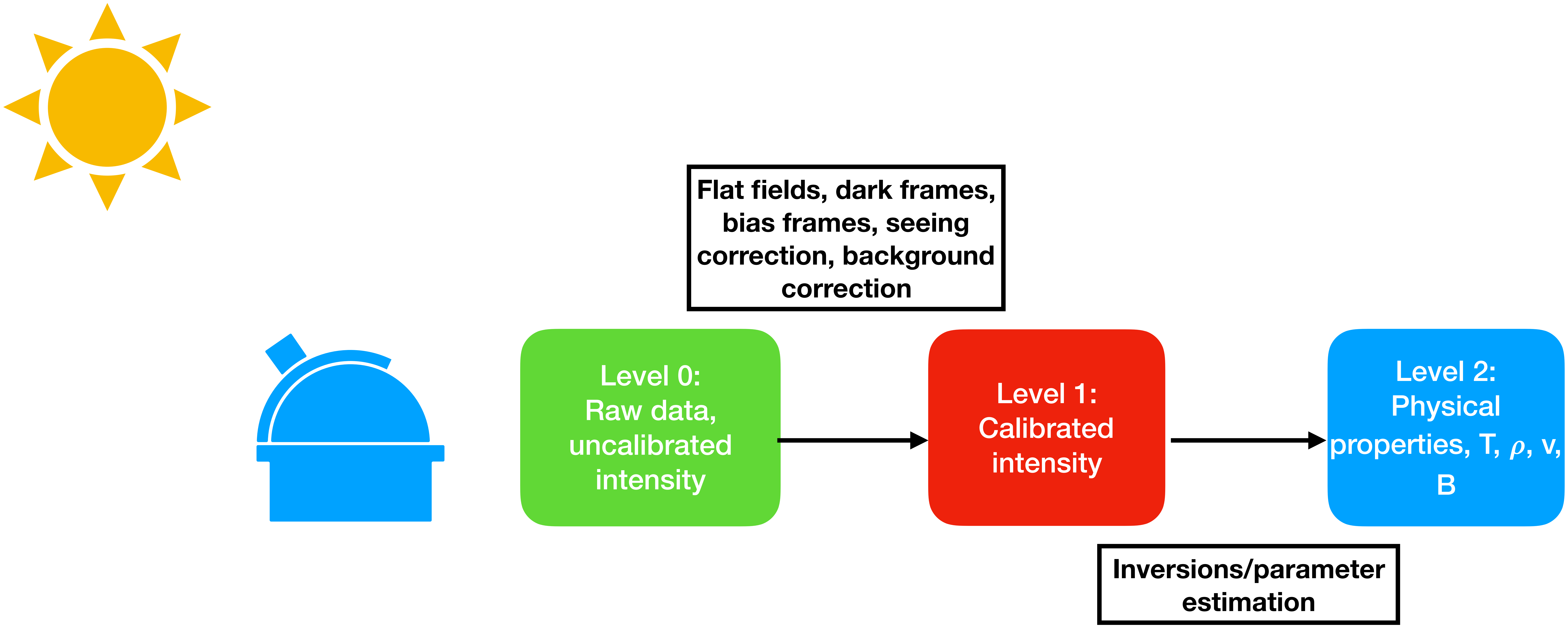
The Future of Inversions Using Invertible Neural Networks

John A. Armstrong
with: C. Osborne and L. Fletcher

http://bit.ly/john_mssl_2019

Data Pipelines



Learning Physics with Inversions

Optically thin inversions — also known as "direct inversions"

- arrived at by assuming thermal equilibrium and the coronal approximation

1. Estimation of temperature structure and electron density using differential emission measures (DEMs)

see Hannah & Kontar 2012, and Cheung et al. 2015 for solutions to the problem

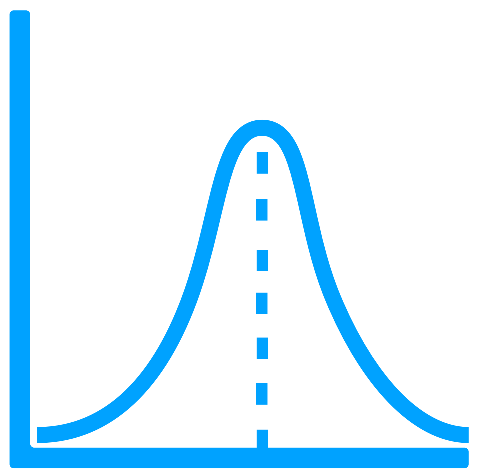
2. Temperature and non-thermal velocity calculations from line shapes

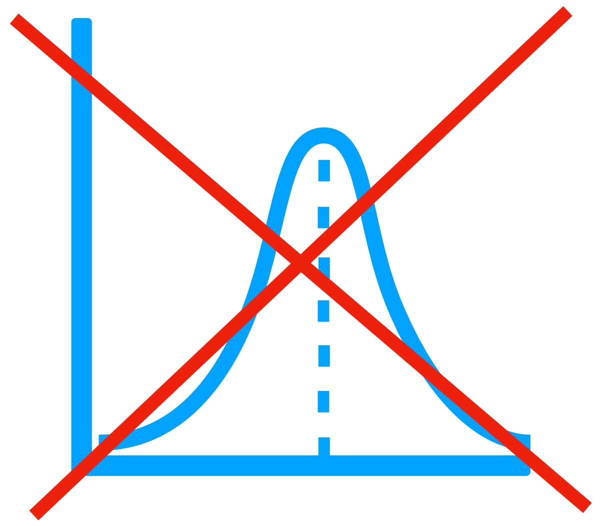
see Jeffrey et al. 2016, 2017

3. Temperature and electron density from line ratios

Two lines from the same species can be used to make estimates of T and n_e e.g. Fe I 6301 & 6302Å doublet

