Hale Collage

Spectropolarimetric Diagnostic Techniques

Rebecca Centeno

March 1 - 8, 2016

Today

- Simple solutions to the RTE

    - Milne-Eddington approximation
    - LTE solution

- The general inversion problem

- Spectral line inversion codes

    - Levenberg-Marquardt techniques
    - Principal Component Analysis

Solutions to the RTE: Milne-Eddington Approximation

Radiative Transfer Equation:

$$\boxed{\frac{d\vec{I}}{d\tau_c} = \mathbf{K}(\vec{I} - \vec{S})}$$

Let's assume:

Polarization due to Zeeman effect.

The elements of **K** are constant: $\mathbf{K} = \mathbf{K_0}$

$\vec{S} = (S_0 + S_1\tau)\ (1, 0, 0, 0)^T$

Then:

$$\vec{I}(0) = \int_0^\infty e^{-\mathbf{K_0}\tau_c}\mathbf{K_0}(\vec{S}_0 + \vec{S}_1\tau_c)d\tau_c = \vec{S}_0 + \mathbf{K_0}^{-1}\vec{S}_1$$

is the Unno-Rachkovsky solution to the RTE and it is analytical in nature!

$$I(0) = S_0 + \Delta^{-1}\eta_I(\eta_I^2 + \rho_Q^2 + \rho_U^2 + \rho_V^2)S_1$$

$$Q(0) = -\Delta^{-1}\left[\eta_I^2\eta_Q + \eta_I(\eta_V\rho_U - \eta_U\rho_V) + \rho_Q(\eta_Q\rho_Q + \eta_U\rho_U + \eta_V\rho_V)\right]S_1$$

$$U(0) = -\Delta^{-1}\left[\eta_I^2\eta_U + \eta_I(\eta_Q\rho_V - \eta_V\rho_Q) + \rho_U(\eta_Q\rho_Q + \eta_U\rho_U + \eta_V\rho_V)\right]S_1$$

$$U(0) = -\Delta^{-1}\left[\eta_I^2\eta_V + \eta_I(\eta_U\rho_Q - \eta_Q\rho_U) + \rho_V(\eta_Q\rho_Q + \eta_U\rho_U + \eta_V\rho_V)\right]S_1$$

with:

$$\Delta = \eta_I^2(\eta_I^2 - \eta_Q^2 - \eta_U^2 - \eta_V^2 + \rho_Q^2 + \rho_U^2 + \rho_V^2) - (\eta_Q\rho_Q + \eta_U\rho_U + \eta_V\rho_V)^2$$

Solutions to the RTE: Milne-Eddington Approximation

Model Parameters:

Line-to-continuum absorption: η_0
Doppler width: $\Delta\lambda_D$
Damping parameter: a
Magnetic field: B, θ, ϕ
Source function: S_0, S_1
LOS velocity: v_{LOS}

Magnetic filling factor: α
Macroturbulent velocity: v_{MAC}

No thermodynamical information

No velocity gradients so no asymmetries
NCP = 0

No magnetic field gradients

If you want to synthesize spectral lines in a Milne-Eddington atmosphere:

http://www.iac.es/proyecto/inversion/online/milne_code/milne.php

Solutions to the RTE: Local Thermodynamical Equilibrium

LTE hypothesis: the plasma is in thermodynamic equilibrium at local values of temperature and density. Hence:

- Maxwellian distribution of velocities
- Saha and Boltzmann give the populations of different atomic species
- Kirchhoff's Law applies: $\vec{j} = B_\nu(T)\,(\eta_I, \eta_Q, \eta_U, \eta_V)^T$

Absorption profiles have the same shape as emission profiles ⇒ complete redistribution of frequencies

- Also assumes hydrostatic equilibrium and Zeeman induced polarization.

And the RTE still looks like this:

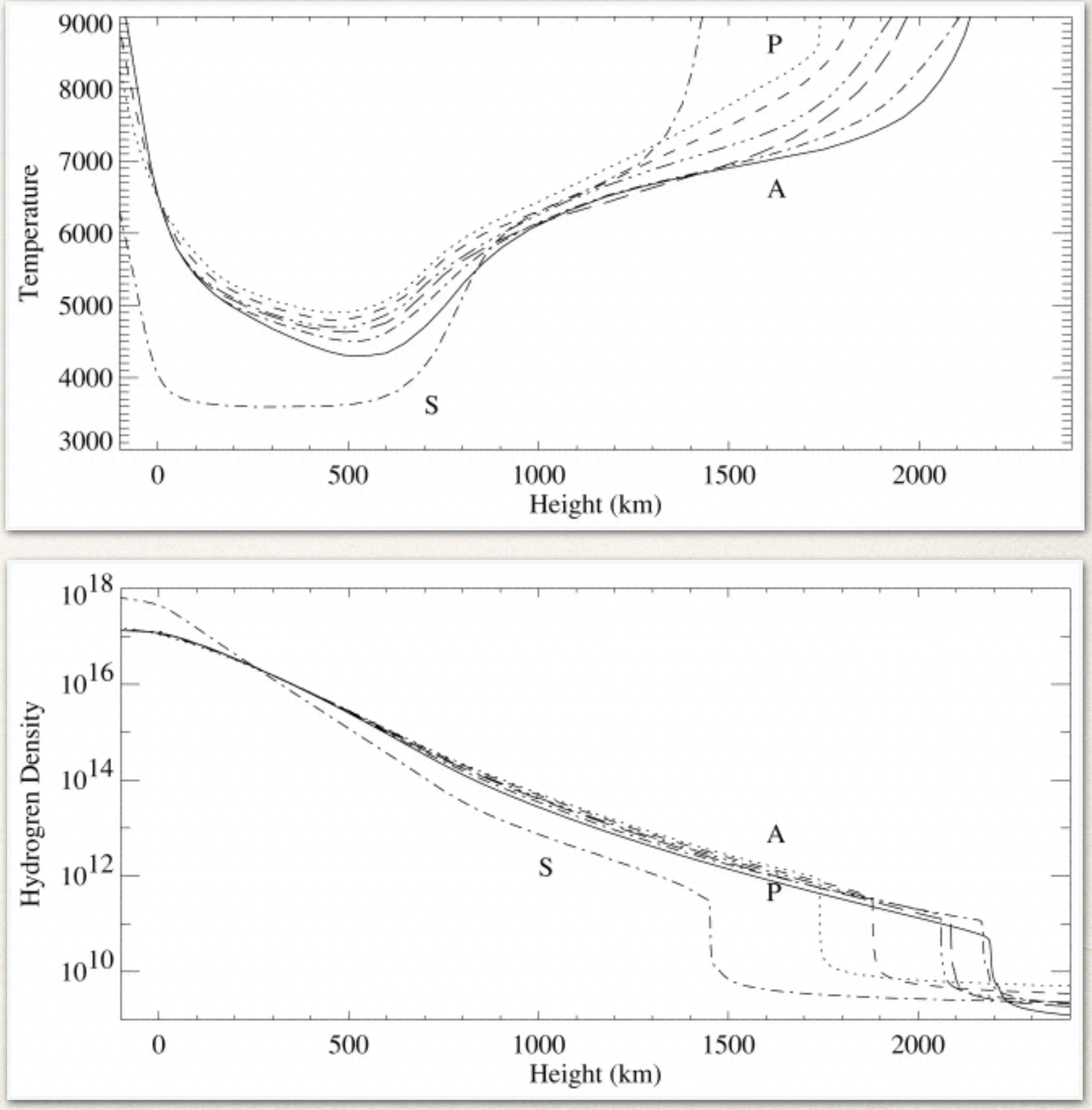
$$\frac{d}{dz}\begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = -\begin{pmatrix} \eta_I & \eta_Q & \eta_U & \eta_V \\ \eta_Q & \eta_I & \rho_V & -\rho_U \\ \eta_U & -\rho_V & \eta_I & \rho_Q \\ \eta_V & \rho_U & -\rho_Q & \eta_I \end{pmatrix}\begin{pmatrix} I - B_\nu(T) \\ Q \\ U \\ V \end{pmatrix}$$

Solutions to the RTE: Local Thermodynamical Equilibrium

Model atmosphere:

