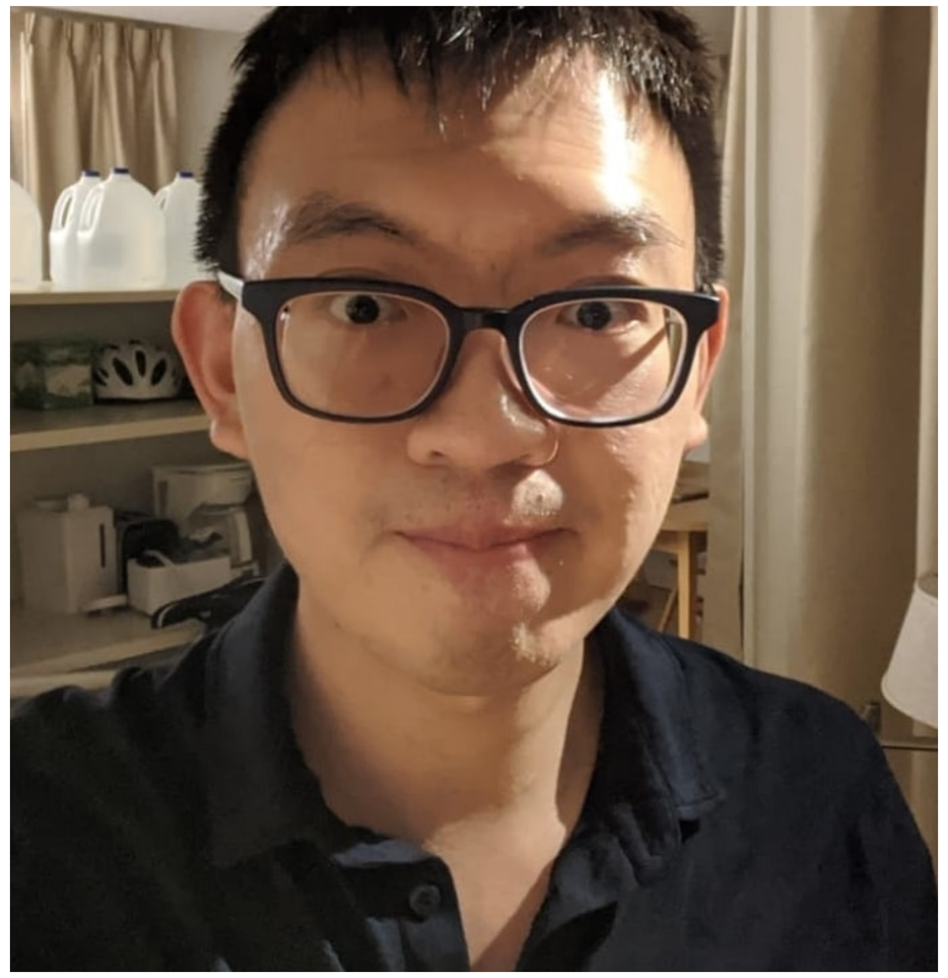
UCL

Australian National University

Learning Stellar Spectra with Deep Normalising Flows

Can generative machine learning find hidden atomic lines?

Ioana Ciucă (ANU & UCL) /w Yuan-Sen Ting (ANU)



Outline

- Brief Introduction to Deep Generative Modeling
- Why/what/how Normalising Flows?
- The unusual case of Galactic Archaeology
- Application I: find *chemical outliers*
- Application II: identify *hidden* atomic line *transitions*
- Summary

Deep Generative Modeling

Which one ain't real?

A

B

C

D

Karras, Laine & Avila 2019

Which one ain't real?

A

B

D

Karras, Laine & Avila 2019

Supervised Learning

Input: Data **X** and label **y**

Goal: Learn how to map **X** to **y**

Examples: Classification, regression

Unsupervised Learning

Input: Just data **X**, no labels

Goal: Learn *underlying structure* of the data

Examples: Dimensionality reduction, clustering

Credit: Ava Soleimany (MIT)

Supervised Learning

Input: Data **X** and label **y**

Goal: Learn how to map **X** to **y**

Examples: Classification, regression

Unsupervised Learning

Input: Just data **X**, no labels

Goal: Learn *underlying structure* of the data

Examples: Dimensionality reduction, clustering

Credit: Ava Soleimany (MIT)

Generative Models

- Solve an unsupervised learning task

- Given a set of input training sample, we want to learn a model that best represents the distribution from which the samples were generated

- Formally, learn the probability distribution over random variable **X** from a set of observed data $\{x_i\}$ with probability density $p_X(x)$ parametrised by θ

The goals of generative modeling

Density Estimation

- *Evaluate* likelihood of new points under the model
- Powerful application for outlier detection
- Build informative priors

Sample Generation

- It's fun to *sample* new faces
- Uncover bias in data & create fairer models by debiasing
- For science, powerful method to identify correlation structure in very high-dimensional datasets

The goals of generative modeling

Density Estimation

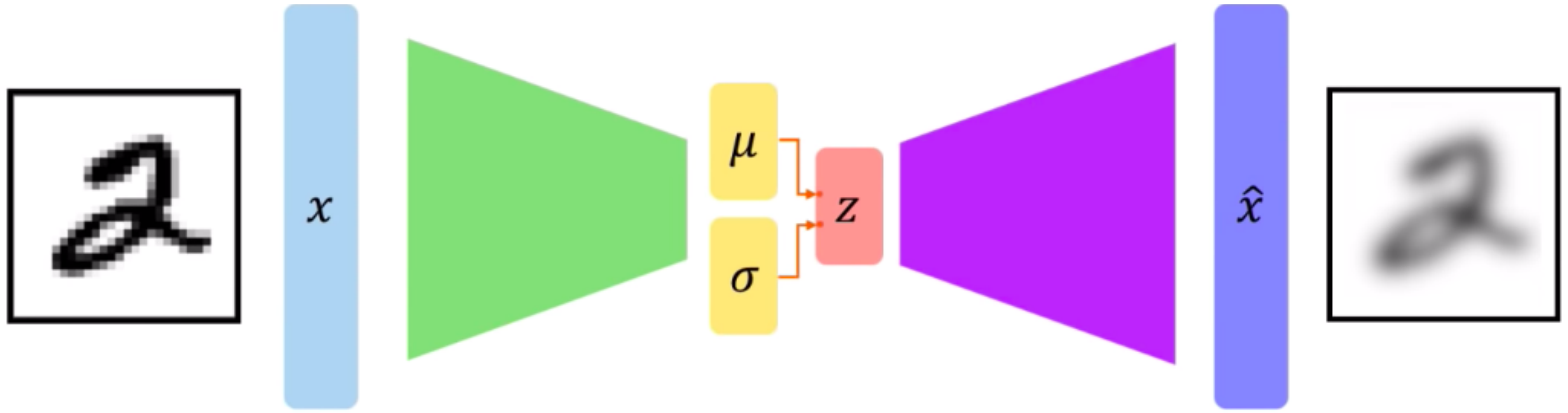
- *Evaluate* likelihood of new points under the model
- Powerful application for outlier detection
- Build informative priors

Sample Generation

- It’s fun to *sample* new faces
- Uncover bias in data & create fairer models by debiasing
- For science, powerful method to identify correlation structure in very high-dimensional datasets

How to learn $P_{model}(x)$ to be as similar to $P_{data}(x)$?