$$I_{ji} = K_{ji}\, \textbf{DEM}(T)$$



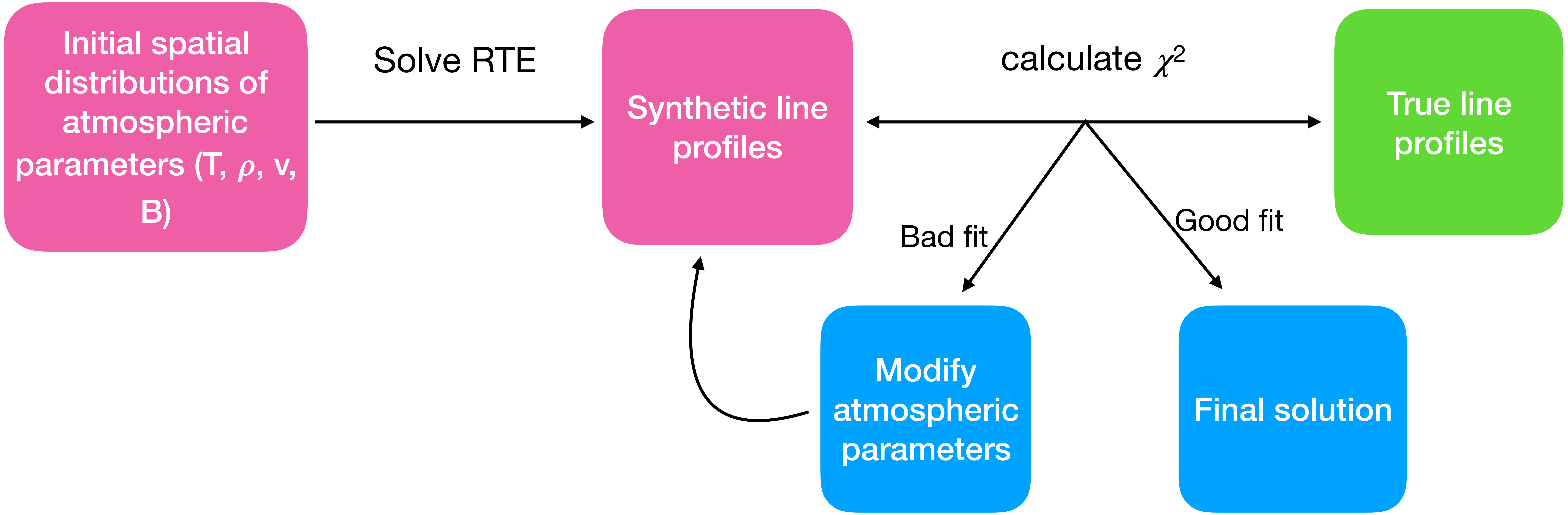


$$\frac{I_{ki}}{I_{ji}}\left(n_e,\ T\right)$$

Learning Physics with Inversions

Optically thick inversions (see Ruiz Cobo & Del Toro Iniesta 2012, Socas-Navarro et al. 2000, Skumanich & Lites 1987, Asensio Ramos et al. 2008, de la Cruz Rodriguez et al. 2019)

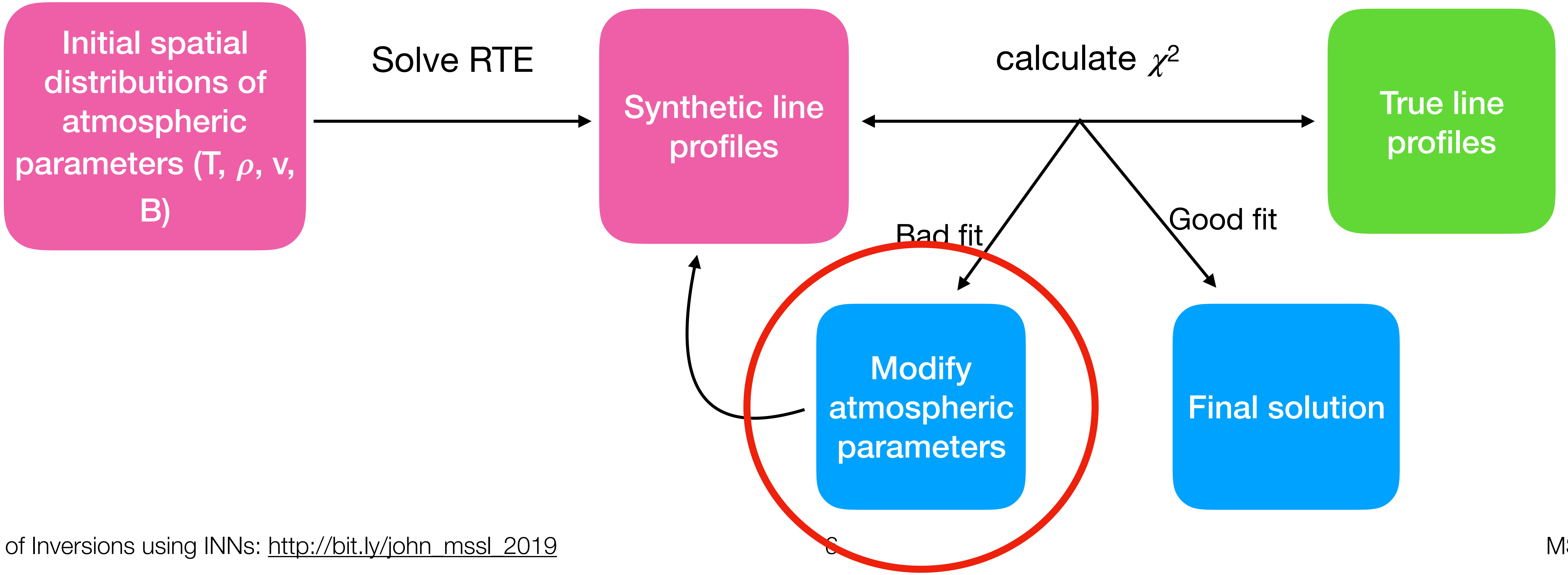
- Solve the full radiative transfer equation (RTE), cannot assume that photons do not interact as they travel through the plasma
- Solving RTE for certain atmospheric conditions gives line profiles
- This leads to the method for doing optically thick inversions: forward modelling



Learning Physics with Inversions

Optically thick inversions (see Ruiz Cobo & Del Toro Iniesta 2012, Socas-Navarro et al. 2000, Skumanich & Lites 1987, Asensio Ramos et al. 2008, de la Cruz Rodriguez et al. 2019)

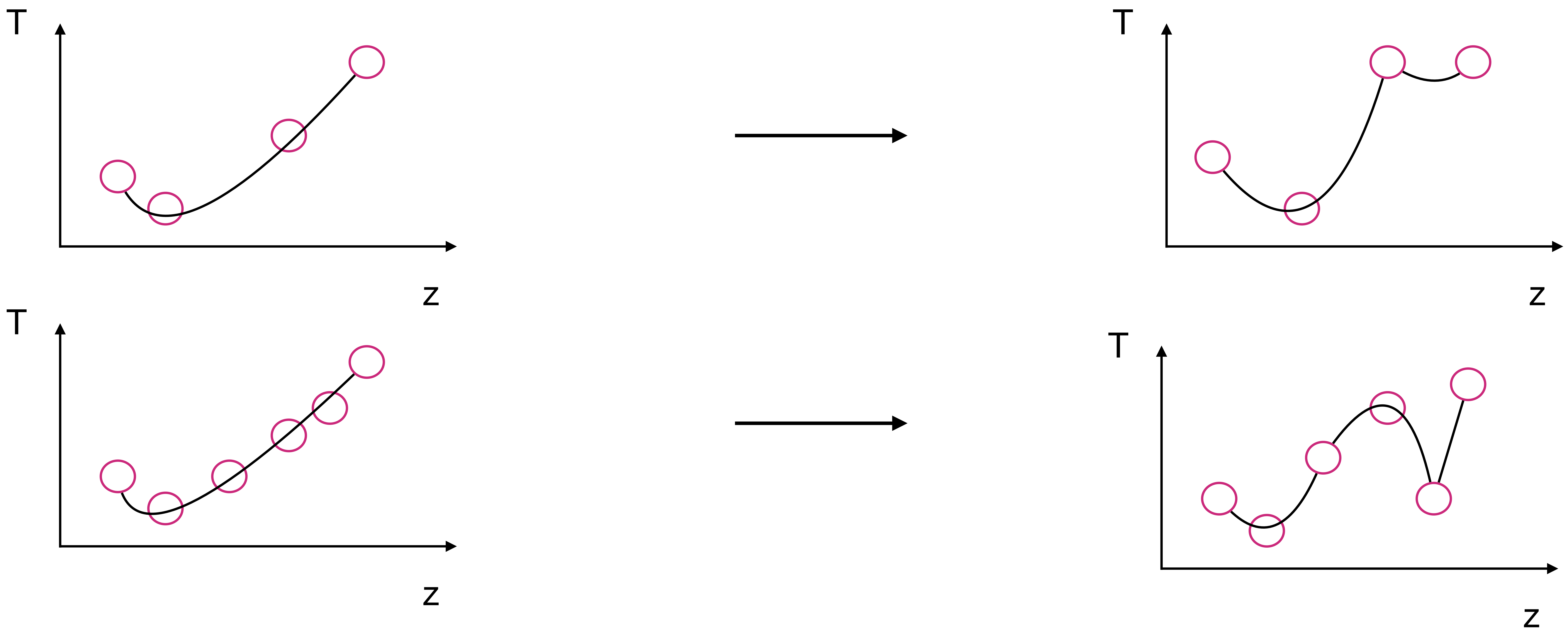
- Solve the full radiative transfer equation (RTE), cannot assume that photons do not interact as they travel through the plasma
- Solving RTE for certain atmospheric conditions gives line profiles
- This leads to the method for doing optically thick inversions: forward modelling