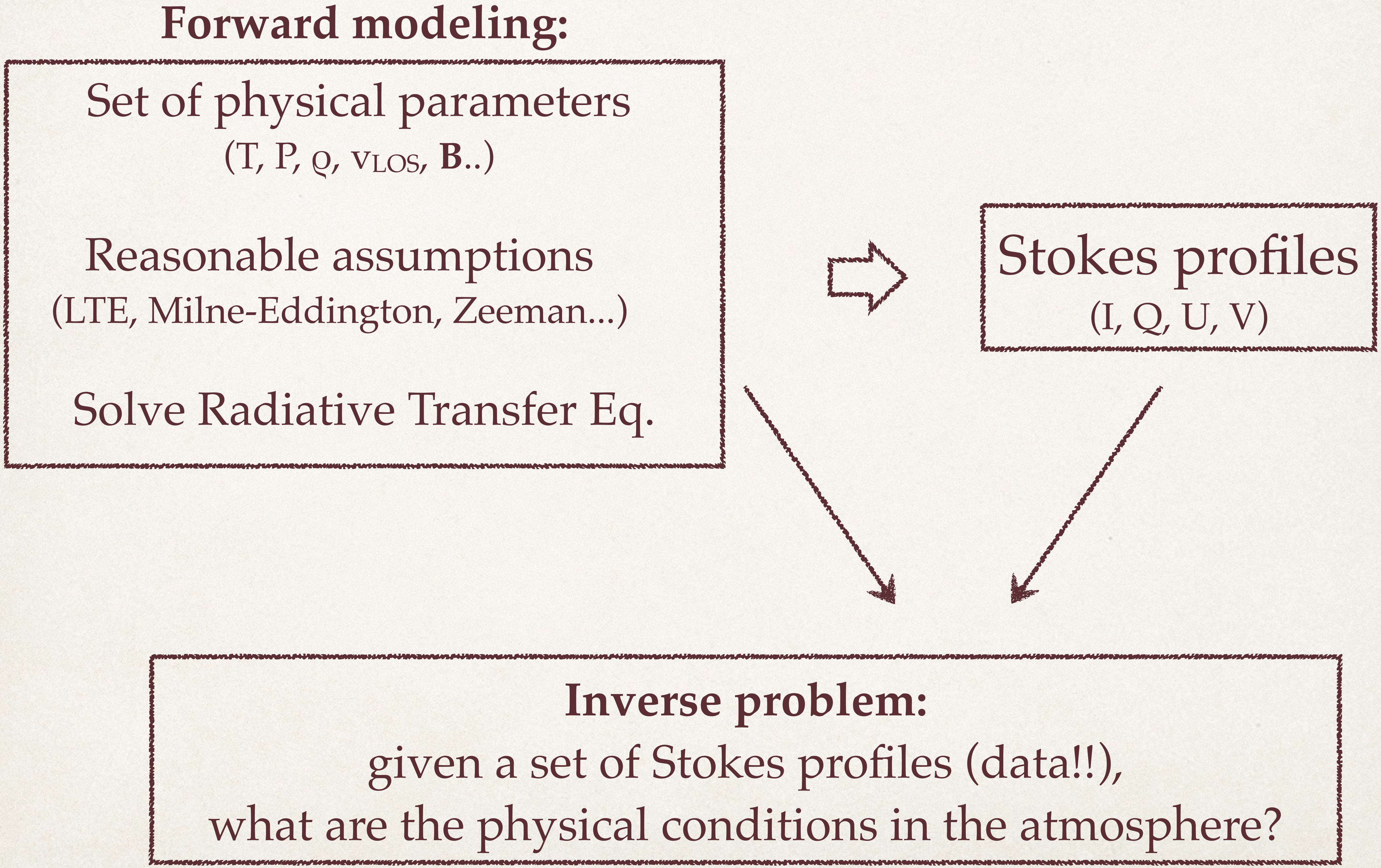
temperature: T(z)
pressure: P(z)
LOS velocity: $v_{LOS}(z)$
Magnetic field: B(z), θ(z), ϕ(z)
Macroturbulent velocity: v_{MAC}
Macroturbulent velocity: v_{MIC}

- 9 physical quantities
- stratified atmosphere:
  9 x Nz free model parameters
- Too many free parameters!
- Need to constrain stratification to a limited number of nodes



(from Fontenla et al 1999)

Spectral Line Inversions



Forward modeling:

Set of physical parameters

(T, P, ϱ, v_{LOS}, **B**..)

Reasonable assumptions

(LTE, Milne-Eddington, Zeeman...)

Solve Radiative Transfer Eq.

Stokes profiles

(I, Q, U, V)

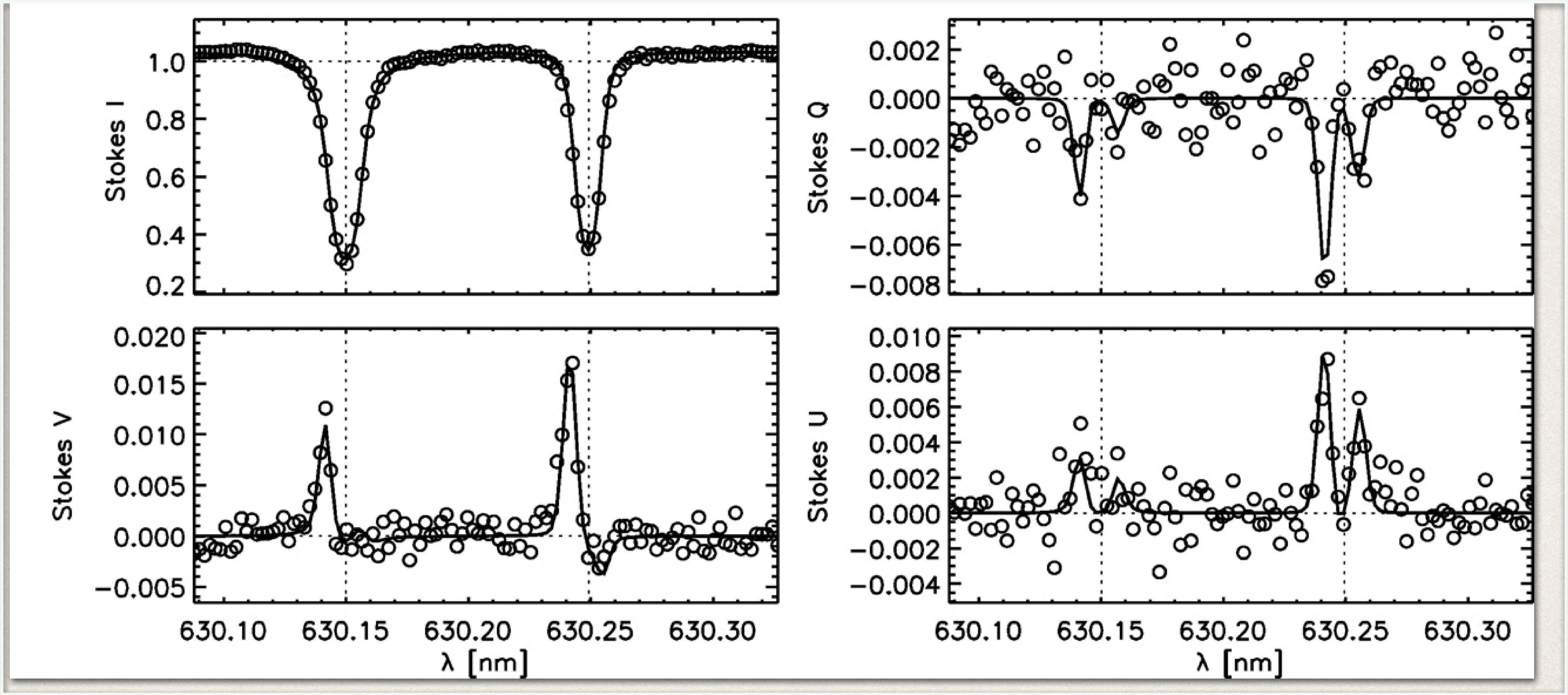
Inverse problem:

given a set of Stokes profiles (data!!),

what are the physical conditions in the atmosphere?

