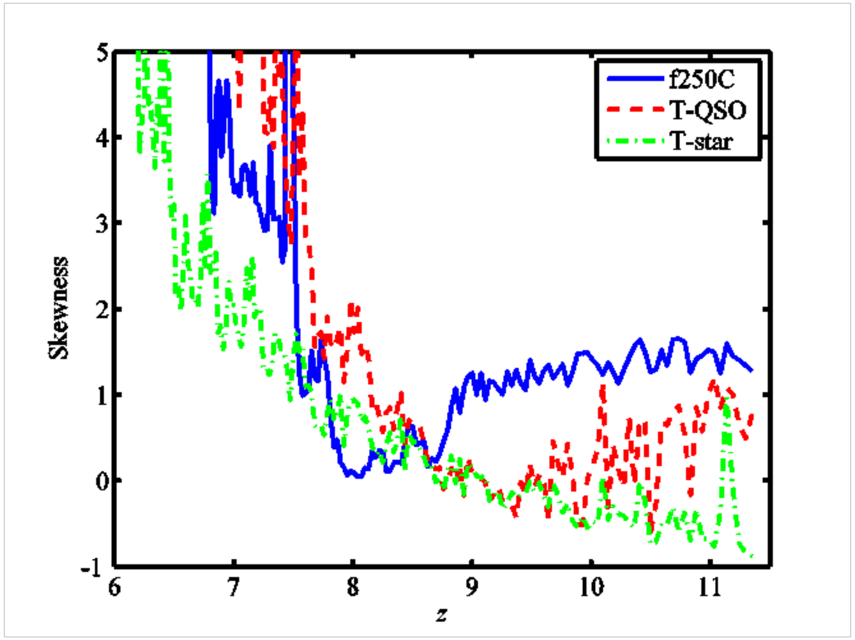


Residuals

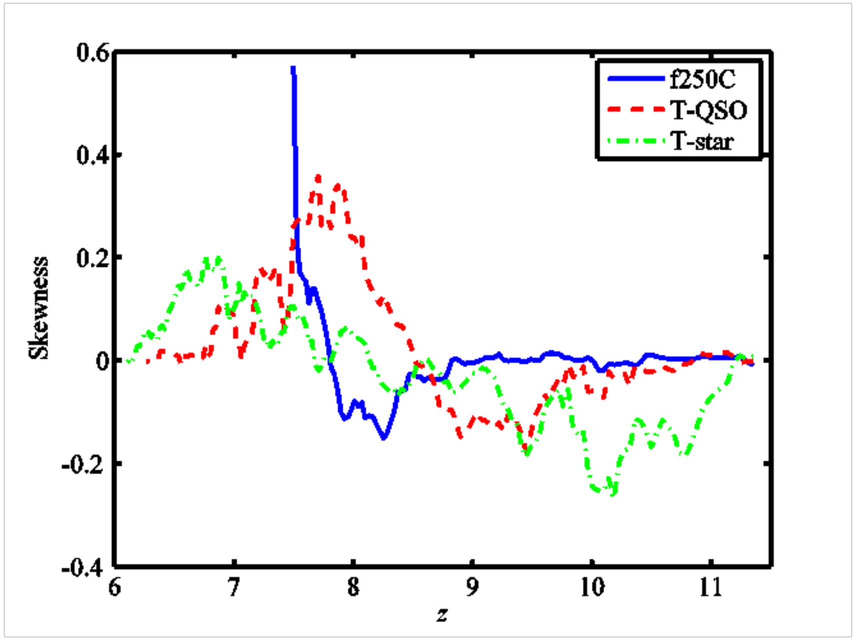


The importance of good foreground fitting

Original simulations

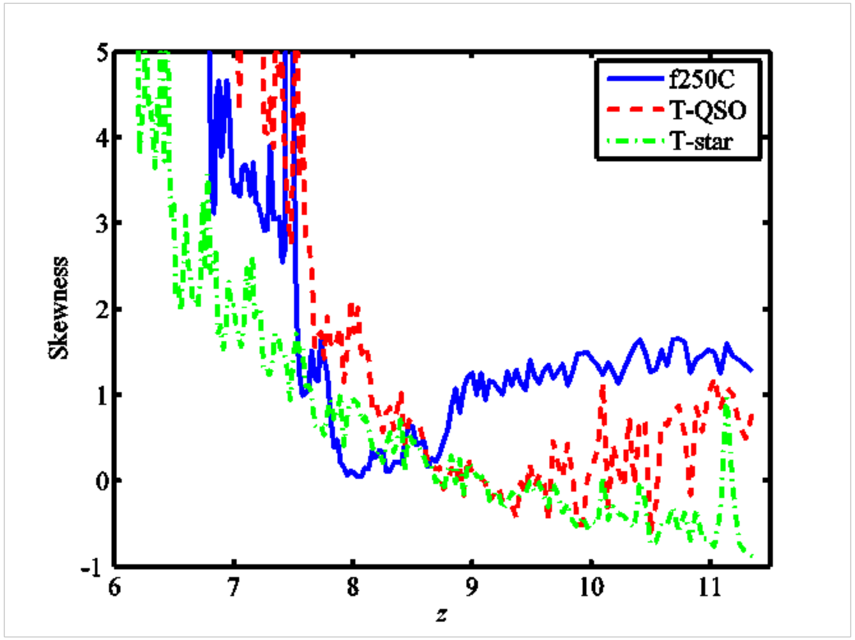


Residuals after Wiener deconvolution

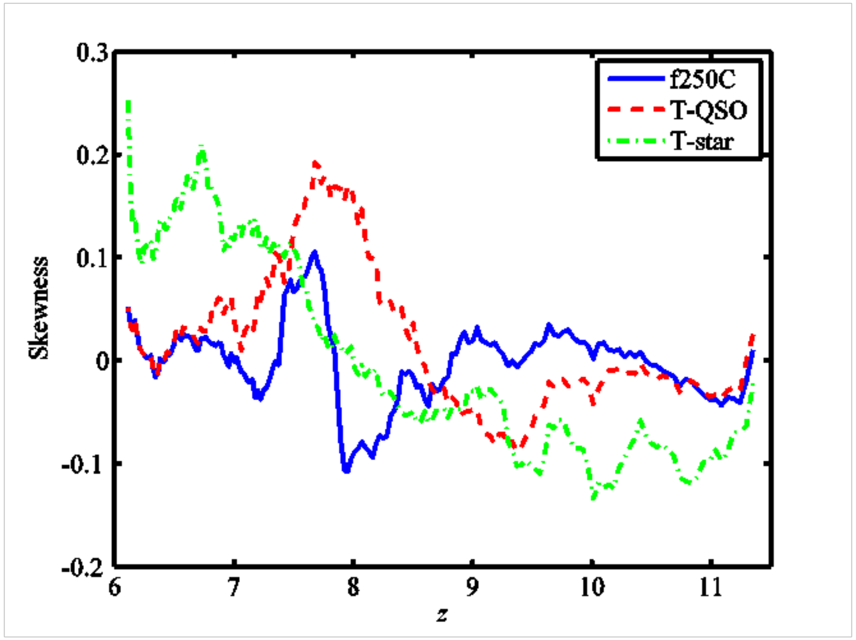


The importance of good foreground fitting

Original simulations



Residuals after perfect foreground subtraction



Wish list for a foreground fitting algorithm

- Accuracy.
- Lack of bias.
- Avoidance of under-fitting or over-fitting.
- Make minimal assumptions about the functional form of the foregrounds; i.e., exploit their smoothness directly.
- Speed (less important if we only wish to subtract the foregrounds once, in post-processing).

Statistical approach

- Model data points (x_i ,y_i) by:

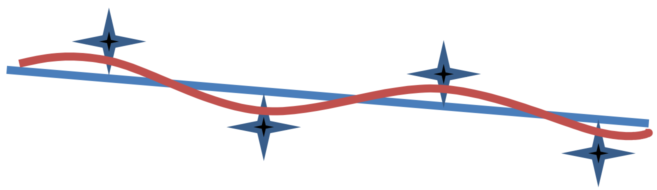
$$y_i = f(x_i) + \varepsilon_i, \;\; i = 1, \ldots, n$$