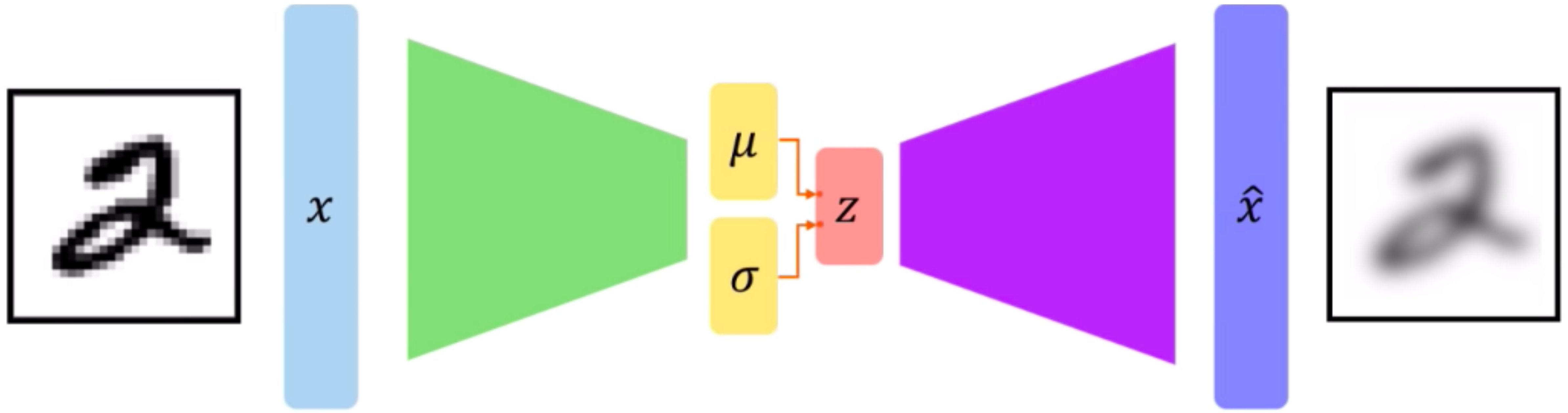
- Variational AutoEncoders (VAEs)



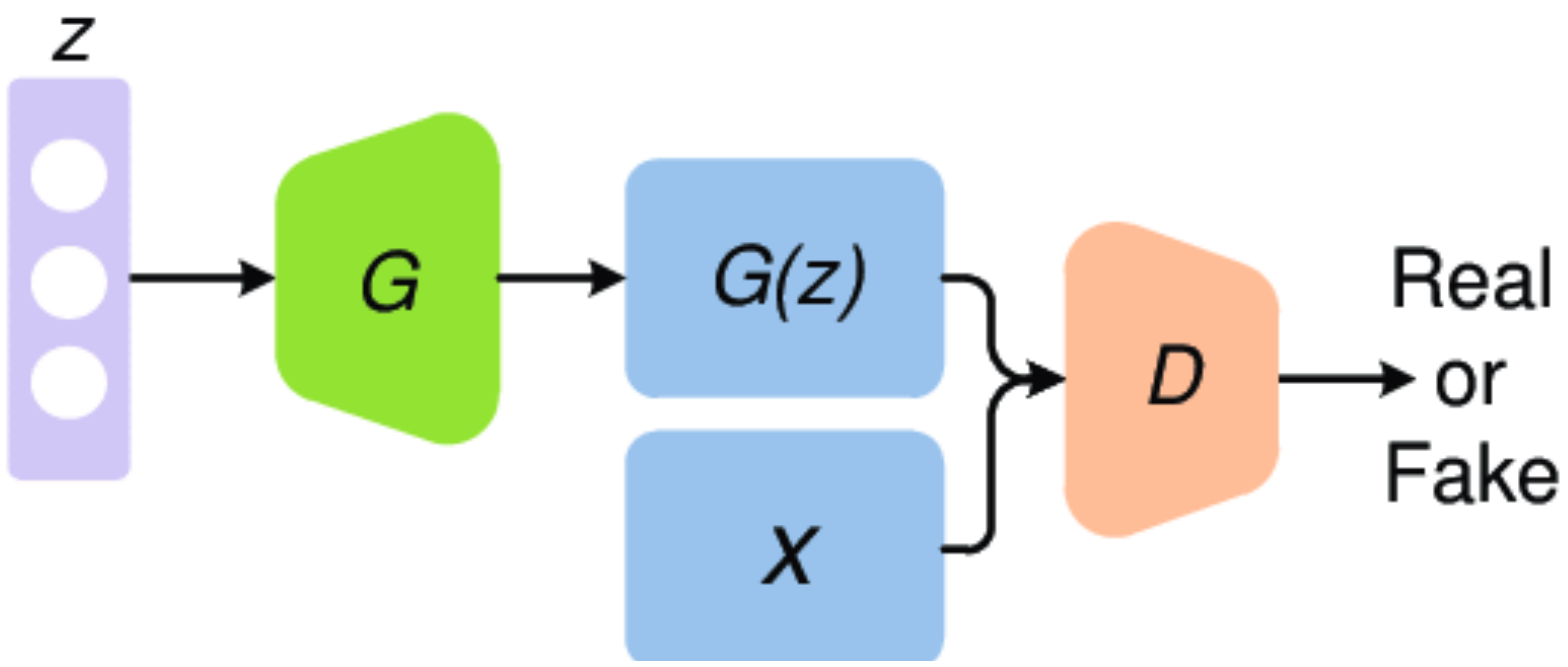
- Generative Adversarial Networks (GANs)

How to learn $P_{model}(x)$ to be as similar to $P_{data}(x)$?

- Variational AutoEncoders (VAEs)



- Generative Adversarial Networks (GANs)