Spectral Line Inversions

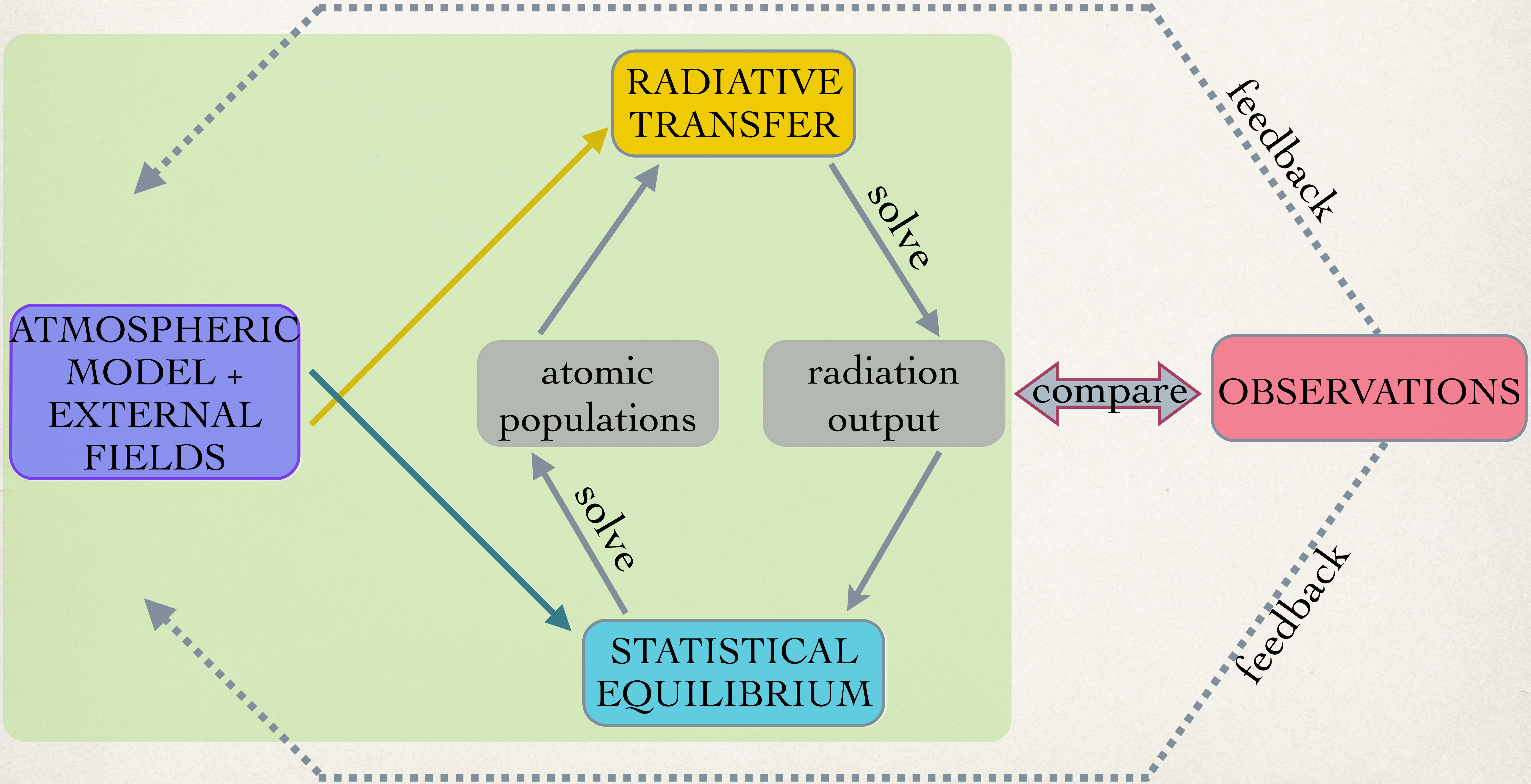
fitting metric +
educated guess



(from Vitticchié et al, 2011)