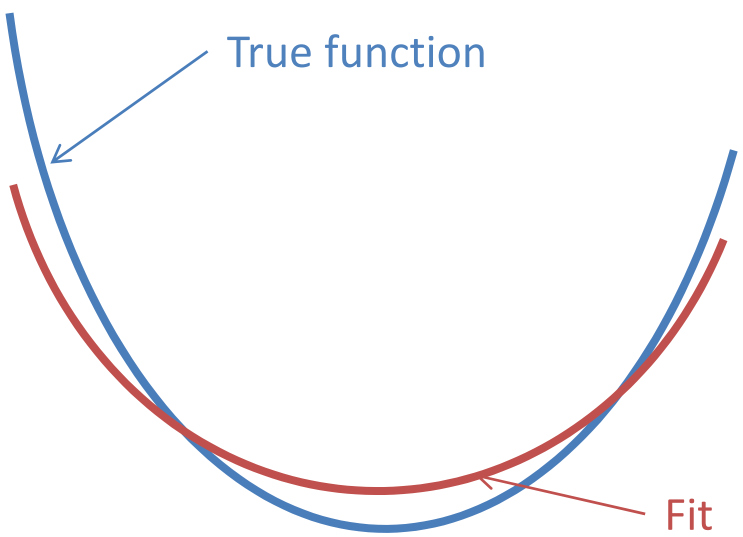
- Then we wish to solve the following problem:

$$\min_f \left\{ \underbrace{\sum_{i=1}^{n} \rho_i(y_i - f(x_i))}_{\text{“Least squares”}} + \underbrace{\lambda R[f]}_{\text{Roughness penalty}} \right\}$$

Choosing a roughness penalty *R*[*f*]

- Require a roughness penalty that stops the curve wiggling towards individual data points, but avoids the problem of attrition.
- 'Smoothing splines' use integrated curvature as the roughness penalty, but in Wp smoothing the integrated *change* of curvature is used instead.





Wp smoothing

- An approximation to the change of curvature, *f'''/f''*, blows up at the inflection points *f''*=0.
- *R*[*f*] measures the change of curvature 'apart from the inflection points', w_i
- Perform the minimization with the position of the inflection points (and s_f) fixed.

$$R[f] = \int_{x_1}^{x_n} h'_f(t)\mathrm{d}t$$

$$f''(x) = p_{\mathbf{w}}(x)e^{h_f(x)}$$

$$\begin{aligned} p_{\mathbf{w}}(x) &= s_f(x-w_1)(x-w_2) \\ &\quad \times \ldots (x-w_{n_w}) \end{aligned}$$

Wp smoothing

- Mächler (1993,1995), who proposed the method, showed that the variational problem leads to the following differential equation:

$$h_f'' = p_{\mathbf{w}} e^{h_f} \left[-\frac{1}{2\lambda} \sum_{i=1}^{n} (x - x_i)_+ \psi_i(y_i - f(x_i)) \right]$$

where a_+ =max(0,a), $\psi_i(\delta) = \frac{\mathrm{d}}{\mathrm{d}\delta}\rho_i(\delta)$, and the boundary conditions are

$$h_f'(x_1) = h_f'(x_n) = \sum_i \psi_i(y_i - f(x_i)) = \sum_i x_i \psi_i(y_i - f(x_i)) = 0$$