Learning Physics with Inversions

Optically thick inversions

- Modifications of atmospheric parameters done via nodes
- Number of nodes kept low to avoid overfitting and unphysical solutions and reduce dimensionality of the problem
- Nodes define grid for atmospheric parameters and are interpolated between

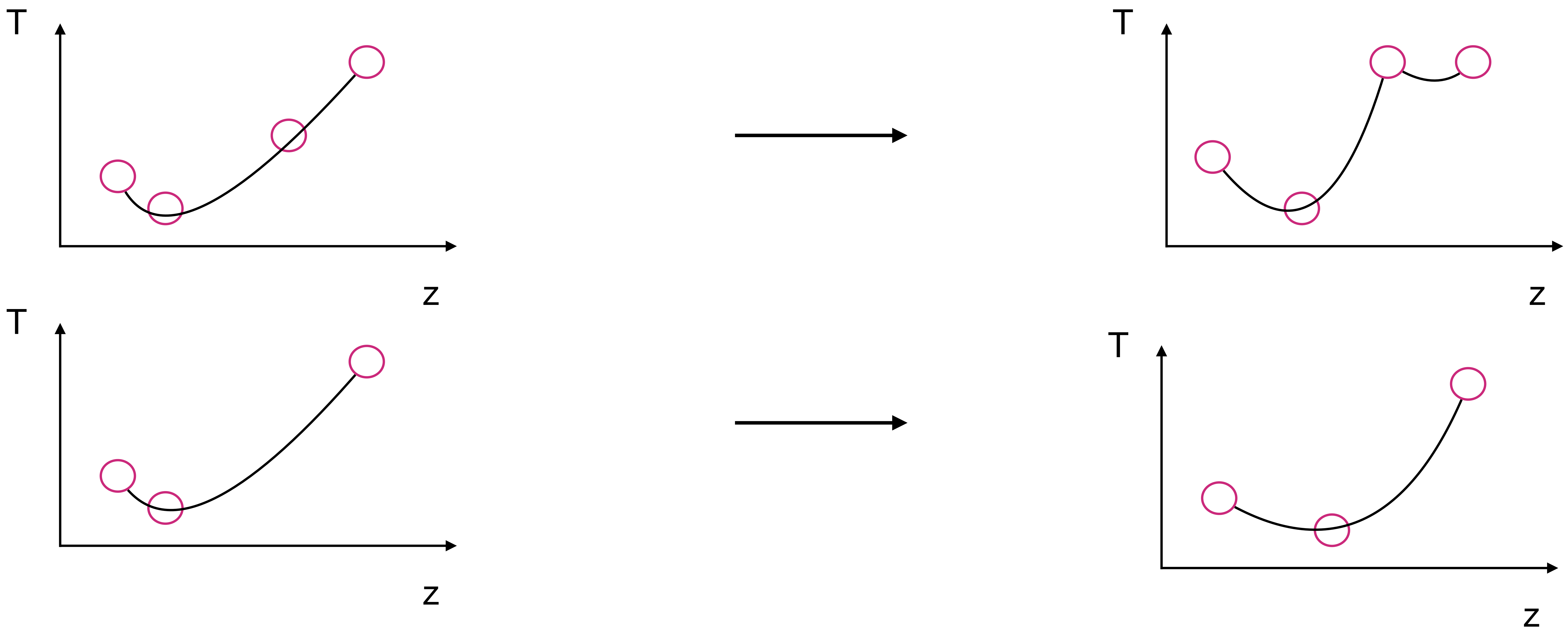


TOO MANY NODES!!

Learning Physics with Inversions

Optically thick inversions

- Modifications of atmospheric parameters done via nodes
- Number of nodes kept low to avoid overfitting and unphysical solutions and reduce dimensionality of the problem
- Nodes define grid for atmospheric parameters and are interpolated between



TOO LITTLE NODES!!

Basics of Deep Learning

- **Machine learning the process of using statistical techniques to give computers the ability to learn how to perform a specific task *without* being explicitly programmed**
- **Deep neural networks are very good function approximators (Cybenko 1989, Lu et al. 2017)**
  - processes that can be expressed by well-defined functions can be learned by a deep neural network

Trained neural network

$$y = f(x; \underbrace{\{\theta_1, \ldots, \theta_n\}}_{\text{Learnable Parameters}})$$

Output

Input

Basics of Deep Learning

- **Machine learning the process of using statistical techniques to give computers the ability to learn how to perform a specific task *without* being explicitly programmed**
- **Deep neural networks are very good function approximators (Cybenko 1989, Lu et al. 2017)**
  - processes that can be expressed by well-defined functions can be learned by a deep neural network

Trained neural network

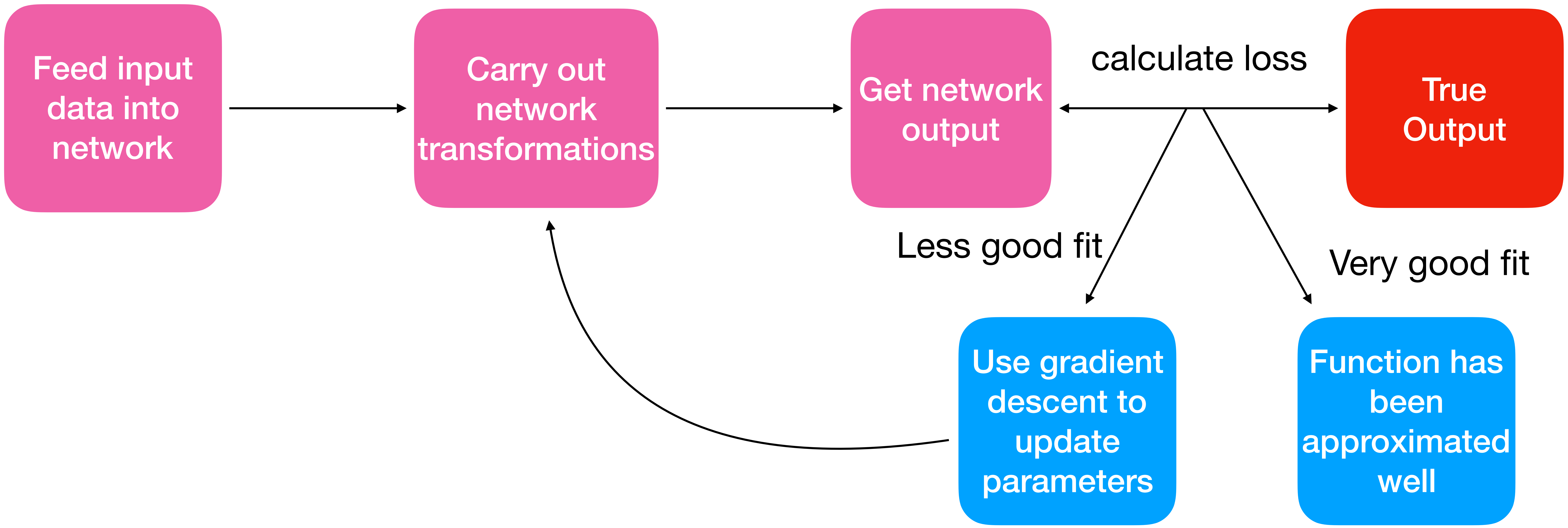
$$y \approx f(x; \underbrace{\{\theta_1, \ldots, \theta_n\}}_{\text{Learnable Parameters}})$$