open circles: observations
solid line: synthetic Stokes profiles

Spectral Line Inversions: General Problem



ATMOSPHERIC MODEL + EXTERNAL FIELDS

RADIATIVE TRANSFER

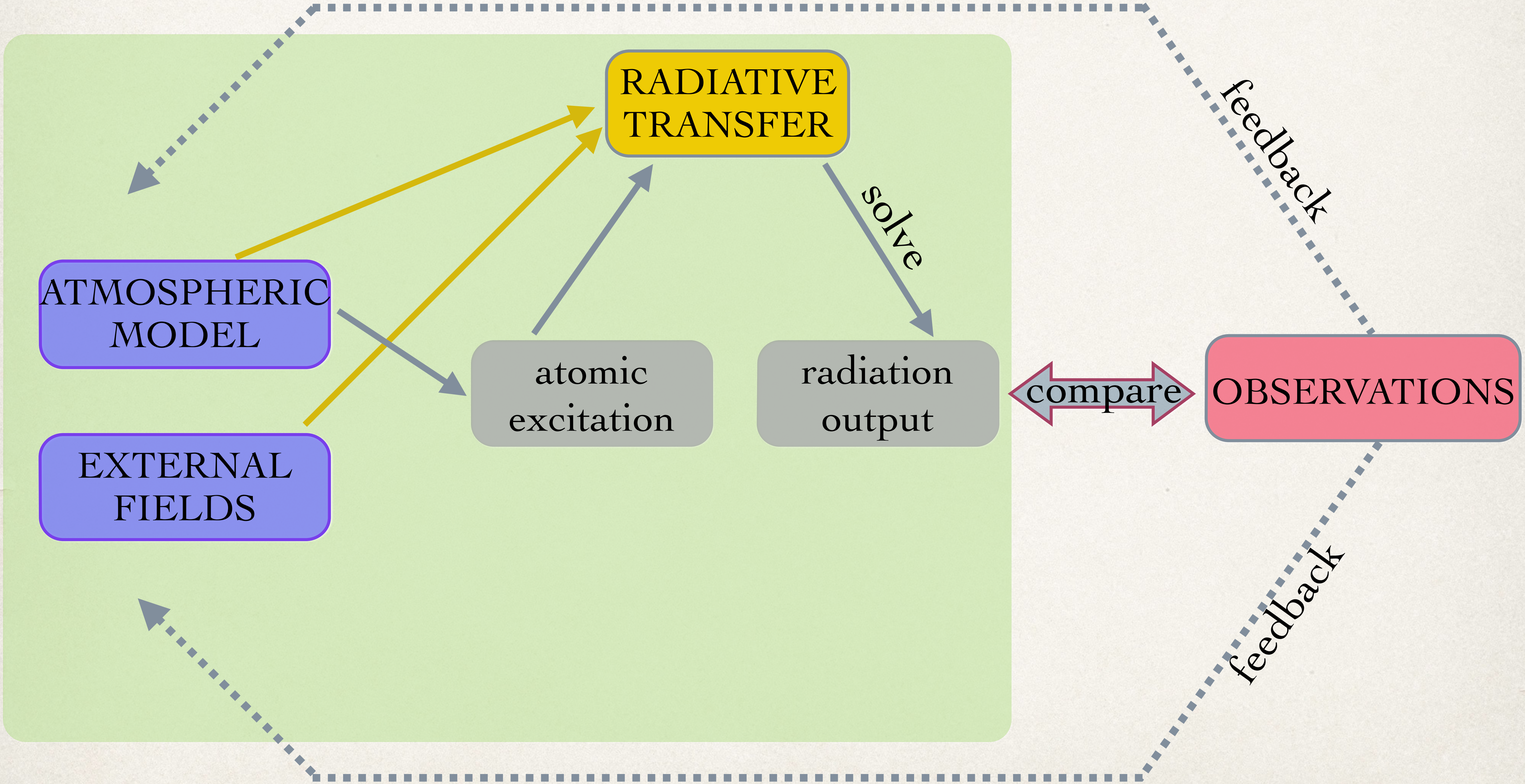
atomic populations

radiation output

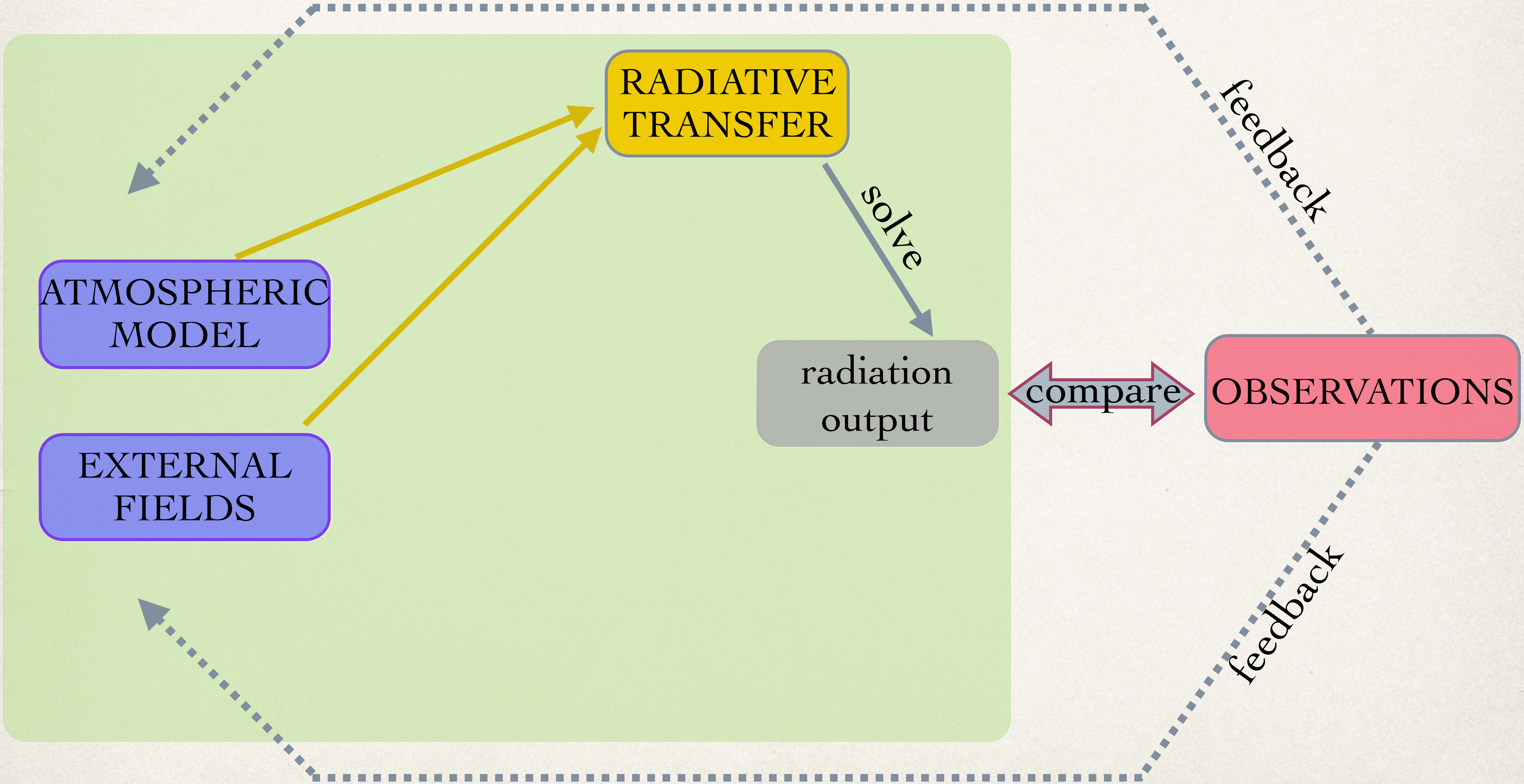
STATISTICAL EQUILIBRIUM

OBSERVATIONS

Spectral Line Inversions: Local Thermodynamic Equilibrium



Spectral Line Inversions: Milne-Eddington Approximation



RADIATIVE TRANSFER

ATMOSPHERIC MODEL

EXTERNAL FIELDS

solve

radiation output

compare

OBSERVATIONS

feedback

feedback

Spectral Line Inversion Methods

Let's assume we know how to solve the RTE.

Inversion methods

Levenberg-Marquardt methods
(least squares fitting)

Principal Component Analysis techniques
(pattern recognition)

Spectral Line Inversions: the merit function

Let's assume our model atmosphere is characterized by a series of N_p parameters, **a**.

The solution to the RTE in the model **a** gives us a set of synthetic Stokes profiles, which we can compare to the observed ones. We can measure the difference using a merit function:

$$\chi^2 = \frac{1}{N_f} \sum_s \sum_\lambda \left[I_s^{\rm obs}(\lambda) - I_s^{\rm syn}(\lambda) \right]^2 \omega_s^2$$

Where the number of degrees of freedom: $N_f = N_s \times N_\lambda - N_p$

I^{syn} and I^{obs} are the synthetic and observed Stokes profiles

ω_s are some weighting factors (related to measurement error).

The sums are over wavelength and Stokes parameters.

Spectral Line Inversions: the merit function

$$\chi^2 = \frac{1}{N_f} \sum_s \sum_\lambda \left[I_s^{\text{obs}}(\lambda) - I_s^{\text{syn}}(\lambda) \right]^2 \omega_s^2$$

χ^2 is a hyper-surface of N_p dimensions.

It quantifies the goodness of the fit
(the distance between the observed and synthetic Stokes vector)
with one number!

The whole inversion problem boils down
to minimizing χ^2

Spectral Line Inversions: Levenberg-Marquardt Techniques

The problem boils down to the minimization of χ^2.

The first derivative is given by:

$$\frac{\partial \chi^2}{\partial a_i} = \frac{2}{N_f} \sum_s \sum_\lambda \left[I_s^{\mathrm{syn}}(\lambda) - I_s^{\mathrm{obs}}(\lambda)\right] \omega_s^2 \frac{\partial I^{\mathrm{syn}}(\lambda)}{\partial a_i} \qquad (\mathrm{i} = 0, \ldots, \mathrm{N_p})$$

And the second derivative of χ^2 is given by:

$$\frac{\partial^2 \chi^2}{\partial a_i \partial a_j} = \frac{2}{N_f} \sum_s \sum_\lambda \omega_s^2 \left(\frac{\partial I_s^{\mathrm{syn}}(\lambda)}{\partial a_j} \frac{\partial I_s^{\mathrm{syn}}(\lambda)}{\partial a_i} + \left[I_s^{\mathrm{syn}}(\lambda) - I_s^{\mathrm{obs}}(\lambda)\right] \frac{\partial^2 I^{\mathrm{syn}}(\lambda)}{\partial a_i \partial a_j} \right)$$

When close to the minimum of χ^2, we can expect $[I^{\mathrm{syn}} - I^{\mathrm{obs}}] \simeq 0$

$$\frac{\partial^2 \chi^2}{\partial a_i \partial a_j} \approx \frac{2}{N_f} \sum_s \sum_\lambda \omega_s^2 \left(\frac{\partial I_s^{\mathrm{syn}}(\lambda)}{\partial a_j} \frac{\partial I_s^{\mathrm{syn}}(\lambda)}{\partial a_i} \right)$$

Spectral Line Inversions: Levenberg-Marquardt Techniques

Let's assume the model **a** is close to the minimum of χ^2, so there is a perturbation **δa** that takes us directly to the minimum.
We can use a quadratic approximation, such that:

$$\chi^2(\overset{\substack{\mathbf{a}_{\text{min}}\\ \downarrow}}{\mathbf{a}+\delta\mathbf{a}}) \approx \chi^{\mathbf{2}}(\overset{\substack{\mathbf{a}_{\text{current}}\\ \downarrow}}{\mathbf{a}}) + \delta\mathbf{a}^{\mathbf{T}}(\nabla\chi^{\mathbf{2}} + \mathbf{H}'\delta\mathbf{a})$$

where

$$H'_{ij} = \frac{1}{2}\frac{\partial^2\chi^2}{\partial a_i \partial a_j}$$

is half the Hessian matrix
(dimensions N_p x N_p)

Spectral Line Inversions: Levenberg-Marquardt Techniques

When one is really close to the minimum, the **second order approximation** is adequate, and we can equate to zero the term in parenthesis:

$$(\nabla\chi^2 + \mathbf{H}'\delta\mathbf{a}) = 0 \longrightarrow \delta\mathbf{a} = -\mathbf{H}'^{-1}\nabla\chi^{\mathbf{2}}$$

This involves inverting the Hessian matrix!!

thus we obtain a better approximation to the minimum of χ^2 by shifting in the parameter space an amount **δa.**

When we're far from the minimum, we can get closer to it following the gradient (**first order approximation**):

$$\delta\mathbf{a} = k\nabla\chi^2$$

with k small enough!

Spectral Line Inversions: Levenberg-Marquardt Techniques

Marquardt had two insights:

- The diagonal elements of the Hessian matrix give us a sense of what good values for k could be (it's a dimensional argument).