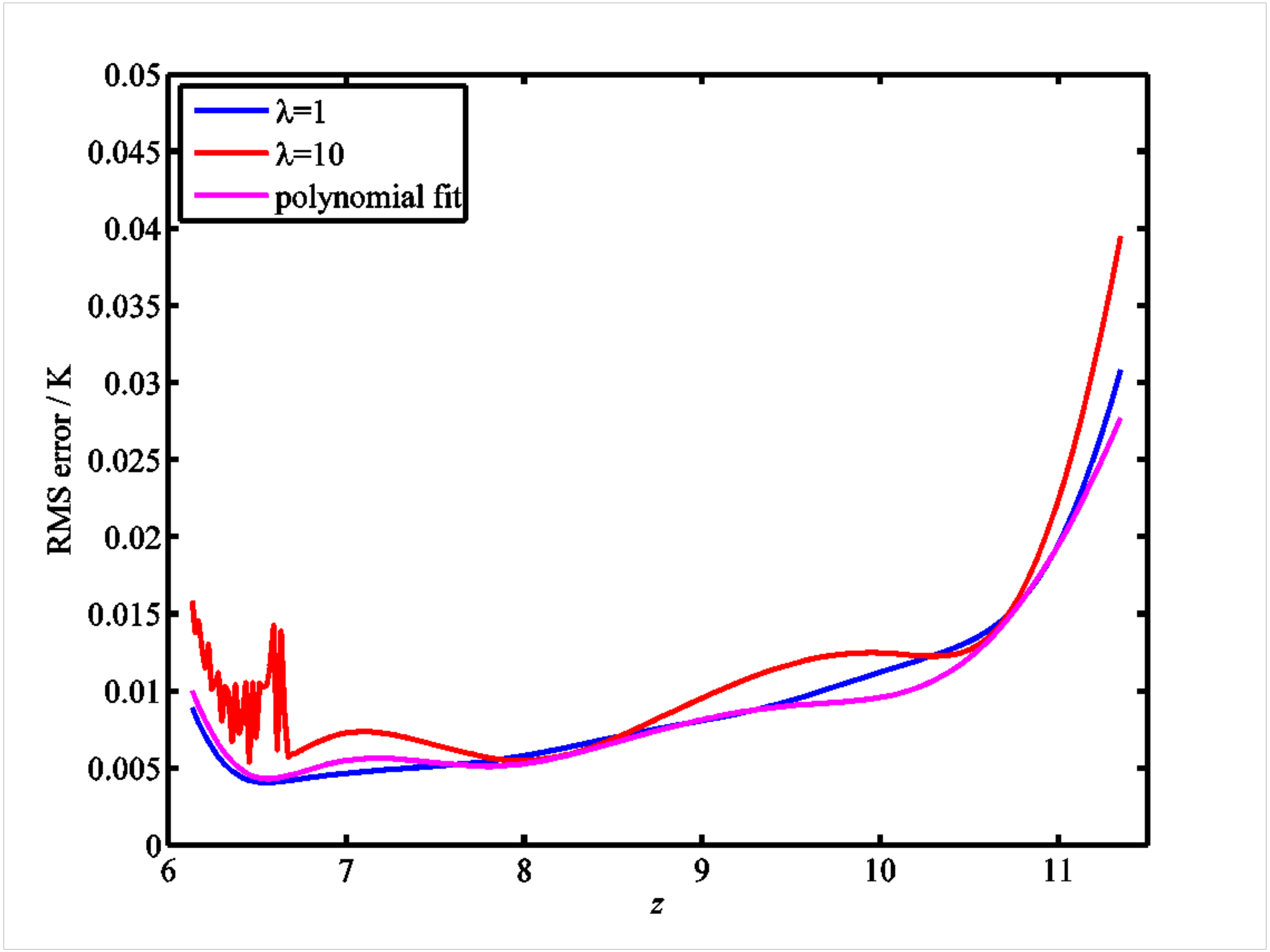
Implementation

- In general we need a method to find the number of inflection points, and need to perform a further minimization over their position.
- For the foreground fitting we find that it works well to have no inflection points (this would be the case anyway for a sum of negative-index power laws).
- The differential equation and the boundary conditions are in a nonstandard form:
  - Can rewrite as a system of 5*n*-4 coupled first-order equations and use a standard BVP solver.
  - Alternatively, convert to a finite difference equation and perform a multidimensional function minimization (seems better so far).
- Either approach requires a reasonable initial guess for the solution; we fit a power law since this has no inflection points.