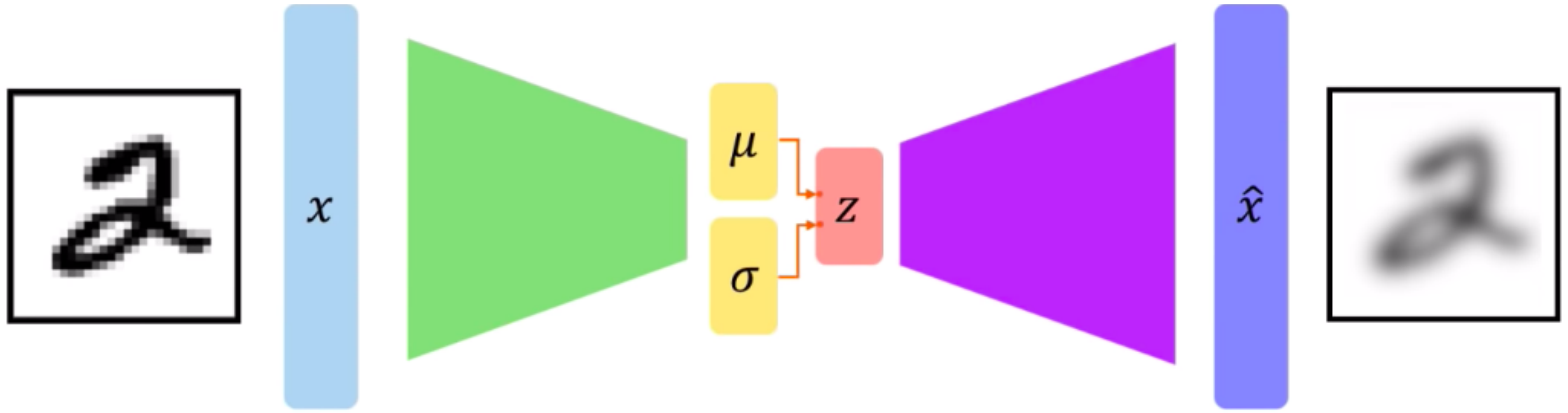


- Normalising Flows

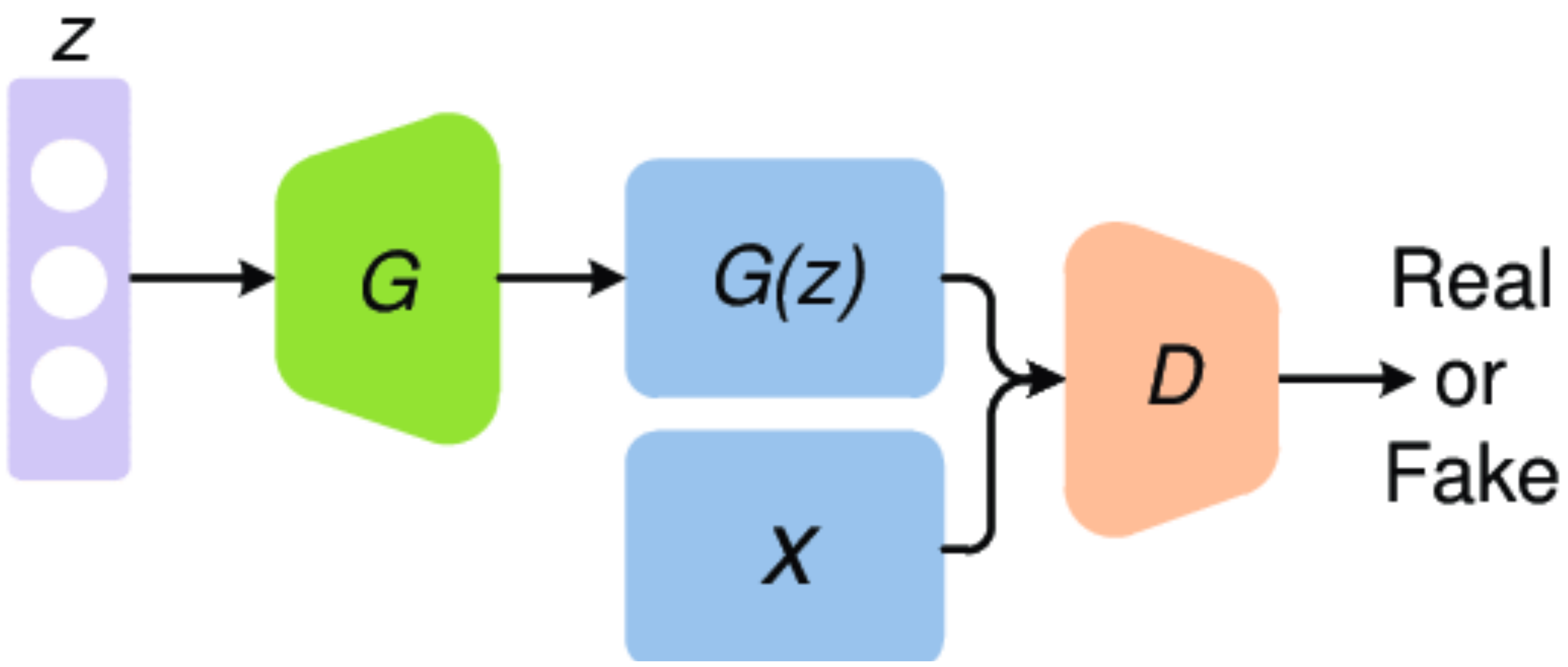

~10,000 dimensions

How to learn $P_{model}(x)$ to be as similar to $P_{data}(x)$?

- Variational AutoEncoders (VAEs)

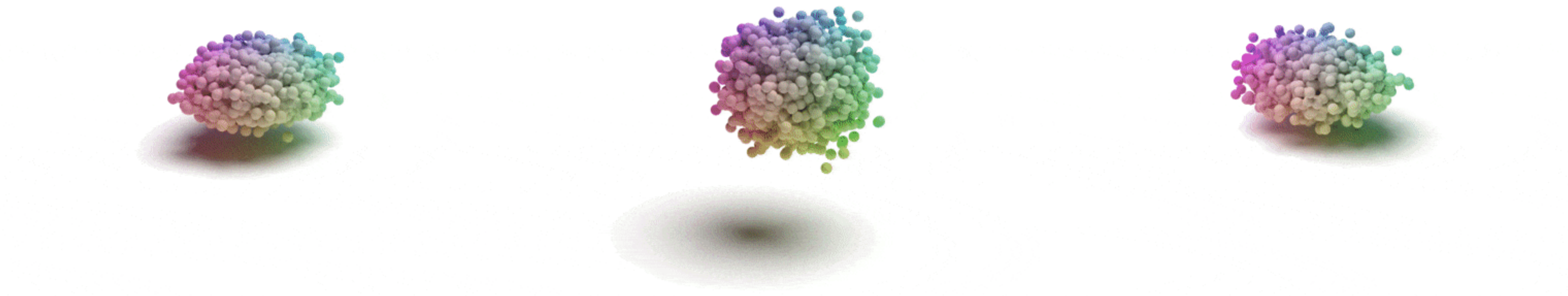


- Generative Adversarial Networks (GANs)