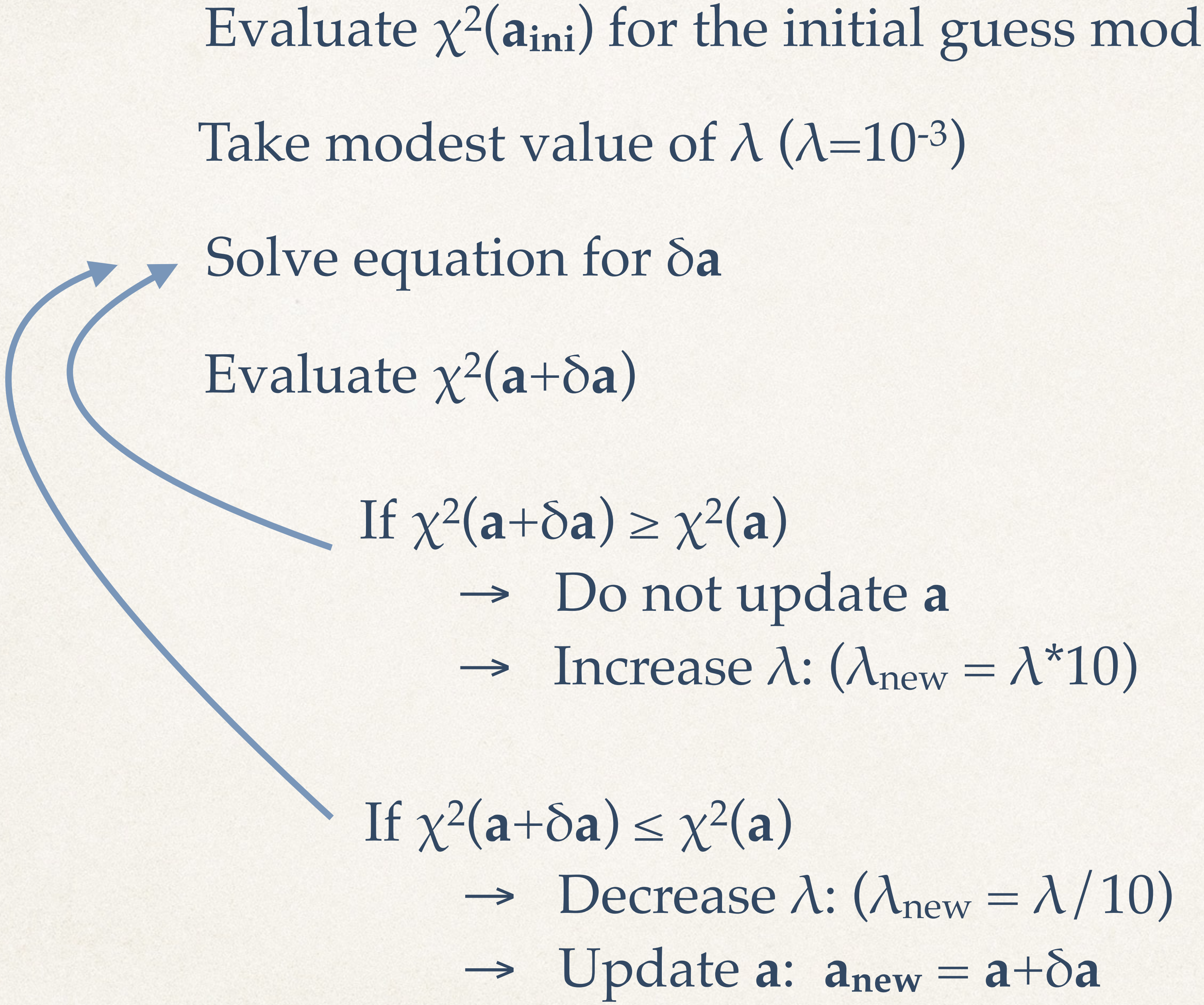
$$\delta a_i = -\overset{=k}{\boxed{\frac{1}{H'_{ii}}}}\nabla\chi^2 \quad \xrightarrow{\text{fudge factor } \lambda} \quad \delta a_i = -\overset{=k}{\boxed{\frac{1}{\lambda H'_{ii}}}}\nabla\chi^2$$

- The two methods can be combined into one equation, that allows to vary smoothly between the gradient and the Hessian approaches:

$$\nabla\chi^2 + \mathbf{H}\delta\mathbf{a} = \mathbf{0} \qquad \text{where} \qquad H_{ij} \equiv \begin{cases} (1+\lambda)H'_{ij} & \text{if } i = j \\ H'_{ij} & \text{if } i \neq j \end{cases}$$

$$\begin{aligned} \lambda\uparrow\uparrow \quad &\Rightarrow \quad \text{gradient (first order) method} \\ \lambda\downarrow\downarrow \quad &\Rightarrow \quad \text{hessian (second order) method} \end{aligned}$$

Spectral Line Inversions: Levenberg-Marquardt Techniques



Evaluate $\chi^2(\mathbf{a}_{\mathbf{ini}})$ for the initial guess model

Take modest value of λ ($\lambda=10^{-3}$)

Solve equation for $\delta\mathbf{a}$

Evaluate $\chi^2(\mathbf{a}+\delta\mathbf{a})$

If $\chi^2(\mathbf{a}+\delta\mathbf{a}) \geq \chi^2(\mathbf{a})$

→ Do not update **a**

→ Increase λ: ($\lambda_{\text{new}} = \lambda*10$)

If $\chi^2(\mathbf{a}+\delta\mathbf{a}) \leq \chi^2(\mathbf{a})$

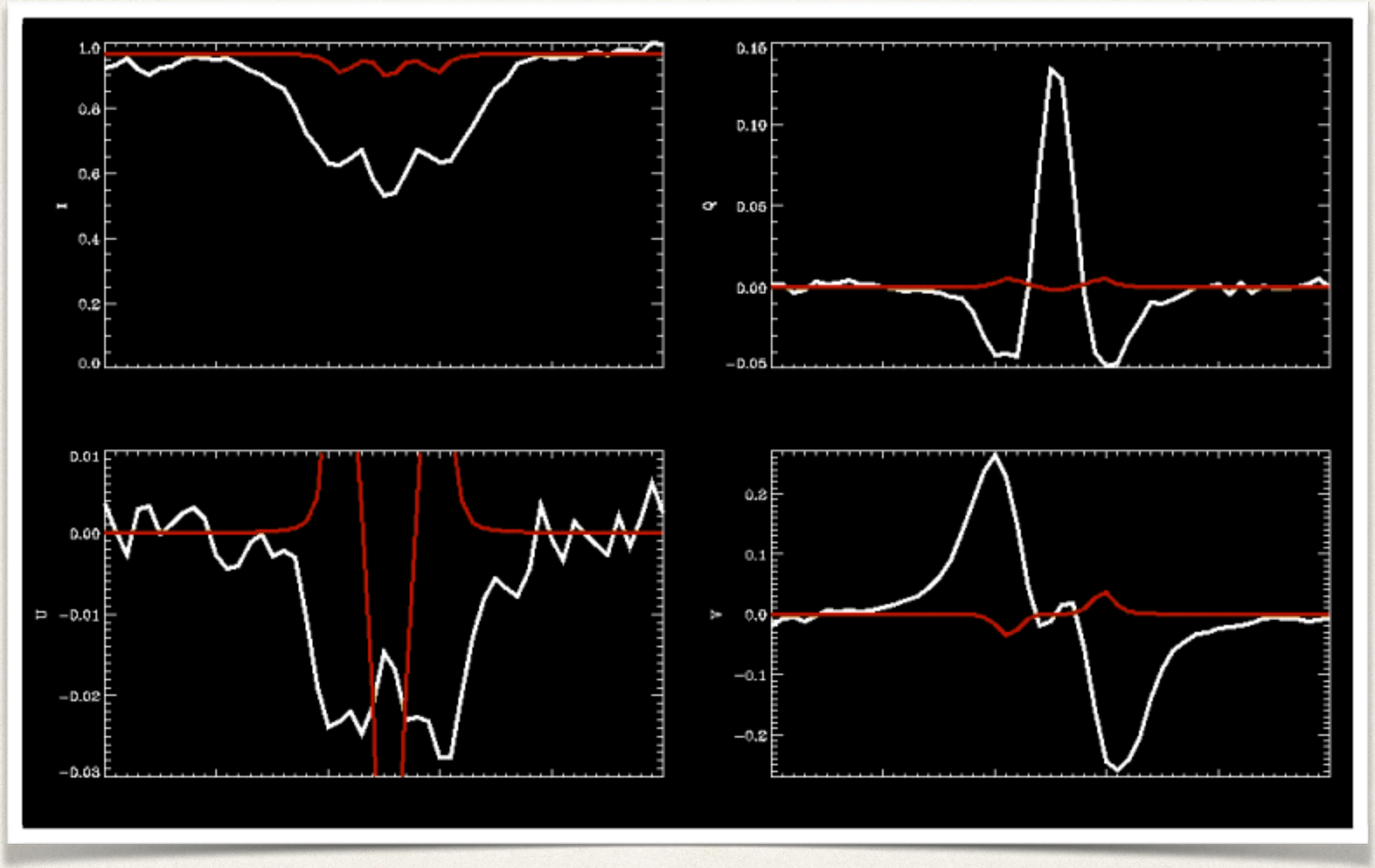
→ Decrease λ: ($\lambda_{\text{new}} = \lambda/10$)

→ Update **a**: $\mathbf{a}_{\mathbf{new}} = \mathbf{a}+\delta\mathbf{a}$

Stop when χ^2 barely decreases once or twice in a row.