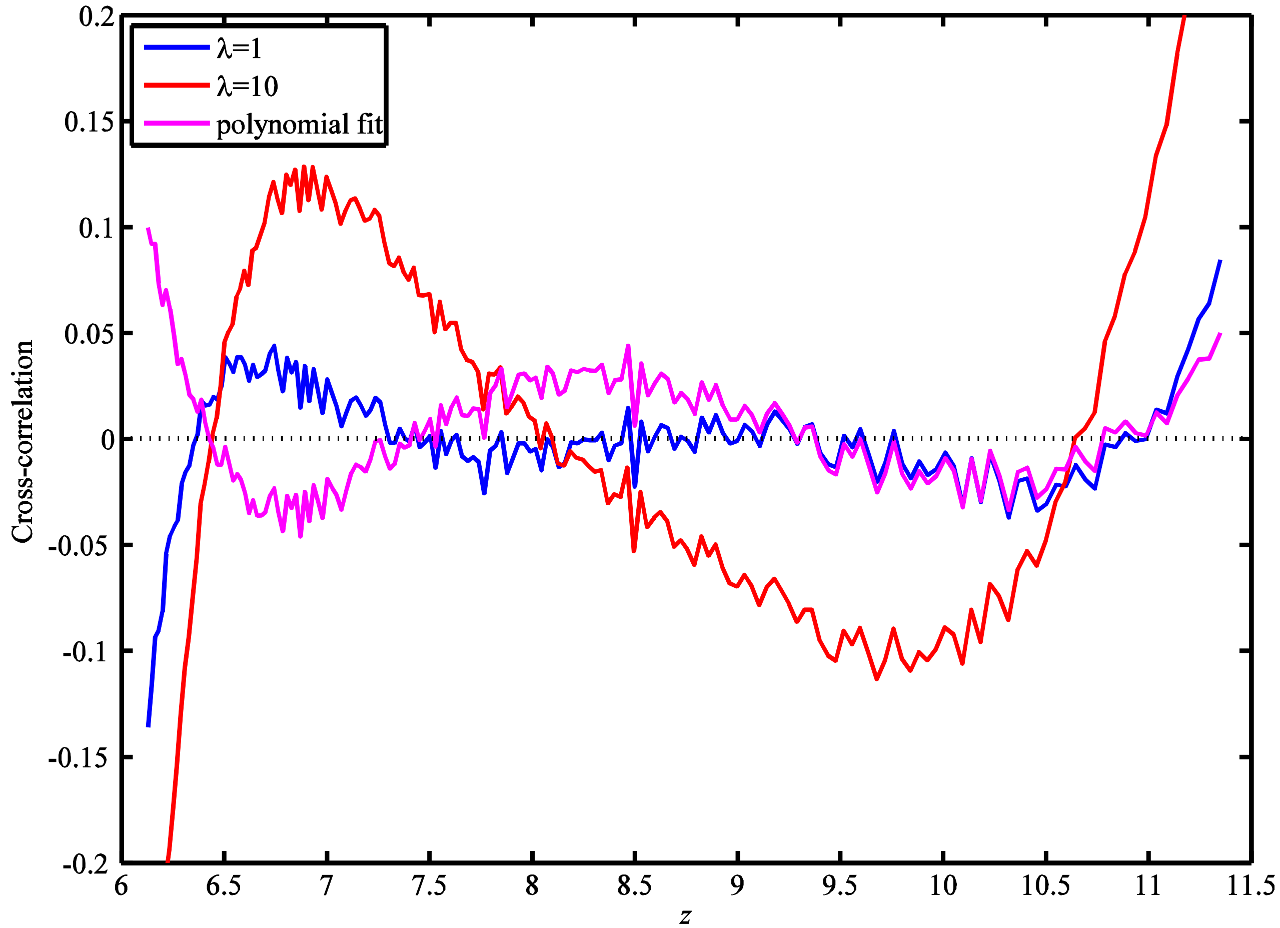
Results

- Approx. 3s of computing time per sightline for 170 points; this depends on the quality of the initial guess.
- *rms* fitting errors small compared to the random noise and comparable to or better than for polynomial or power law fitting (where we have to have assumed a functional form).
- Better cross-correlation properties with the (known, simulated) foregrounds compared to polynomial fitting.

RMS fitting error



Cross-correlation of residuals with foregrounds



Ongoing work

- Find the best value for λ.
- What's the effect of using more or fewer bins?
- Ways to alleviate the problems at the ends of the range (change weighting scheme?); can we deal with gaps?
- Generalize and speed up the Wp algorithm (another use for GPUs?).
- Does the improved foreground fitting allow us to relax the assumptions we make when processing the foreground-subtracted images (e.g. the signal correlation matrix in Wiener deconvolution)?
- Power spectrum estimation; discriminating between models.
- Other statistics.

Conclusions

- Accurate and unbiased foreground fitting is a crucial part of our signal extraction.
- Non-parametric methods do not require us to specify a particular functional form for the foregrounds.
- Wp smoothing, which penalizes the integrated change of curvature (apart from inflection points) is a promising method.
- Implementations are computationally expensive at the moment but not unreasonable.
- We find it gives accurate and unbiased estimates of the simulated foregrounds making only general assumptions about smoothness, especially in the middle of the frequency range.