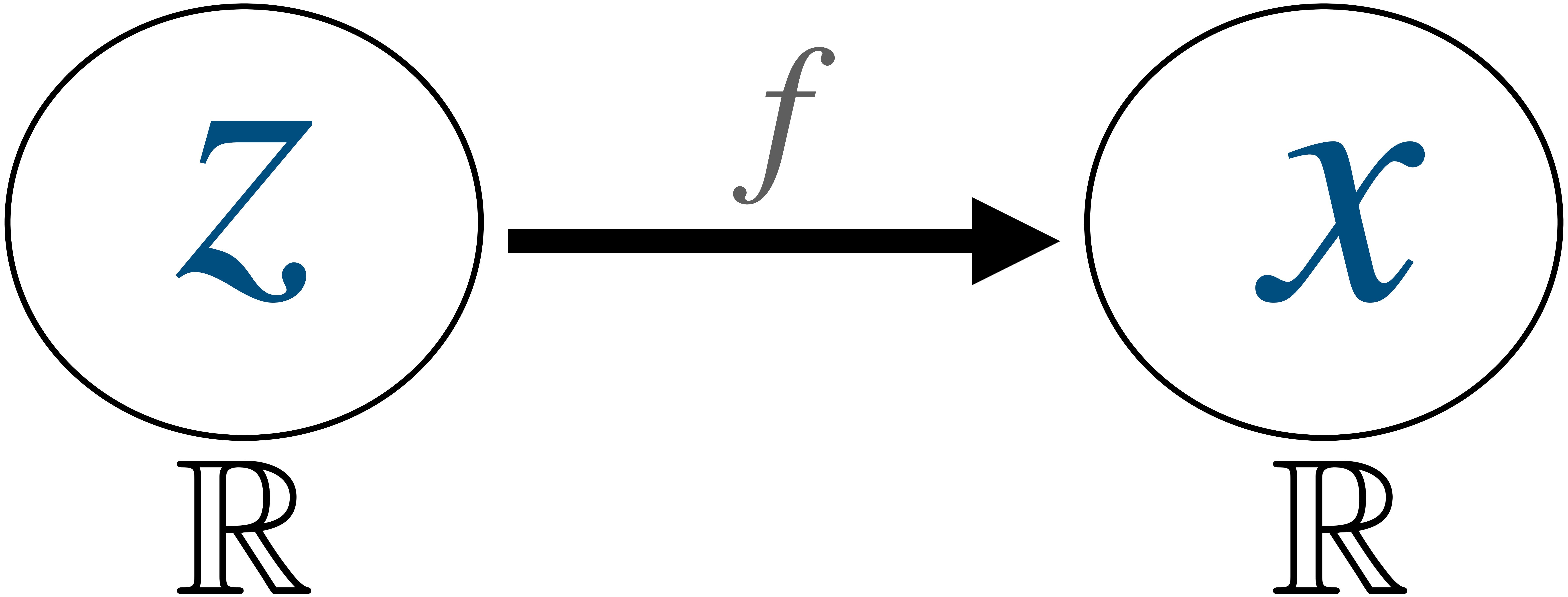


Normalising Flows

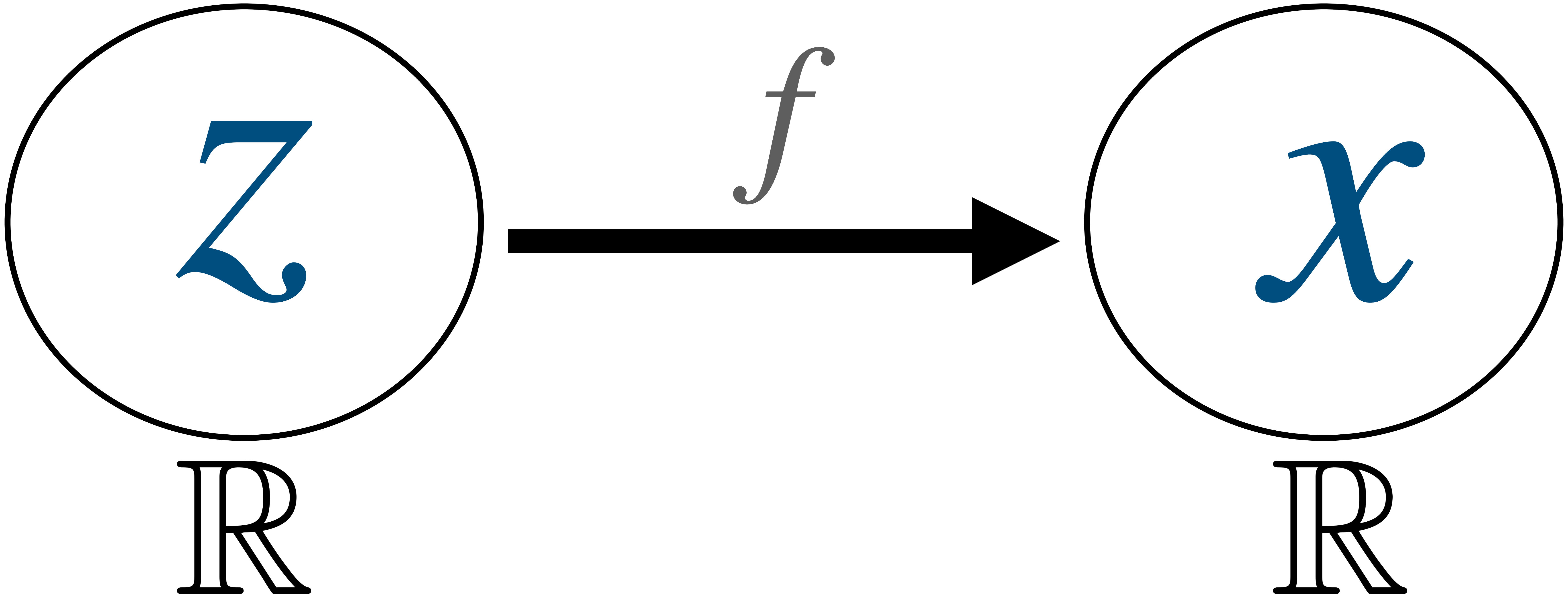
- Incredibly powerful generative modelling (e.g., Kingma & Dhariwal, 2018)
- Straightforward to both *sample* and *evaluate* new samples



Credit: Guandao Yang (Cornell)