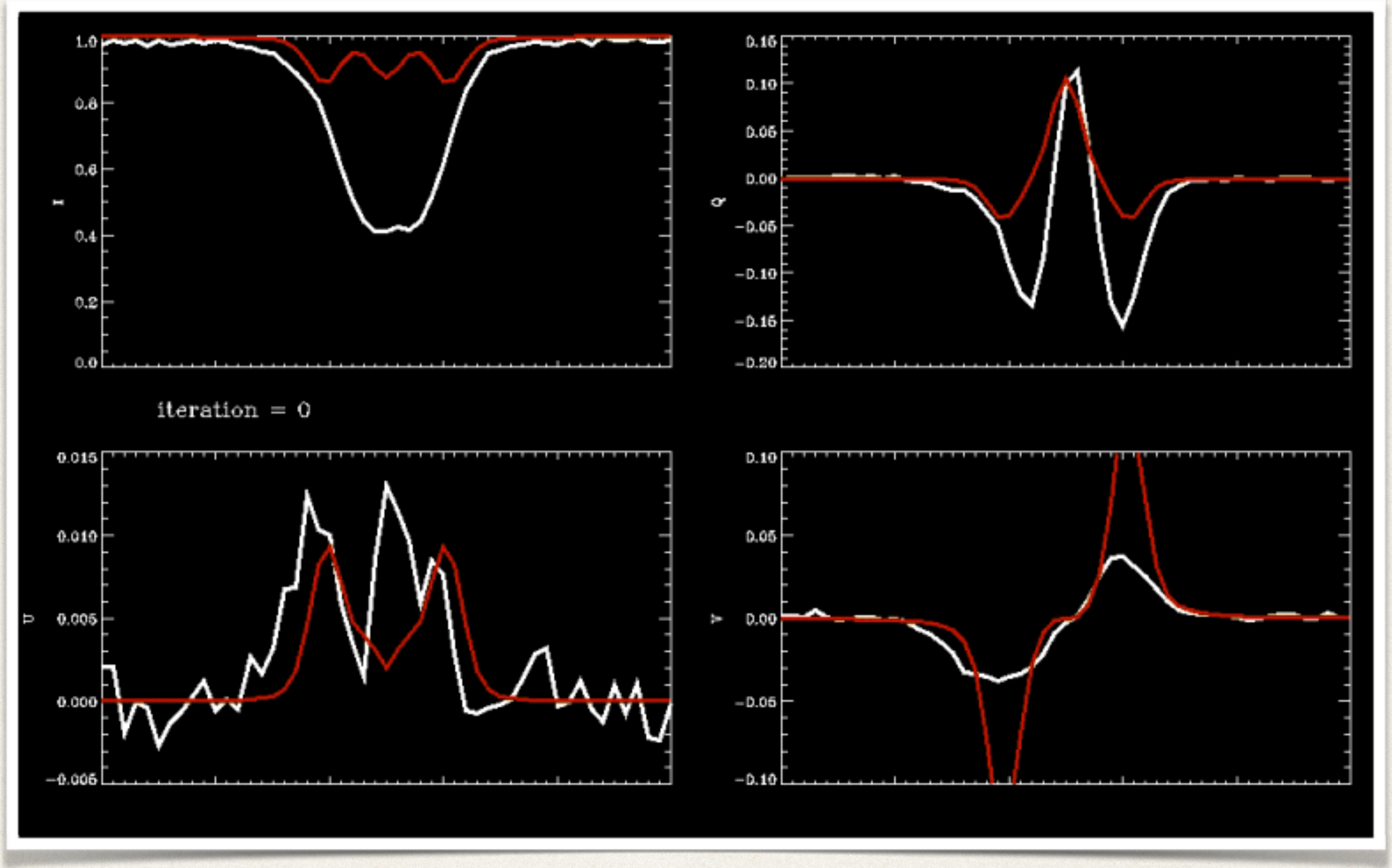
This algorithm is explained in detail in "Numerical Recipes" by Press et al. !

Spectral Line Inversions: Levenberg-Marquardt Techniques



white = observations red = synthetic fit

Spectral Line Inversions: Levenberg-Marquardt Techniques



white = observations red = synthetic fit

Levenberg-Marquardt Techniques: Issues

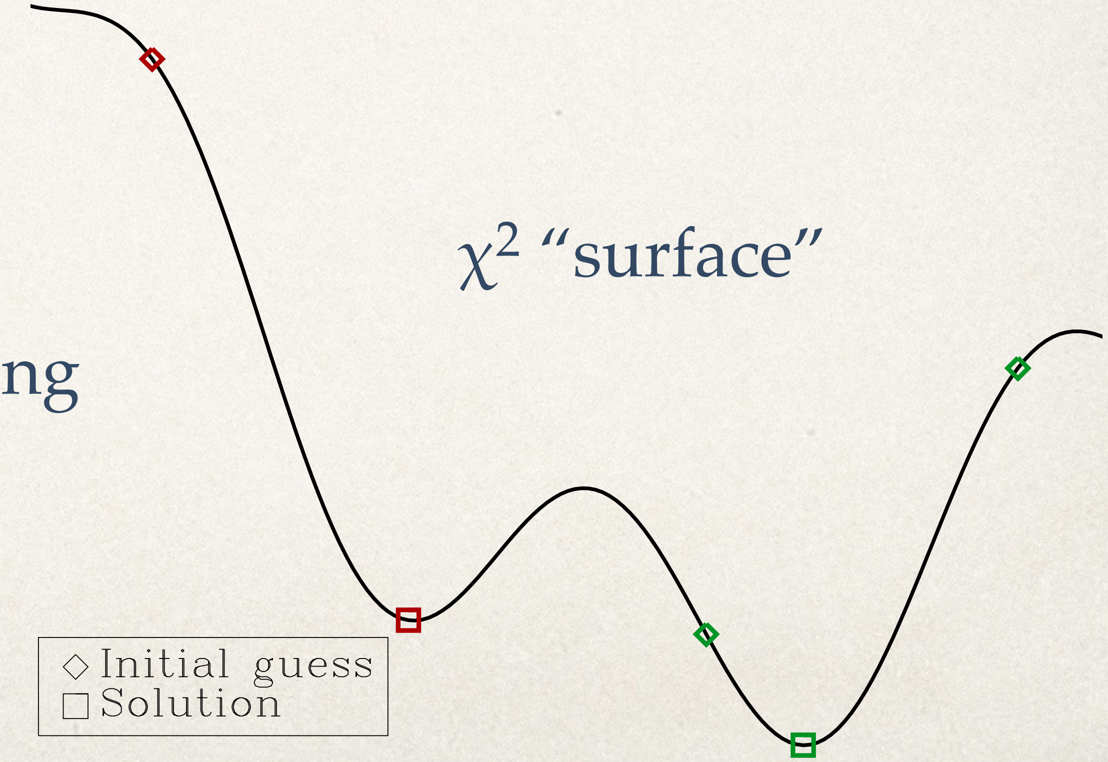
H can be quasi-singular due to different sensitivity of χ^2 to the various model parameters. But it has to be inverted!
Singular Value Decomposition methods:
H is real and symmetric ⇒ $\exists\ \mathbf{Y}$ such that: $\mathbf{H} = \mathbf{Y}^T\,\mathbf{W}\,\mathbf{Y}$ and $\mathbf{Y}\mathbf{Y}^T = \mathbf{Y}^T\mathbf{Y} = \mathbf{1}$

$$\text{So } \mathbf{H}^{-1} = \mathbf{Y}^T\,\mathbf{W}^{-1}\,\mathbf{Y}$$

If $W_k \downarrow\downarrow \Rightarrow$ we set $1/W_k = 0$, so a_k does not contribute to the model perturbation.

Global vs. local minima of χ^2

Levenberg-Marquardt techniques can lead to local (rather than global) minima depending on the location of the initial guess



Principal Component Analysis (PCA) Techniques

Let's assume we have a set of observations

$$\mathbf{S}_{ij} = S_j(\lambda_i) \ , \qquad i = 1, \ldots, N \ ; \quad j = 1, \ldots, M$$

S — Stokes vector (I,Q,U,V)
N — wavelengths
M — pixels

They are independent realizations of the Stokes profile S(λ), so the average:

$$\bar{S}(\lambda_i) = \frac{1}{M} \sum_{j=1}^{M} \mathbf{S}_{ij} \ , \qquad i = 1, \ldots, N$$

We can define a covariance matrix:

$$\mathbf{C}_{ij} = \sum_{l=1}^{M} [\mathbf{S}_{il} - \bar{S}(\lambda_i)][\mathbf{S}_{jl} - \bar{S}(\lambda_j)] \ , \qquad i, j = 1, \ldots, N$$

real and symmetric!

