Abundances from direct mid-resolution spectroscopy of exoplanets

Mostly taken from Konopacky et al. 2013 and Line et al. 2014
The data

5.5hr NIR spectroscopy from OSIRIS (@Keck) R~4000
First step:
Cross correlation with theoretical line lists:
\( \text{H}_2\text{O}, \text{CO}, \text{but no CH}_4 \)
Taking on abundances

Second Step:
High pass filtering to highlight high resolution features such as lines
Taking on abundances

Third step:
Model fitting to find the « best » C/O
Line et al 2014 approach

« Data driven approach rationale »

\[ X : \text{state vector (i.e. Teff, log g, abundance}_1, \ldots \text{ abundance}_n...) \]
\[ Y : \text{Observation vector (i.e. fluxes at given wavelength)} \]
\[ Y = F(X) \]

Usually \( F(X) \) is the model and we compare \( F(X) \) to \( Y \) to retrieve \( X \)

The idea is to invert the problem and directly find \( X \) analytically

« Bayesian approach »

They formulate the problem as:

Finding \( P(X|Y) \)

*Bayes theorem*: \[ P(X|Y) = P(Y|X) \cdot P(X) \]

Where \( P(X) \) is called the prior probability of having given state vector
Minimising the cost function

After some algebra it comes that the most likely $X$ is the one which minimises this:

$$\chi^2(x) = (y - F(x))^T S_e^{-1}(y - F(x)) + (x - x_a)^T S_a^{-1}(x - x_a)$$

$X$ : state vector (i.e. $T(z1)...T(zn)$, log $g$, abundance1, ... abundance n...)
$X_a$ : prior state vector
$S_a$ : Prior allowed variability or confidence (covariance matrix)
$Y$ : Observation vector (i.e. fluxes at given wavelength)
$S_e$ : gaussian observational error (diagonal if uncorrelated)
$F(X) : « forward model »$
Advantages

- *Minimisation issues are well known*

- *Errors on retrieved X are consistently « given » by the covariance matrix*

- *In theory this is a very optimised approach which provides nice quantitative results in a consistent way.*
Weakness

- **Huge dependance on the forward model**:
- Directly in the minimisation « F(X) »
- By its jacobian in the error determination

*But what is this forward model ???*  
- It needs to take the form of a matrix operator
- One dimension is the number of wavelength bin
- The other the number of input parameters (=length of the state vector)
The model atmosphere is parameterized with **five retrievable gases**: H2O, CH4, CO, CO2, and NH3. H2/He continuum absorption is also included where the H2/He mole fraction (with He/H2=0.193) is computed by subtracting the latter molecules from unity. **All mixing ratios are assumed to be uniform with altitude**.

The Freedman et al. (2008) cross section database was used with the updates to the ammonia and H2 collision-induced opacities described in Saumon et al. (2012). **Alkali metals, metal oxides or hydrides are not included** in this investigation.

The temperature profile is not parameterized, rather the temperature at each model slab is retrieved. However, some smoothing is implemented through the a-priori covariance matrix to prevent overfitting and unphysical oscillations in the profiles.

**Is it really a crude model?**

**Is it because any new parameter increases the dimension of the problem?**

**Could we forward-model using BT-SETTL?**