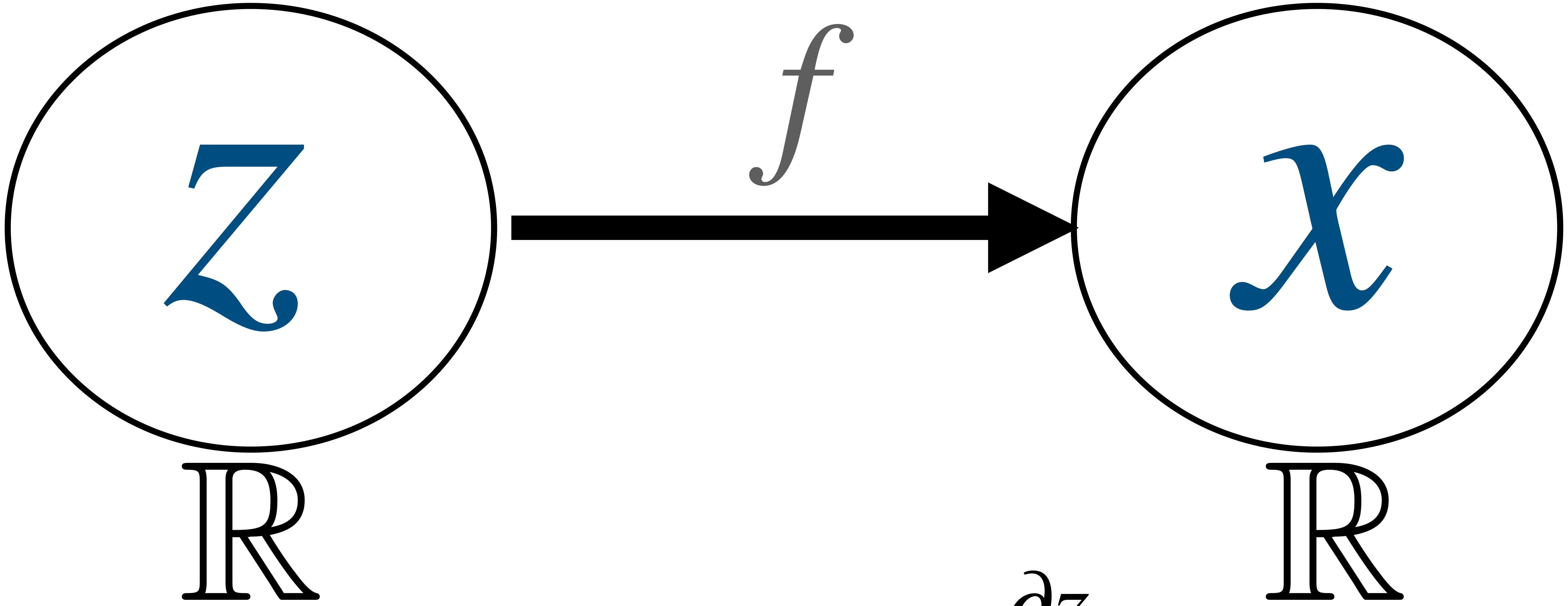
$$p_X(x)\,|\,dx\,| = p_Z(z)\,|\,dz\,|$$



Conservation of mass

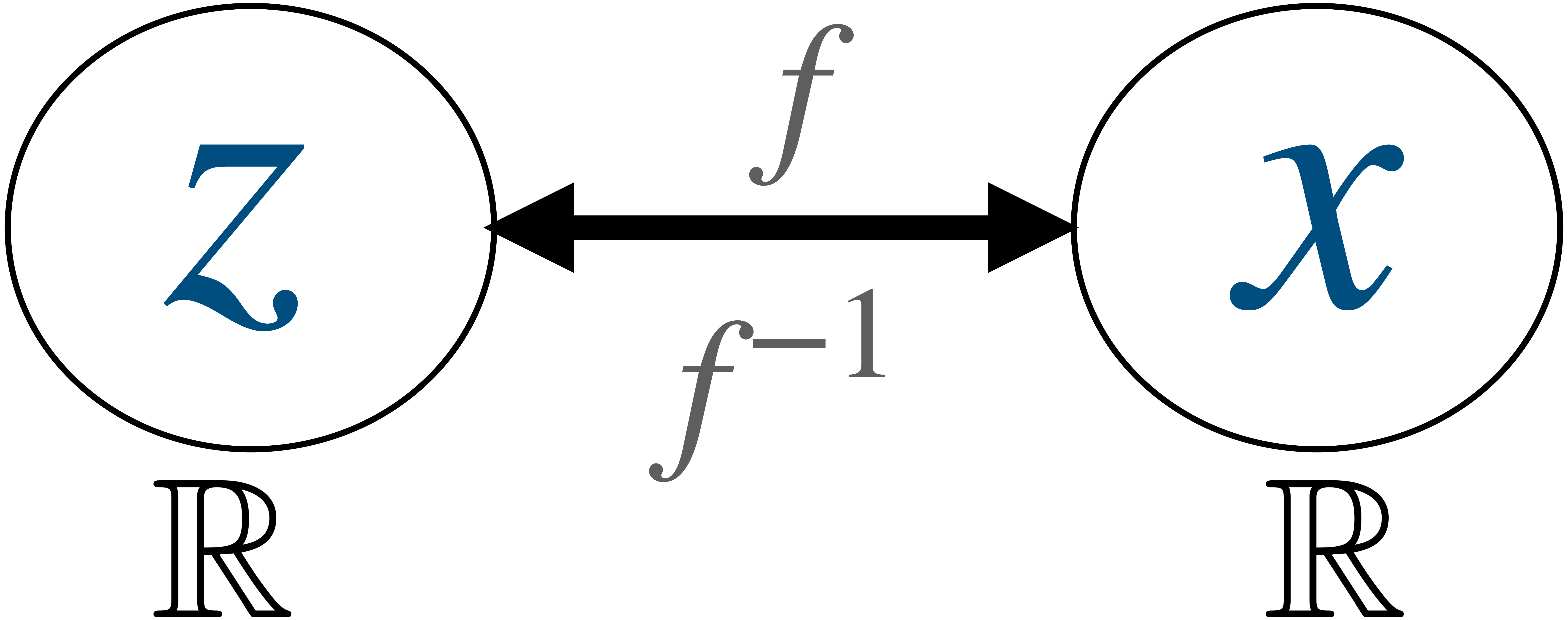
$$p_X(x)\,|\,dx\,| = p_Z(z)\,|\,dz\,|$$

Fundamental insight is *change of variable* formula:
use neural networks as change of variables