Which can be diagonalized by an orthogonal transformation:

$$\mathbf{C}\boldsymbol{f}^{(k)} = e^{(k)} \boldsymbol{f}^{(k)} \ , \qquad k = 1, \ldots, N$$

where $f^{(k)}$ are the eigenvectors that form a basis for the residual $S_j(\lambda)$ - $\bar{S}(\lambda)$

So that:

$$S_j(\lambda) - \bar{S}(\lambda) \doteq \sum_{k=1}^{N} c_j^{(k)} \boldsymbol{f}^{(k)}$$

Principal Component Analysis (PCA) Techniques

average

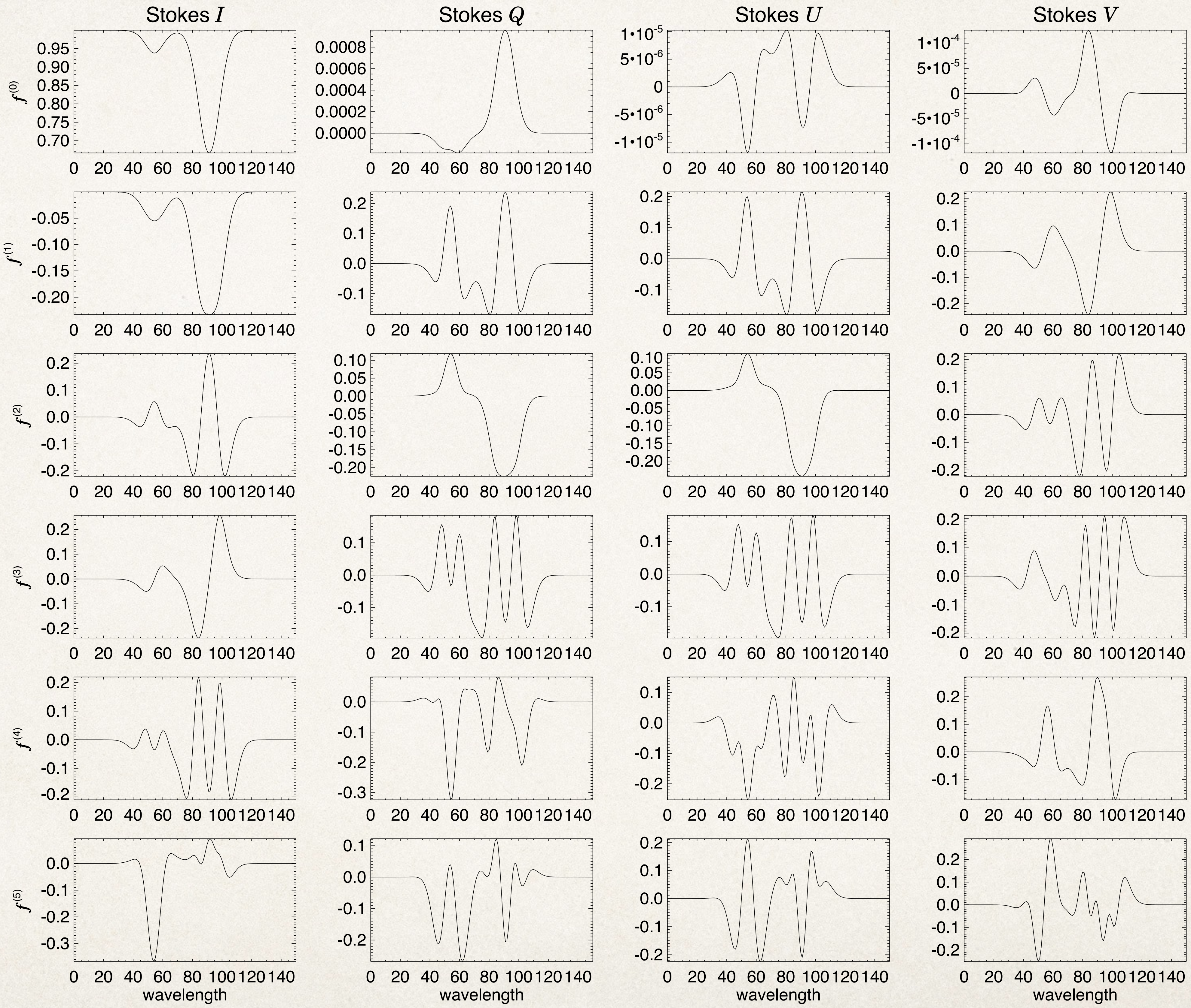
1st PC

2nd PC

3rd PC

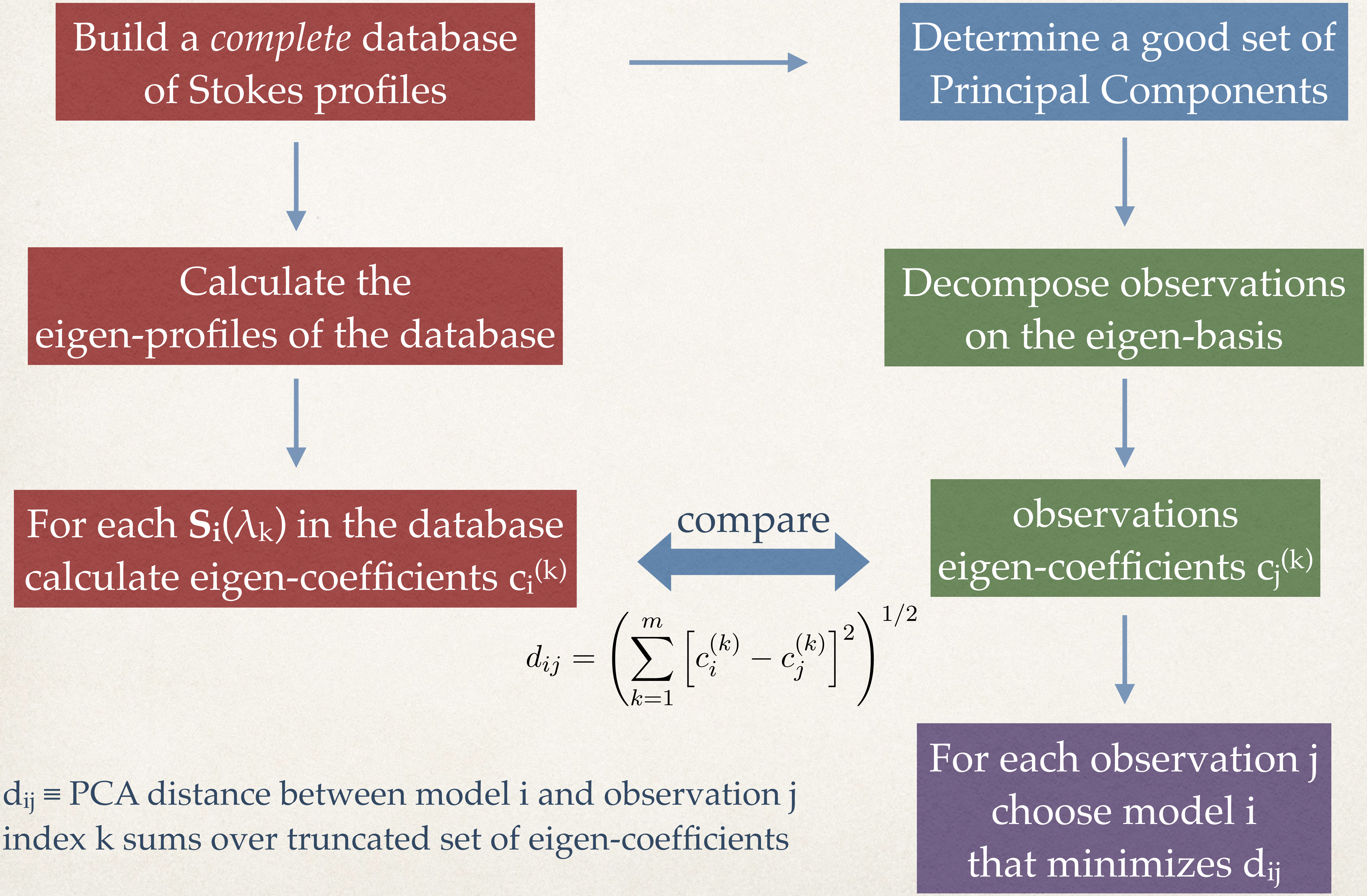
4th PC

5th PC



Casini et al 2012

Principal Component Analysis (PCA) Techniques



Build a *complete* database of Stokes profiles

Determine a good set of Principal Components

Calculate the eigen-profiles of the database

Decompose observations on the eigen-basis

For each $\mathbf{S}_i(\lambda_k)$ in the database calculate eigen-coefficients $c_i^{(k)}$

compare

observations eigen-coefficients $c_j^{(k)}$

$$d_{ij} = \left(\sum_{k=1}^{m} \left[c_i^{(k)} - c_j^{(k)} \right]^2 \right)^{1/2}$$