Output

Input

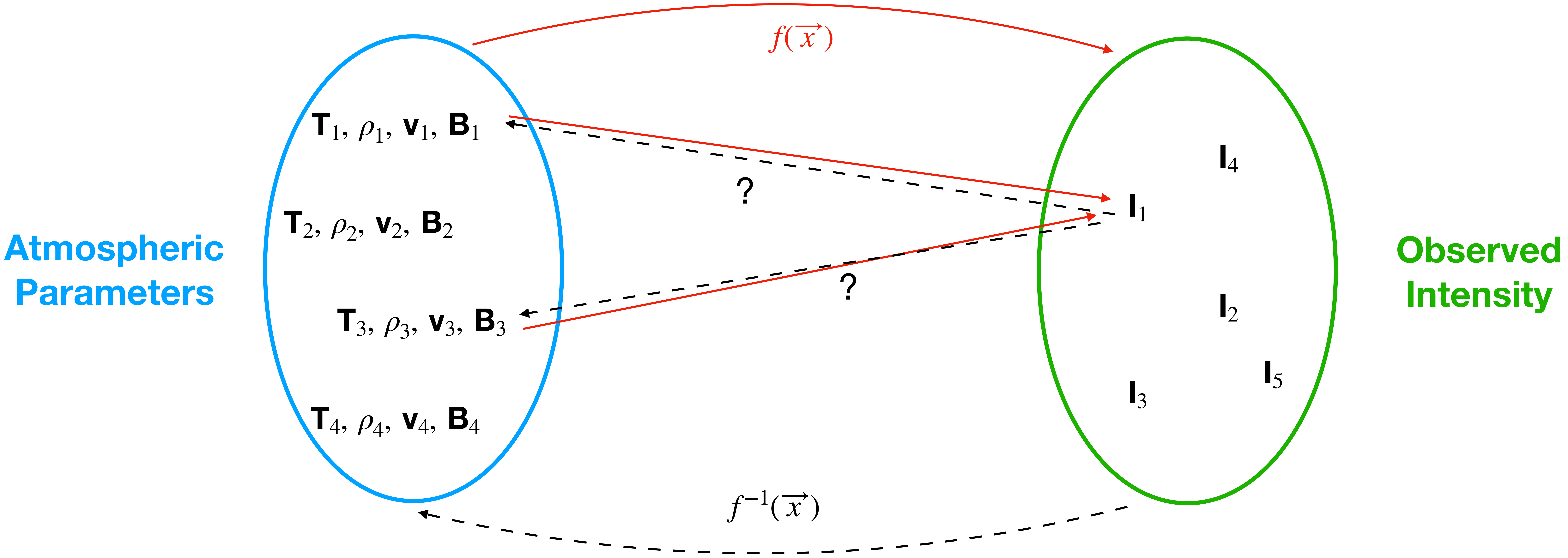
Basics of Deep Learning

- **The optimisation (training) takes place like a feedback loop similar to how the inversion process works**



Back to Inversions

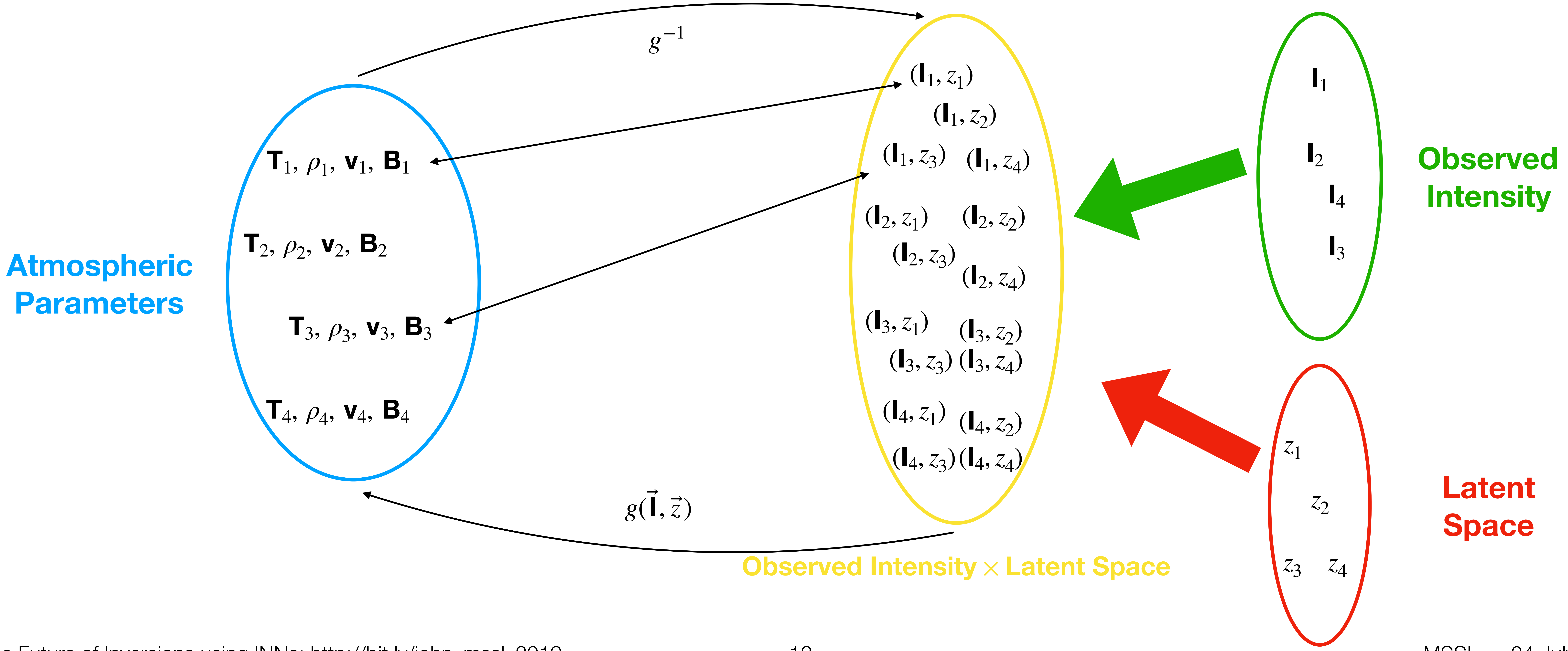
- **Inversions, however, are not well-defined functions**
  - there are many combinations of atmospheric parameters that can produce the same line profiles
  - there is information lost about the physics in the forward process