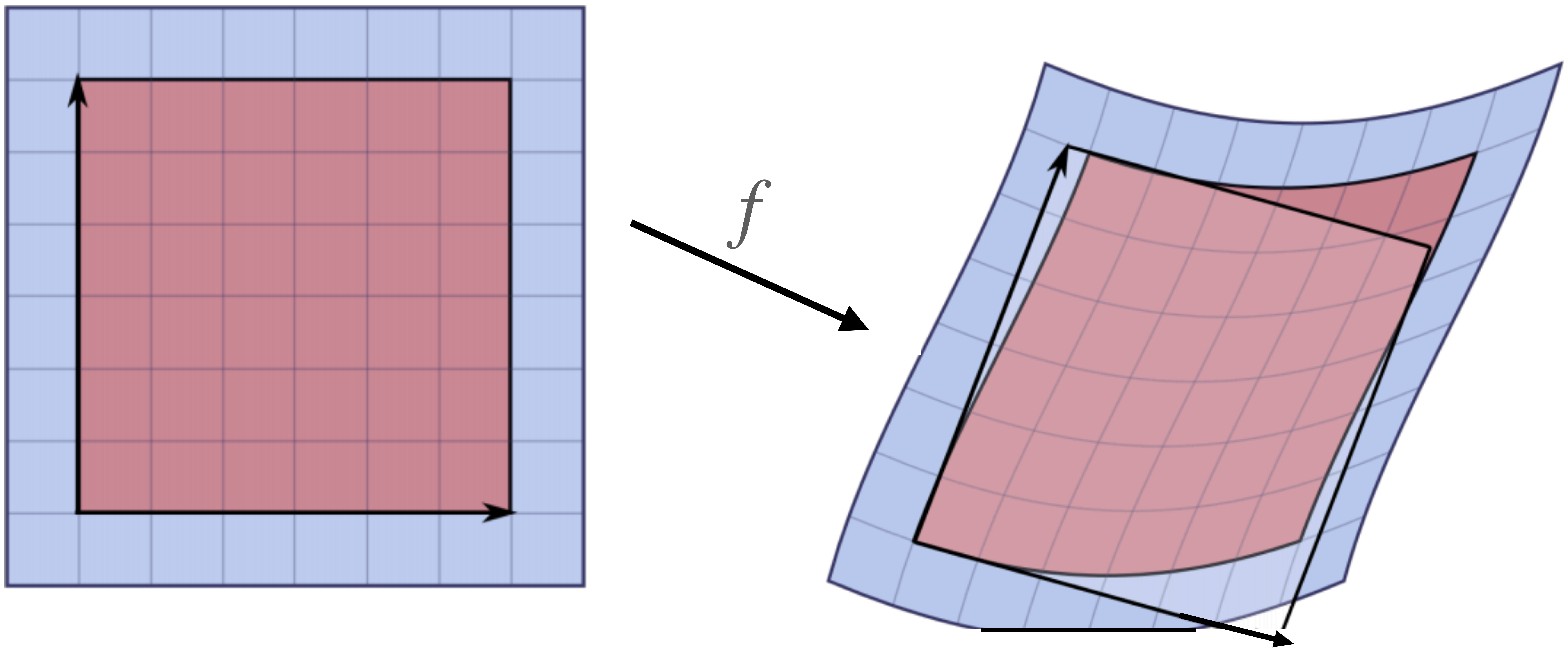


$$p_X(x) = p_Z(z)\,|\,(\frac{\partial z}{\partial x})\,|$$

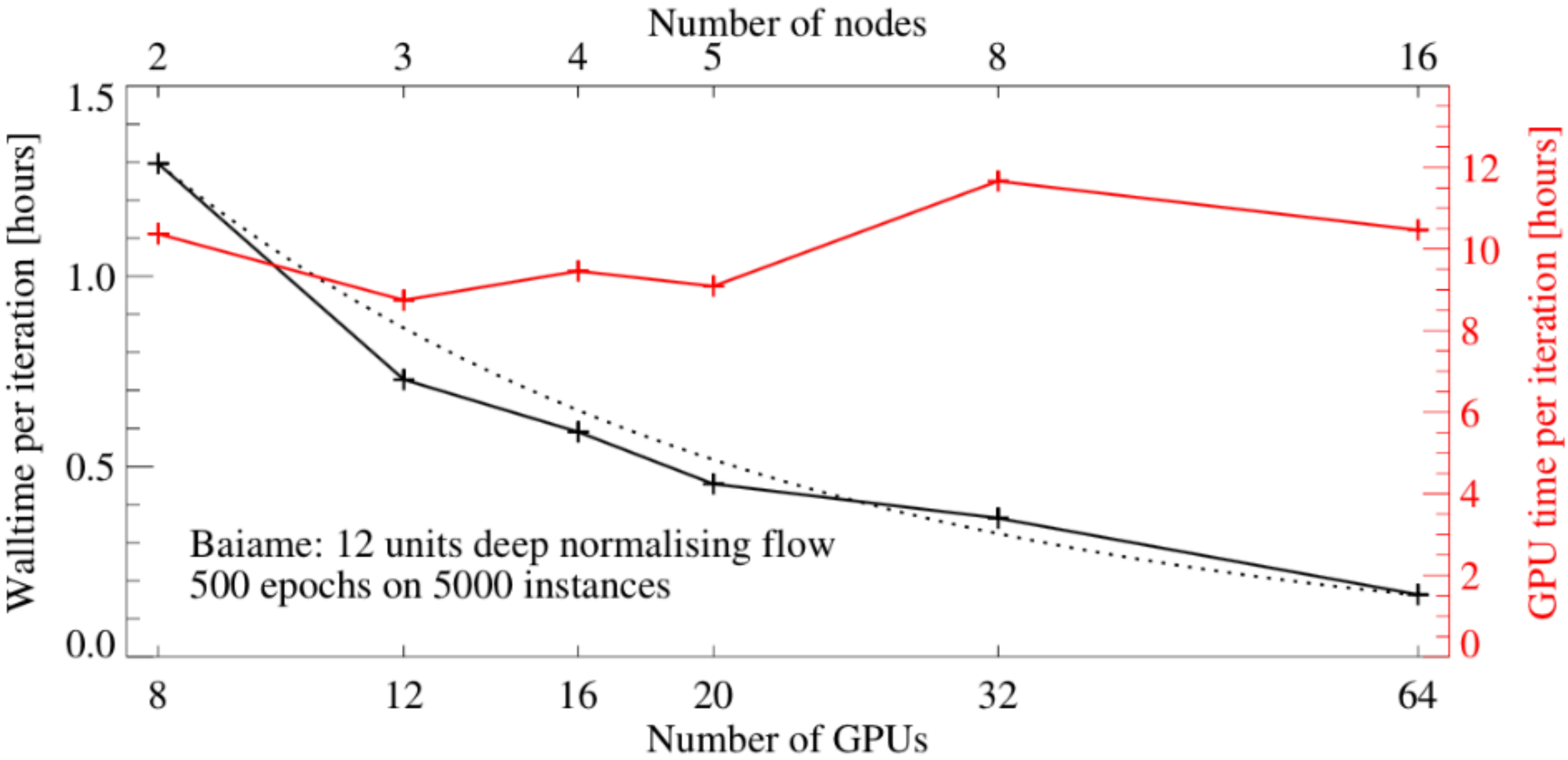
2 criteria for NF: NN has to be *invertible*.



2 criteria for NF: The determinant of the Jacobian can be computed easily.