For each observation j choose model i that minimizes d_{ij}

d_{ij} ≡ PCA distance between model i and observation j
index k sums over truncated set of eigen-coefficients

Principal Component Analysis: Pros and Cons

Pros

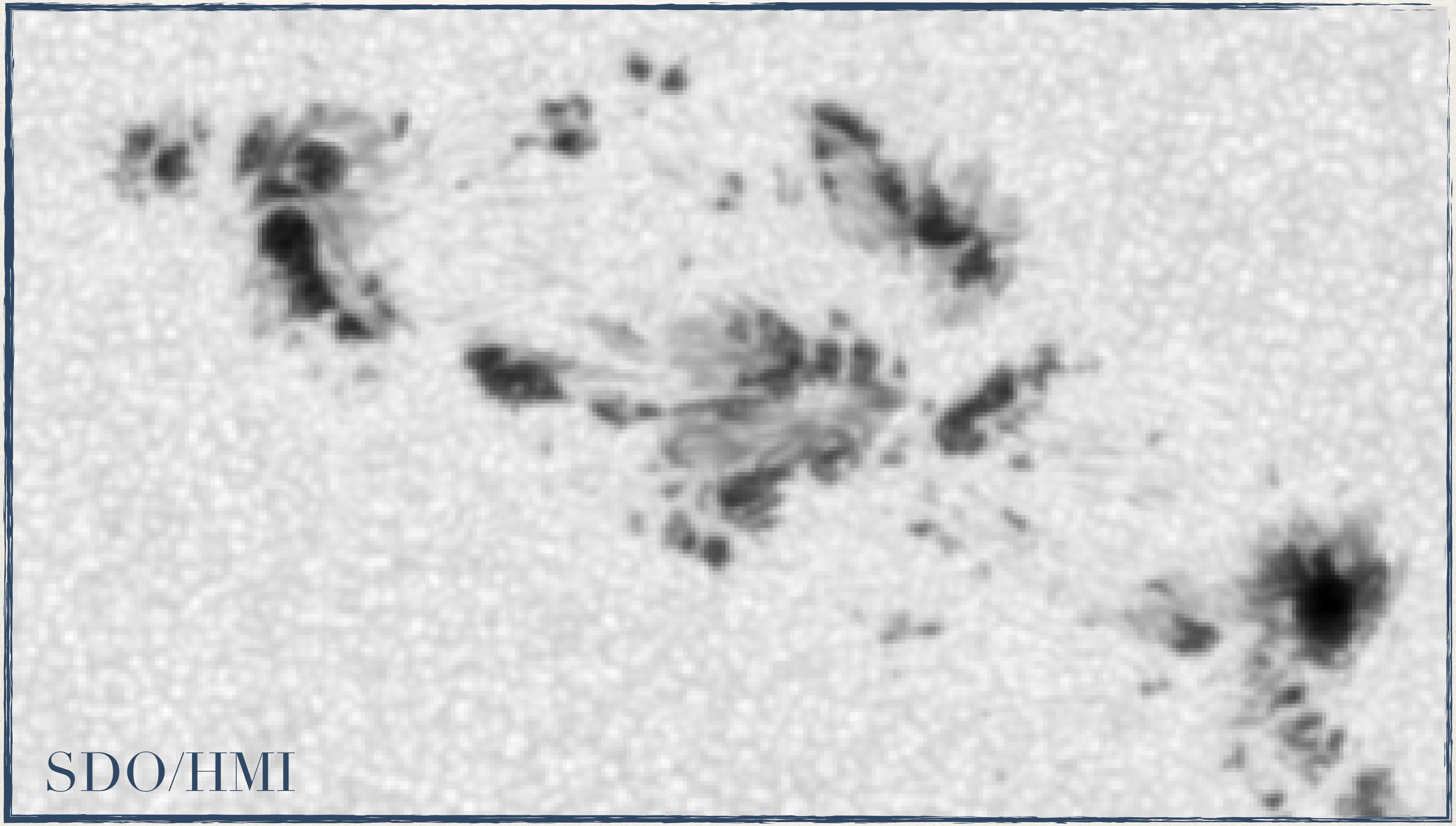
- fast (searches best fit in a pre-built database of models)
- stable (always finds best fit: no problems of local minima)
- model independent (universal search/minimization algorithm)

Cons

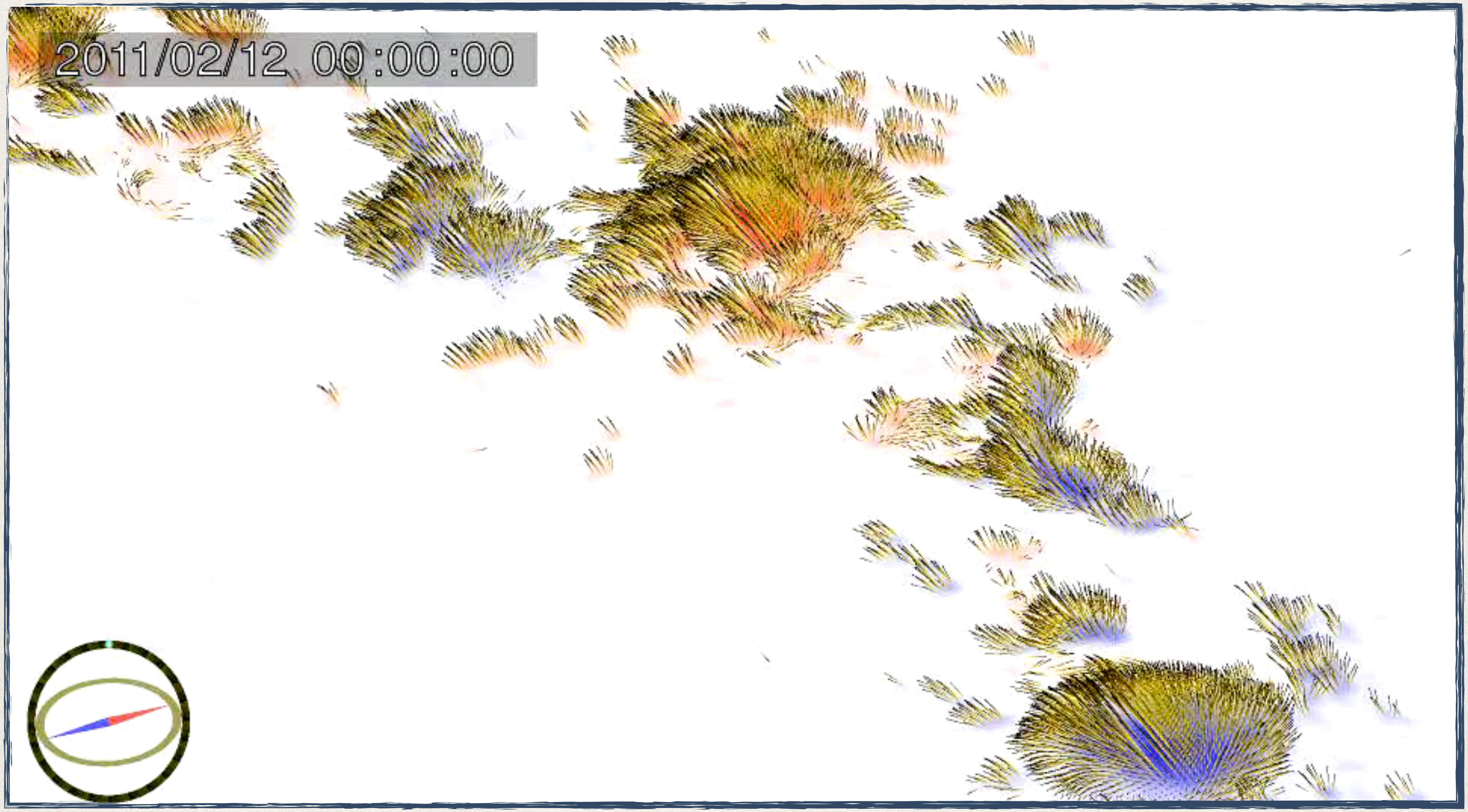
- no solution refinement (can be fixed by increasing the density of the database)
- database can become unmanageably large (dimensionality of parameter space, parameter ranges; partial mitigation from optimally sampling the parameter space)

All PCA stuff is from Roberto Casini

Spectral line inversions from HMI



Spectral line inversions from HMI



Courtesy: R. Bogart and K. Hayashi