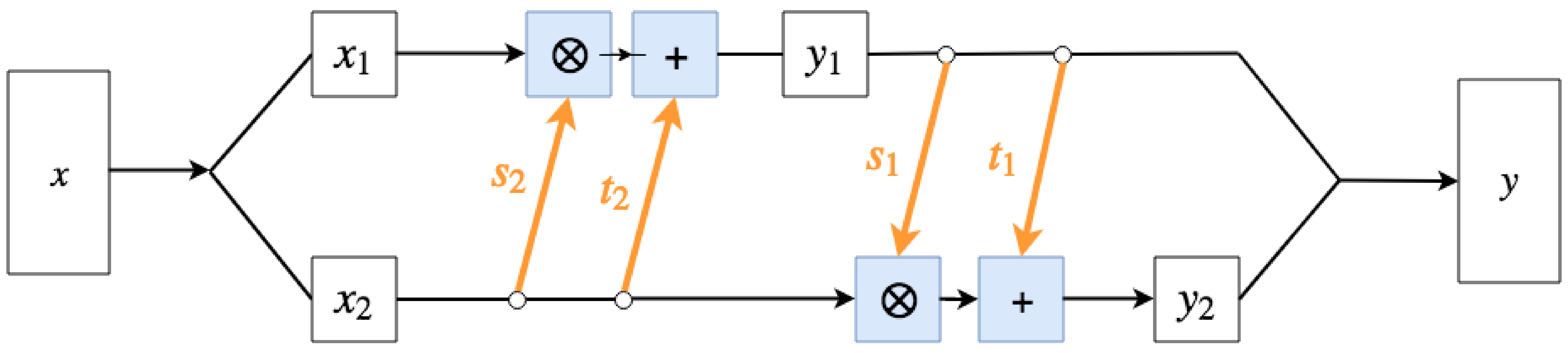
Cannot be modelled with traditional deep learning

Back to Inversions

- **How do we formulate the inverse process in such a way as to make it well-defined?**
  - Introduce a latent space, z, which contains the information lost in the forward process

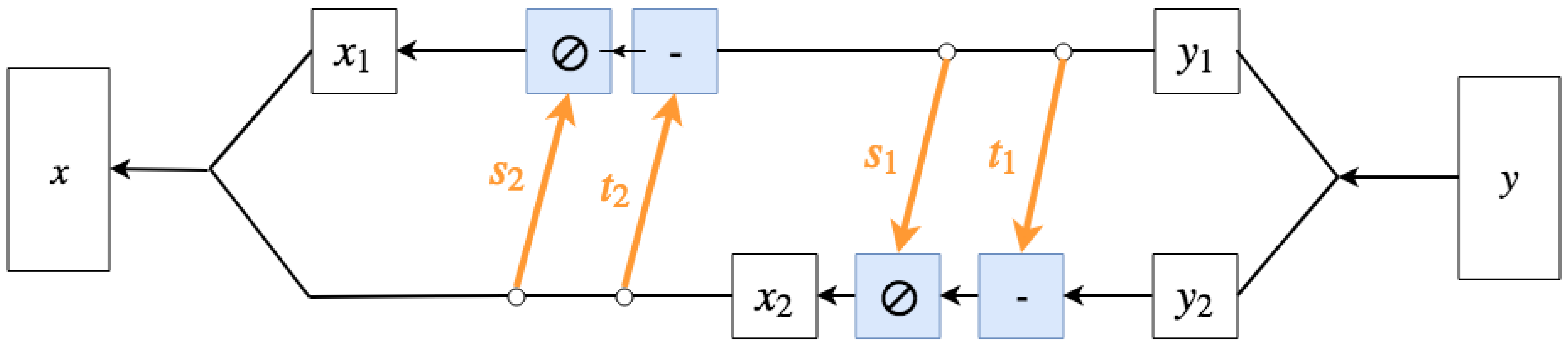


Affine-Coupling Layers



$$y_1 = x_1 \otimes \exp\left(s_2\left(x_2\right)\right) + t_2\left(x_2\right)$$

$$y_2 = x_2 \otimes \exp\left(s_1\left(y_1\right)\right) + t_1\left(y_1\right)$$

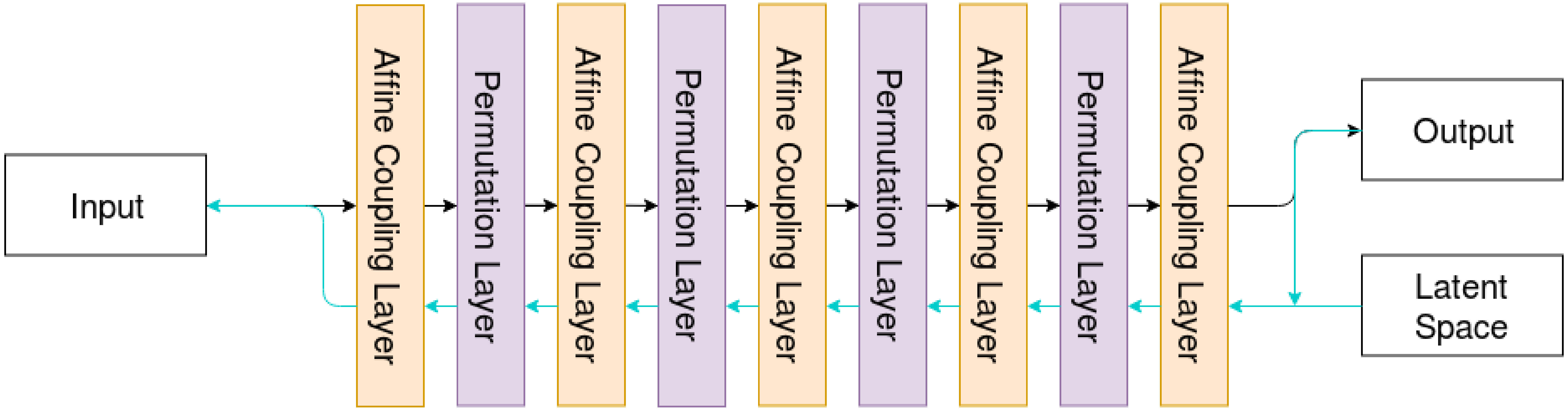


$$x_2 = \left(y_2 - t_1\left(y_1\right)\right) \oslash \exp\left(s_1\left(y_1\right)\right)$$

$$x_1 = \left(y_1 - t_2\left(x_2\right)\right) \oslash \exp\left(s_2\left(x_2\right)\right)$$