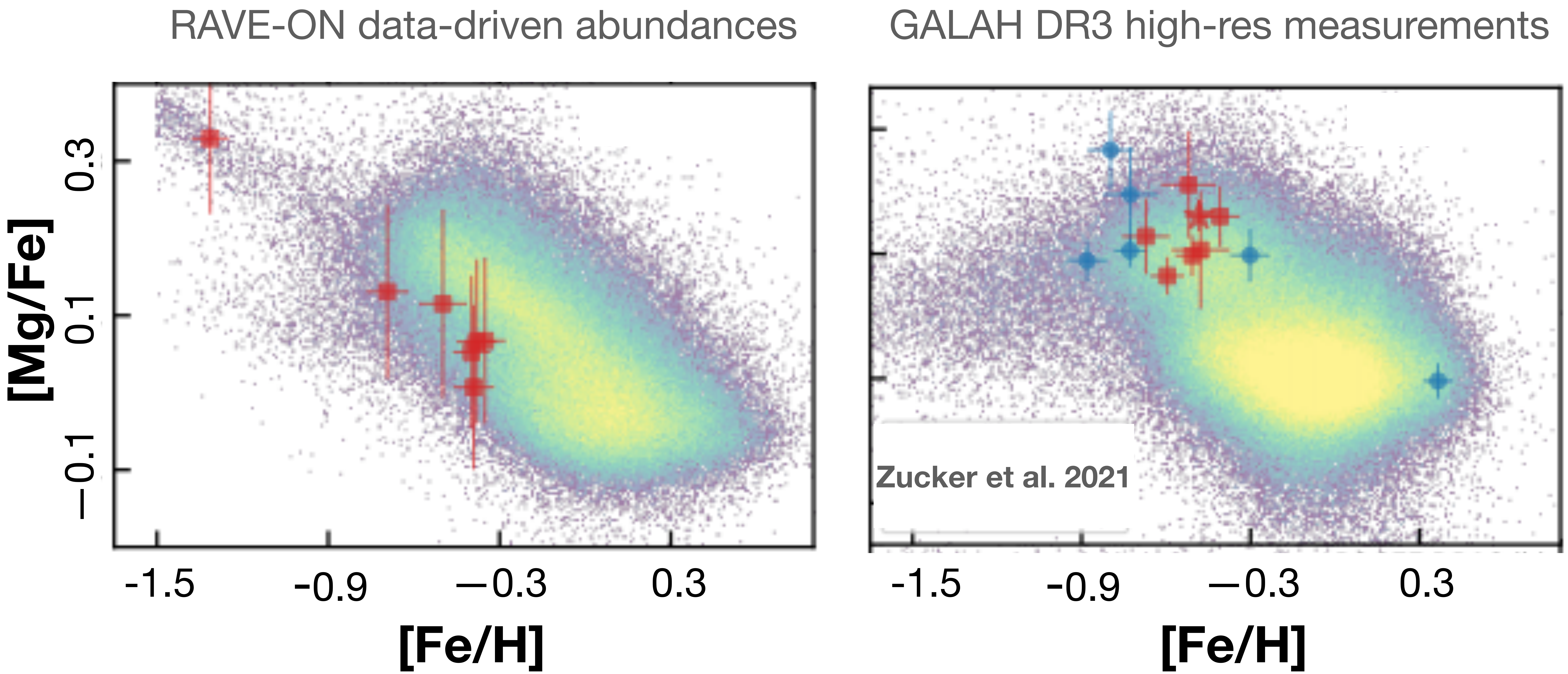


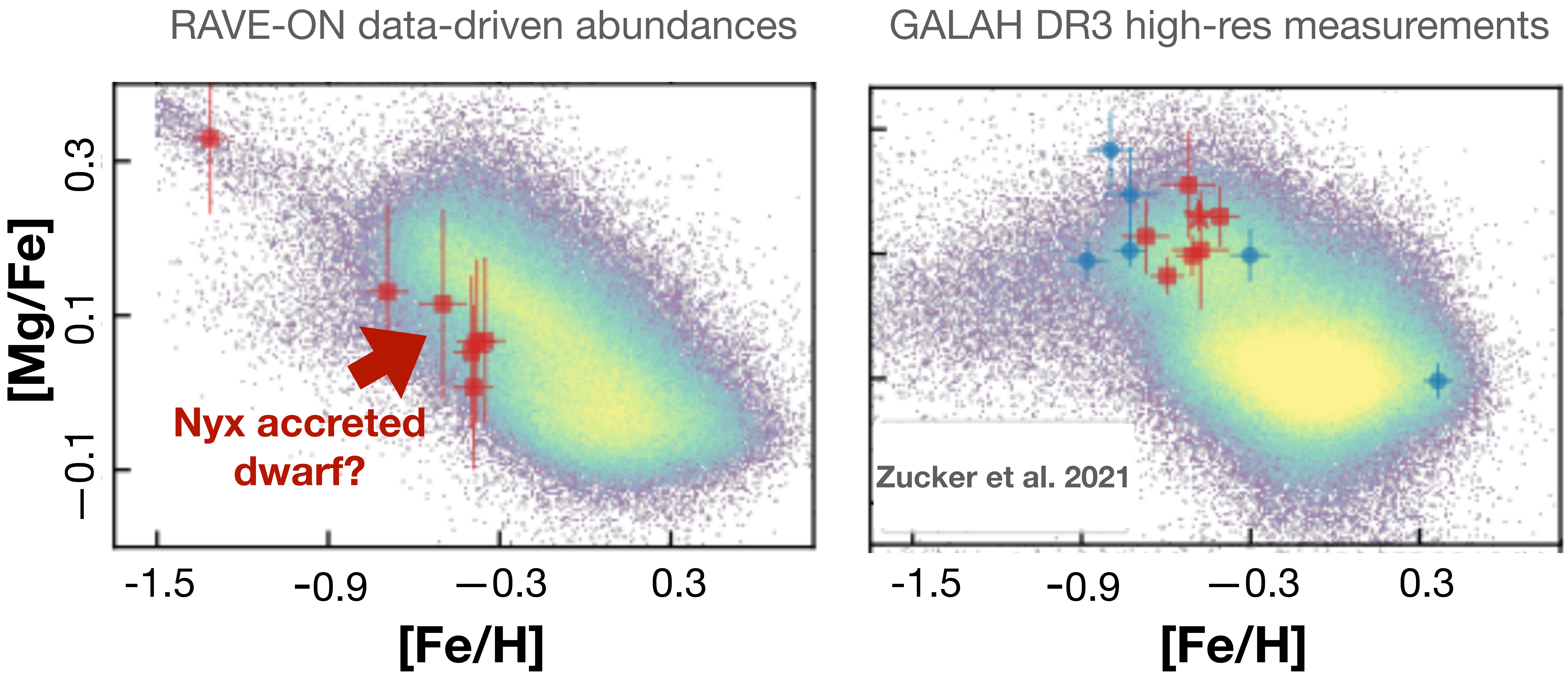
Training: PyTorch and supercomputers



The case for Galactic Archaeology

- ML spectroscopy has largely employed supervised learning, i.e. mapping *observed spectra* ➔ *chemical composition*
- **Dangerous**: Data-driven abundances could lead to the wrong conclusions
- **Powerful**: By inferring chemical abundances for > 100,000 stars we can examine the dimensionality of the chemical space of the Milky Way, which is key to understanding galactic chemical evolution