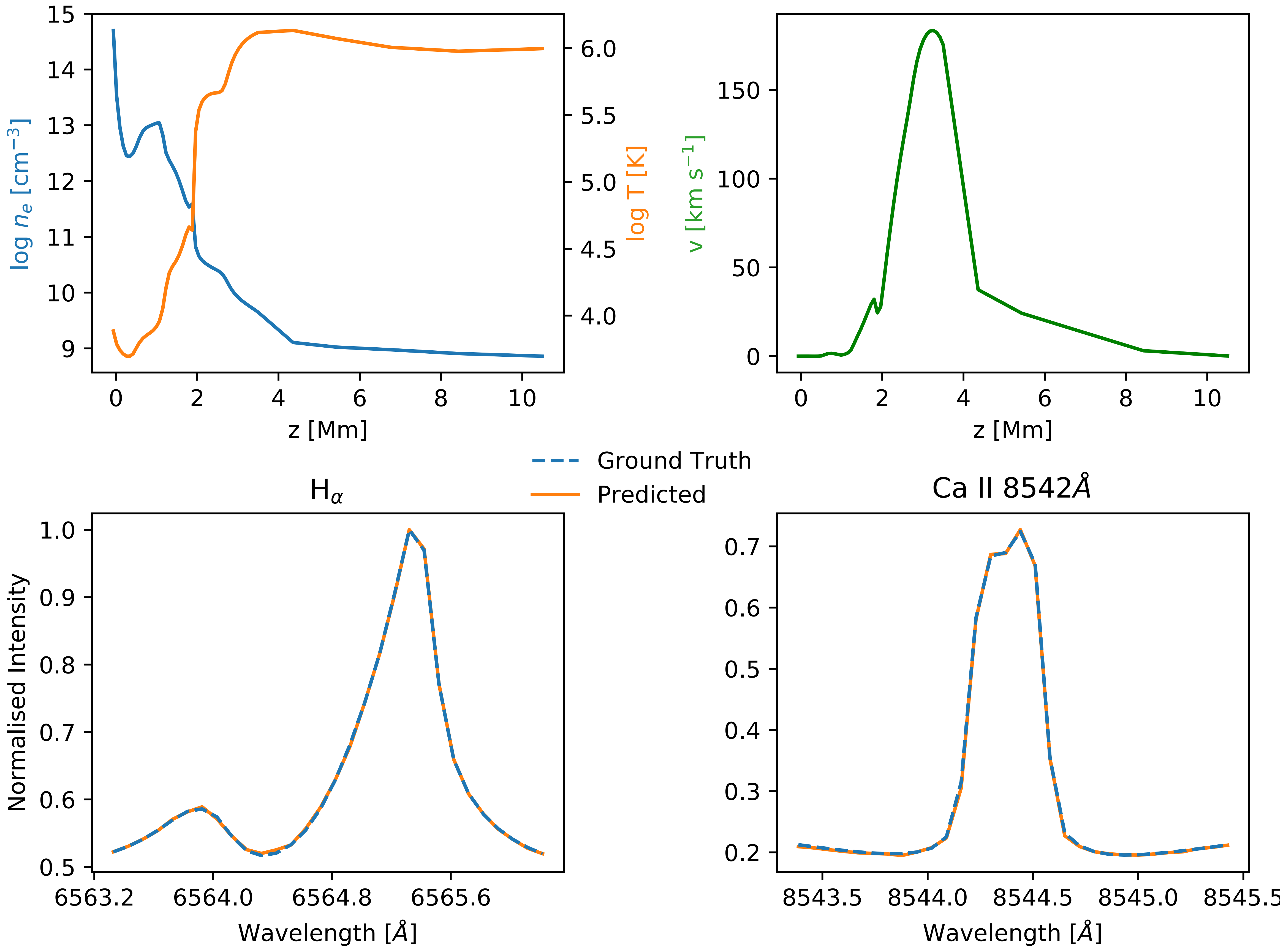
See Dinh et al. 2014, 2017 & Ardizzone 2018 for more details

Invertible Neural Networks (INNs)

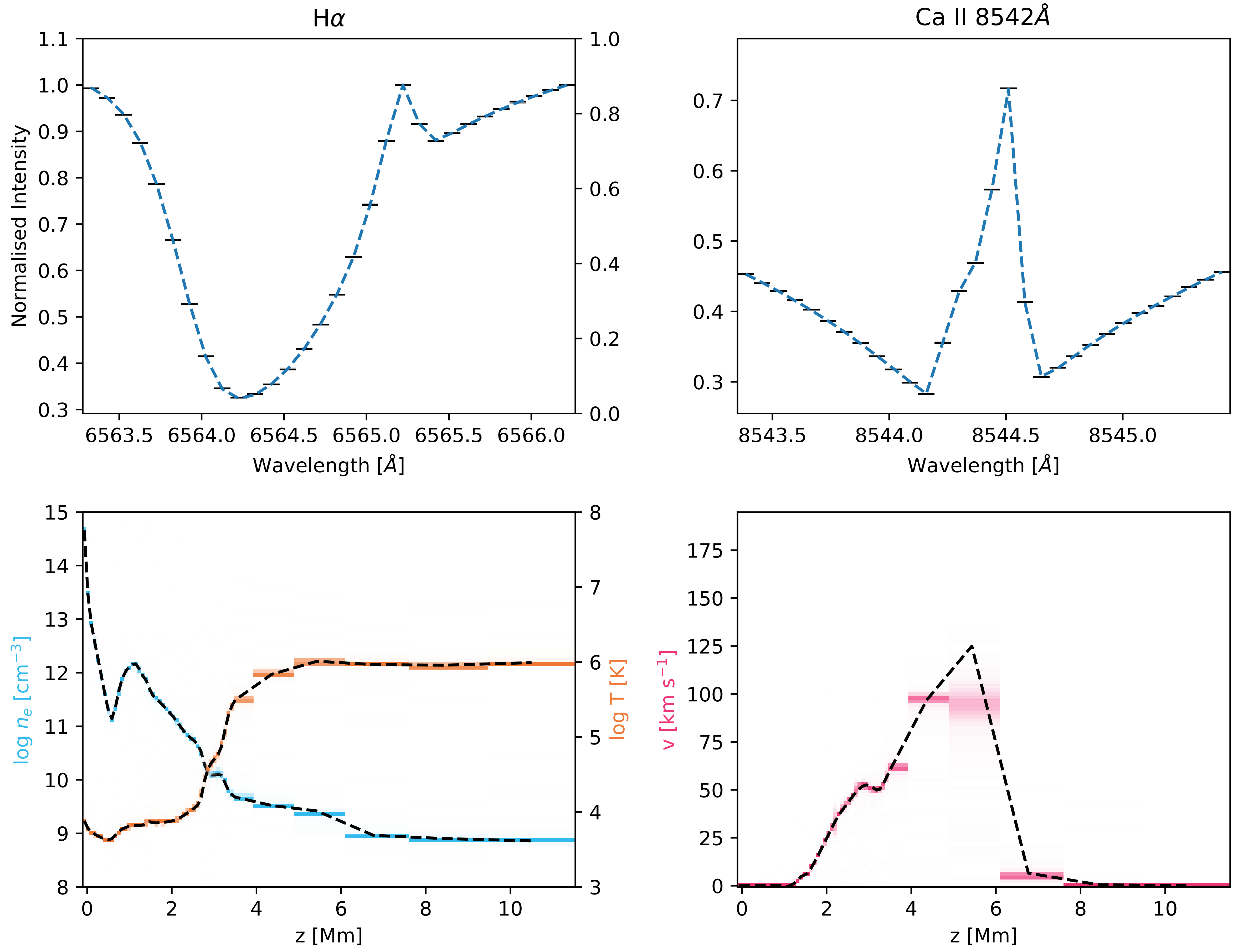


Training data: https://star.pst.qub.ac.uk/wiki/doku.php/public/solarmodels/start

Invertible Neural Networks (INNs)



Forward Model Test



Inversion Test

Inversion of Real Data

- **M1.1 flare SOL20140906T17:09 NOAA AR12157**
- **Observed by SST/CRISP in Hα and Ca II λ8542**
- **Wavelength sampling: 15 points for Hα, 25 for Ca II**

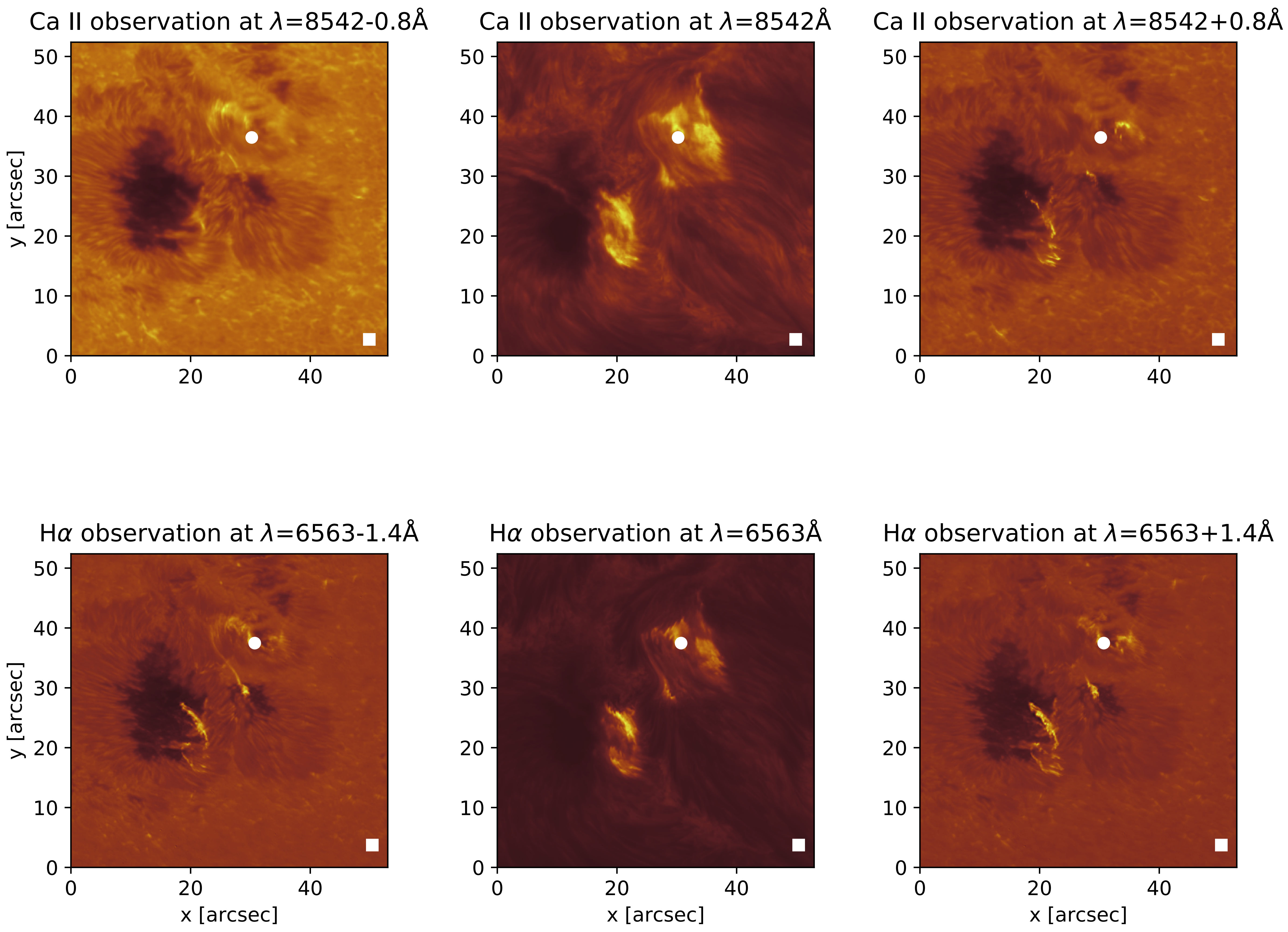
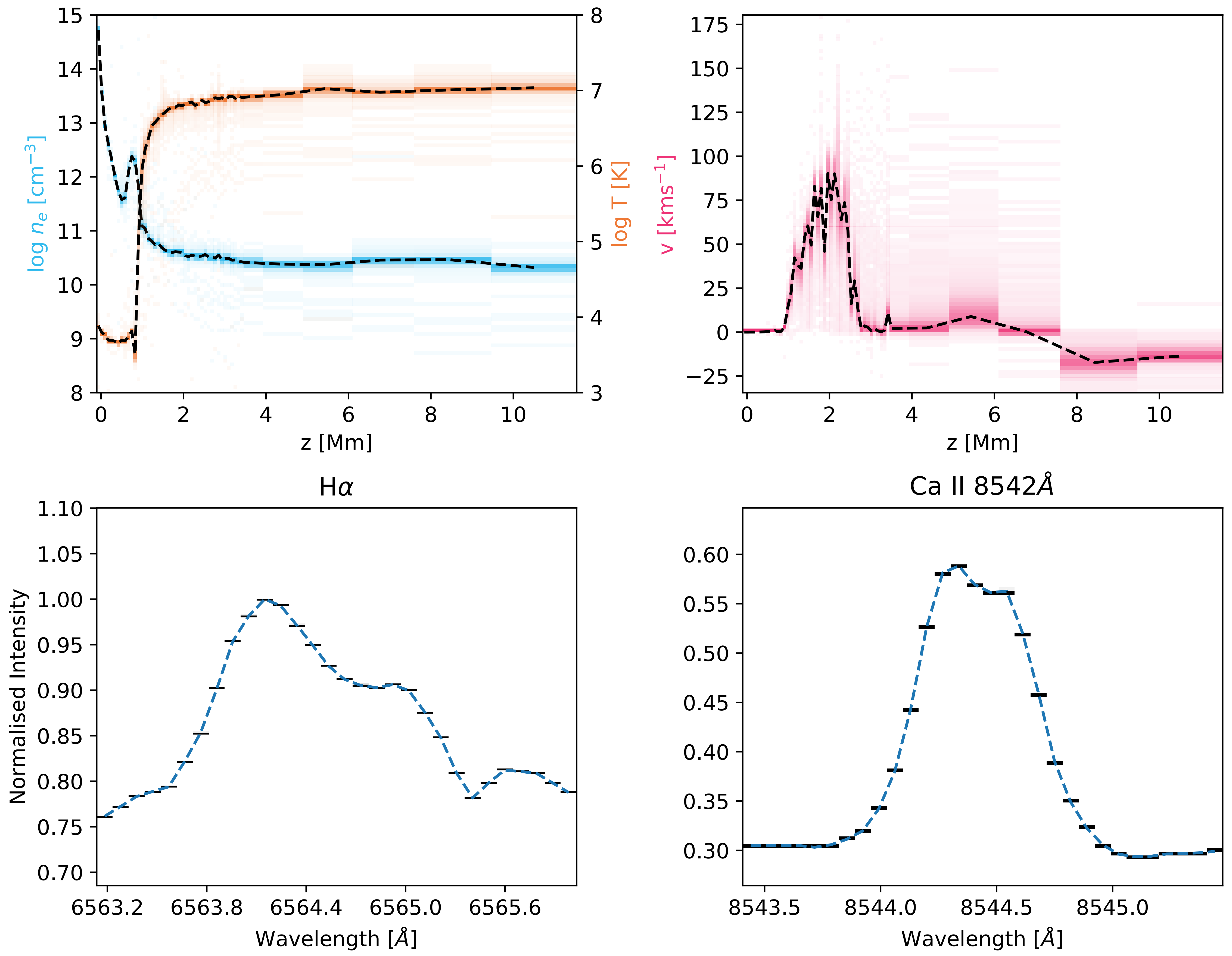
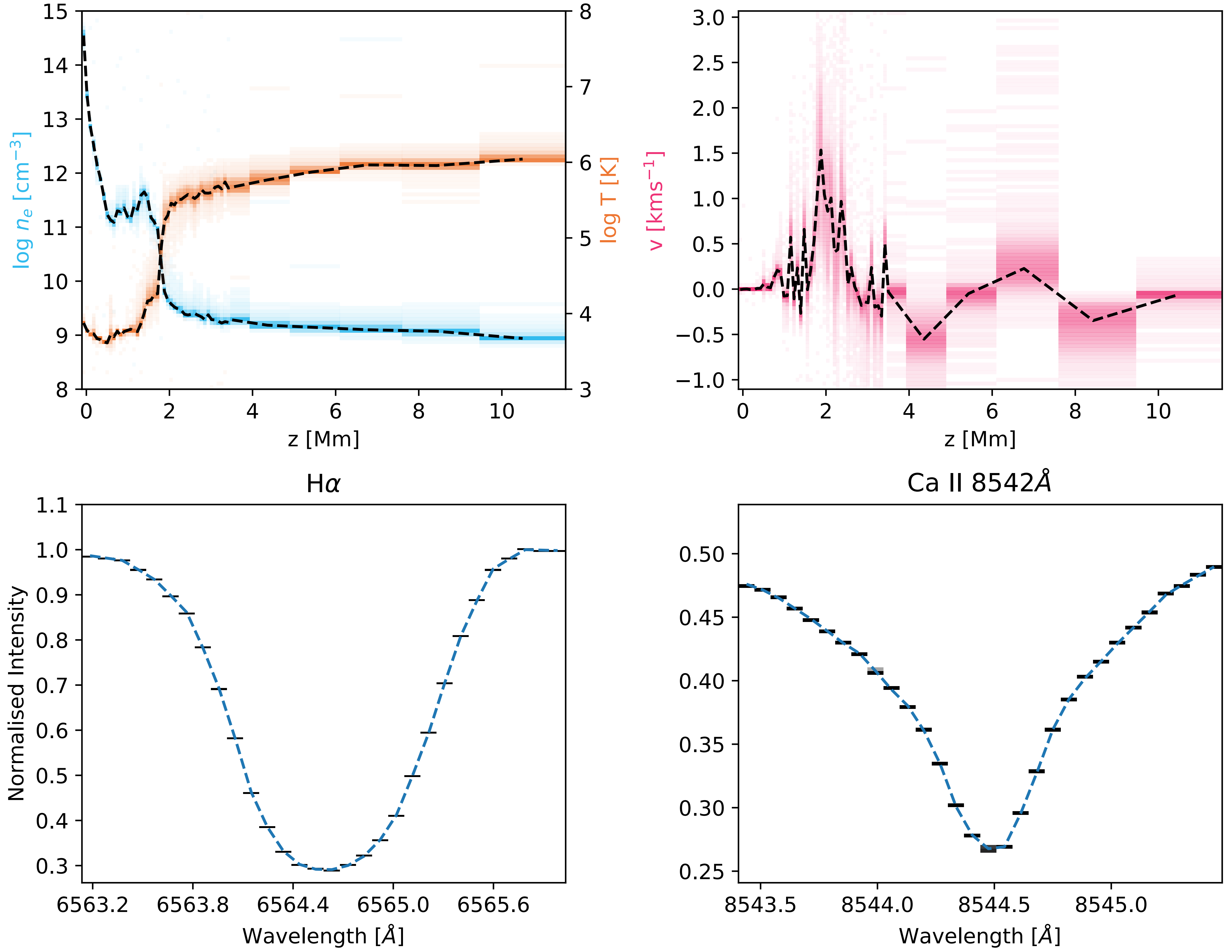


Figure: Observations in both wavelengths just after the flare onset. The circular point is a point on the flare ribbon.
The square point is a point off the flare ribbon.

Single Pixel Inversions



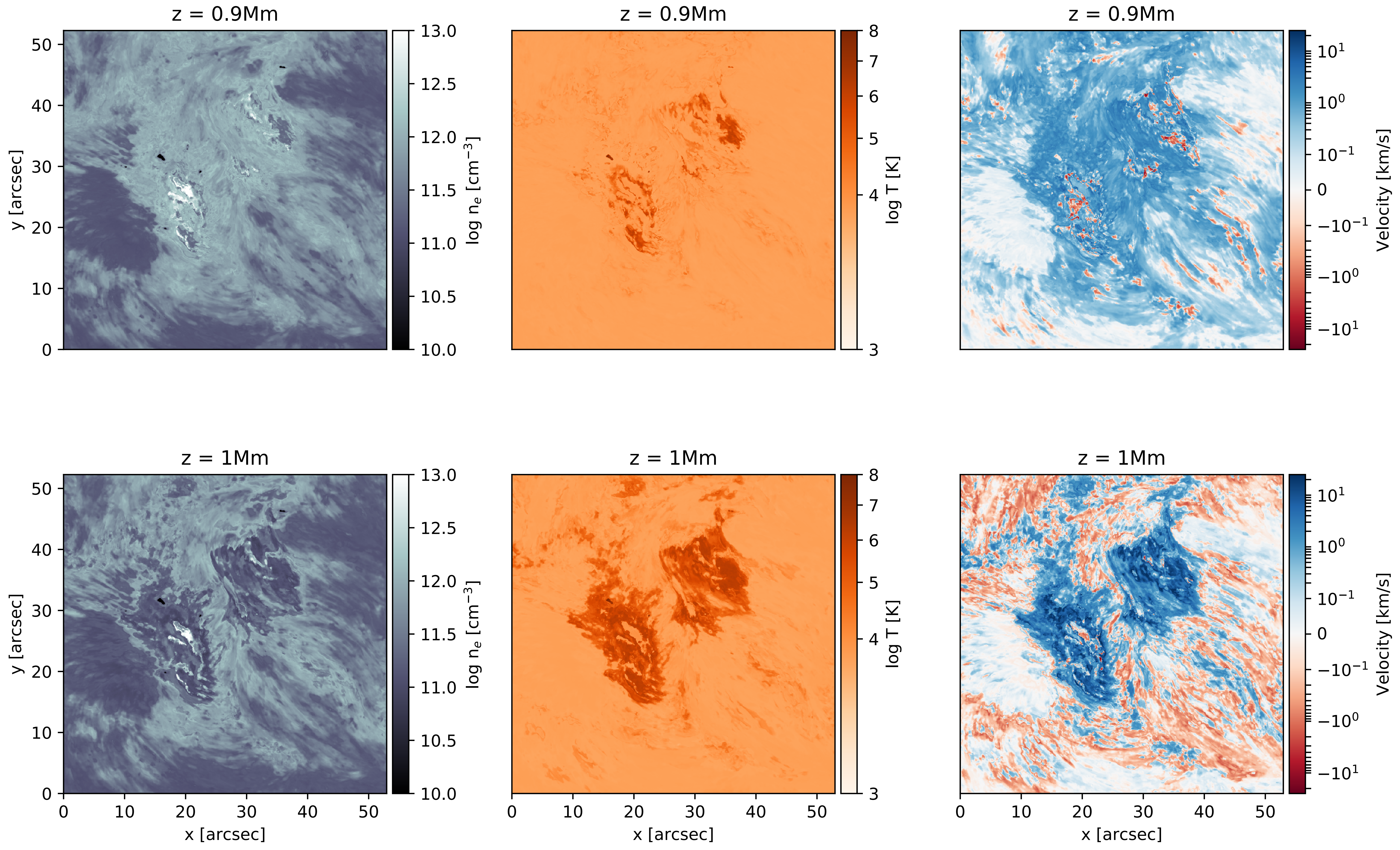
Inversion of circular point



Inversion of square point

See Osborne, Armstrong & Fletcher, 2019

Whole Image Inversions



Future Work

- **Analyse whole image inversions and see what our inversions say about the flaring chromosphere**
- **Calculate errors on the inversions**
  - The whole image inversions come with standard error calculated on the median solutions and so have error maps — use this to quantify uncertainty in our inversions
- **Apply RADYNVERSION to 6 September 2017 X9.3 flare data (and see how much it breaks)**
- **RADYNVERSION using a log τ grid is in prep for more insights into analysis**
- **Add more spectral lines (Mg II, Ca II H&K,…)**
- **Ideally add polarimetry once a forward model is available (see Osborne thesis)**
- **Apply to quiet polarimetric data (colab with Andrés and Carlos)**

Some Useful Links

- Code: http://bit.ly/radynversion_code
- Paper: http://bit.ly/radynversion_paper
- My GitHub: http://bit.ly/gh_john