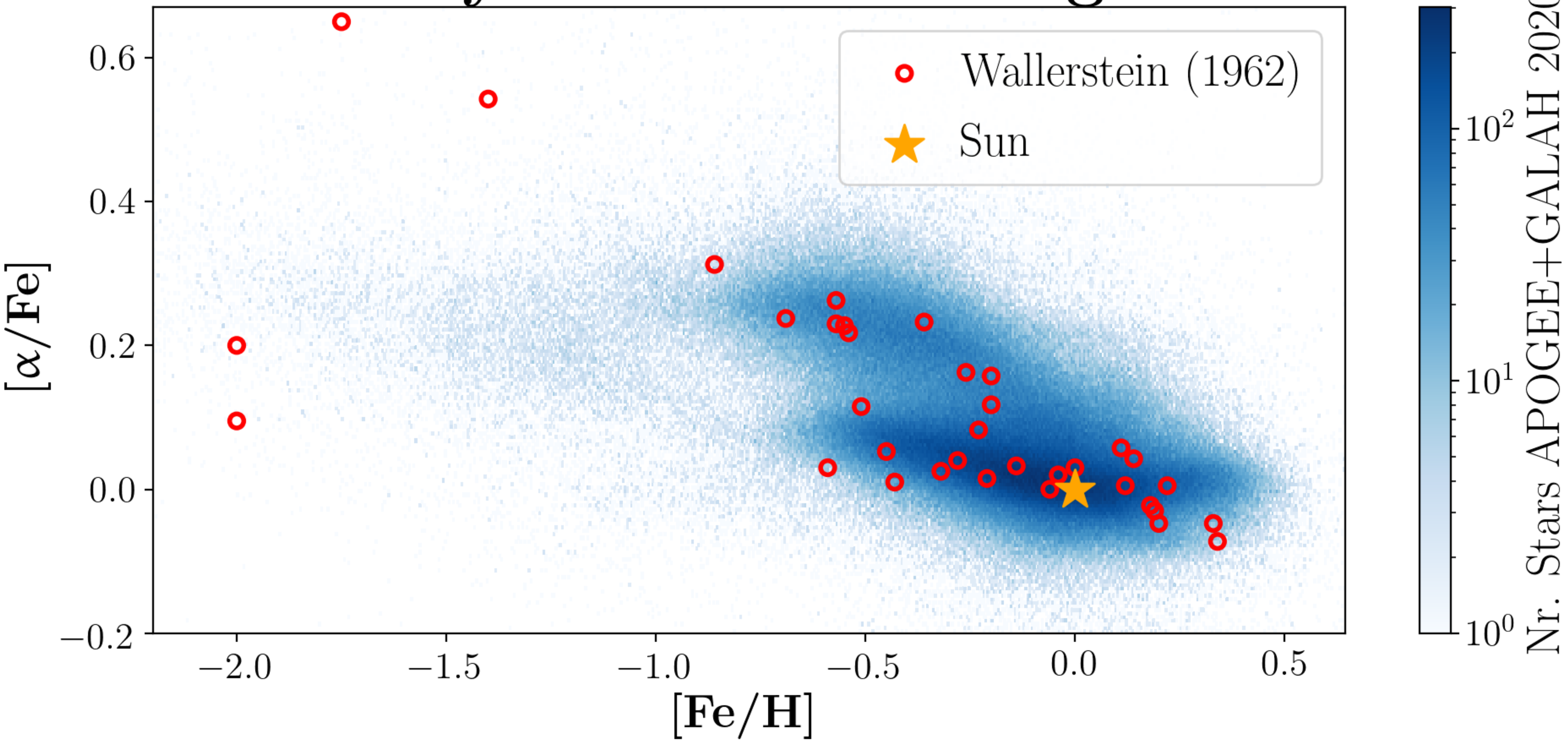

RAVE-ON data-driven abundances
GALAH DR3 high-res measurements
[Mg/Fe]
0.3
0.1
−0.1
-1.5
-0.9
−0.3
0.3
[Fe/H]
Zucker et al. 2021
-1.5
-0.9
−0.3
0.3
[Fe/H]


RAVE-ON data-driven abundances
GALAH DR3 high-res measurements
[Mg/Fe]
0.3
0.1
−0.1
Nyx accreted
dwarf?
Zucker et al. 2021
-1.5
-0.9
−0.3
0.3
[Fe/H]
-1.5
-0.9
−0.3
0.3
[Fe/H]

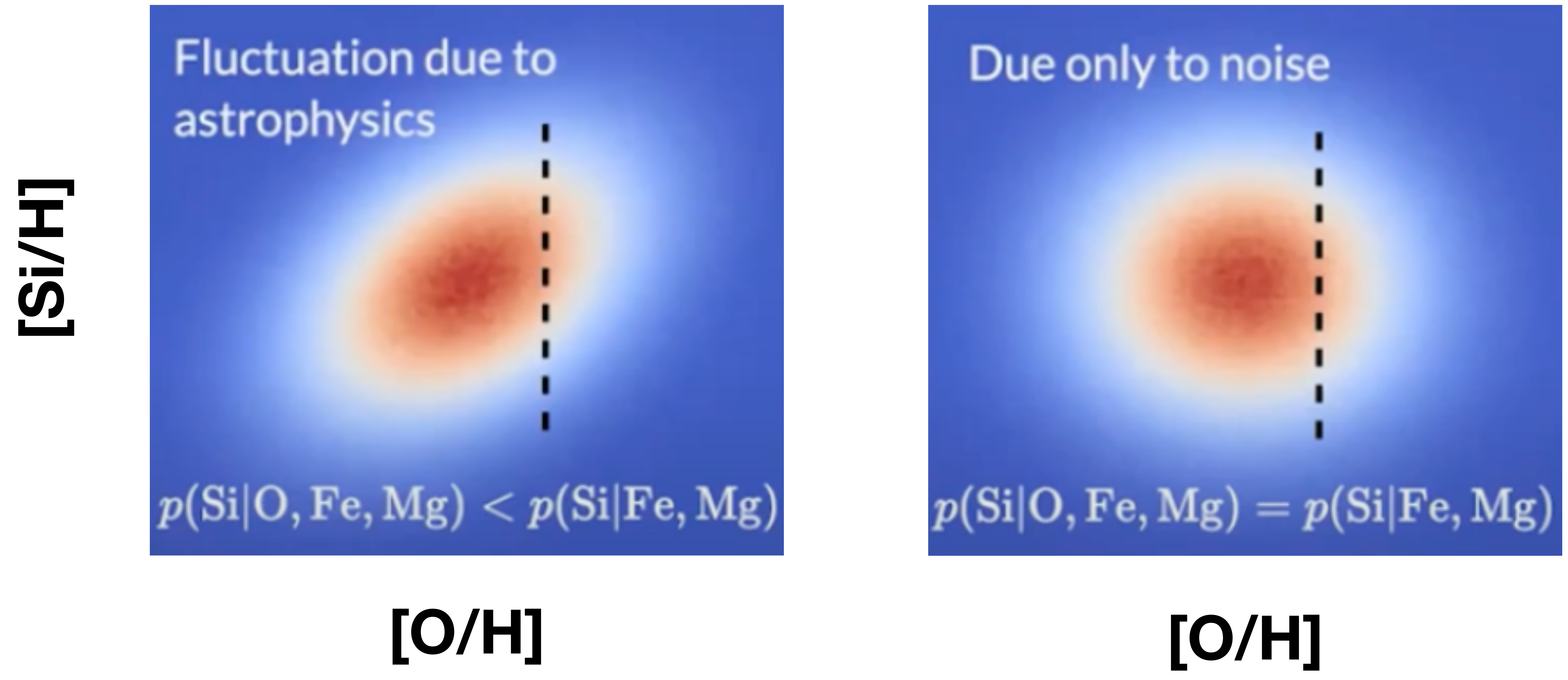
The case for Galactic Archaeology

- ML spectroscopy has largely employed supervised learning, i.e. mapping *observed spectra* → *chemical composition*
- **Dangerous**: Data-driven abundances could lead to the wrong conclusions
- **Powerful**: By inferring chemical abundances for > 100,000 stars we can examine the dimensionality of the chemical space of the Milky Way, which is key to understanding galactic chemical evolution

Tinsley-Wallerstein Diagram



“How many elements matter?” (YST & D. Weinberg 2021)



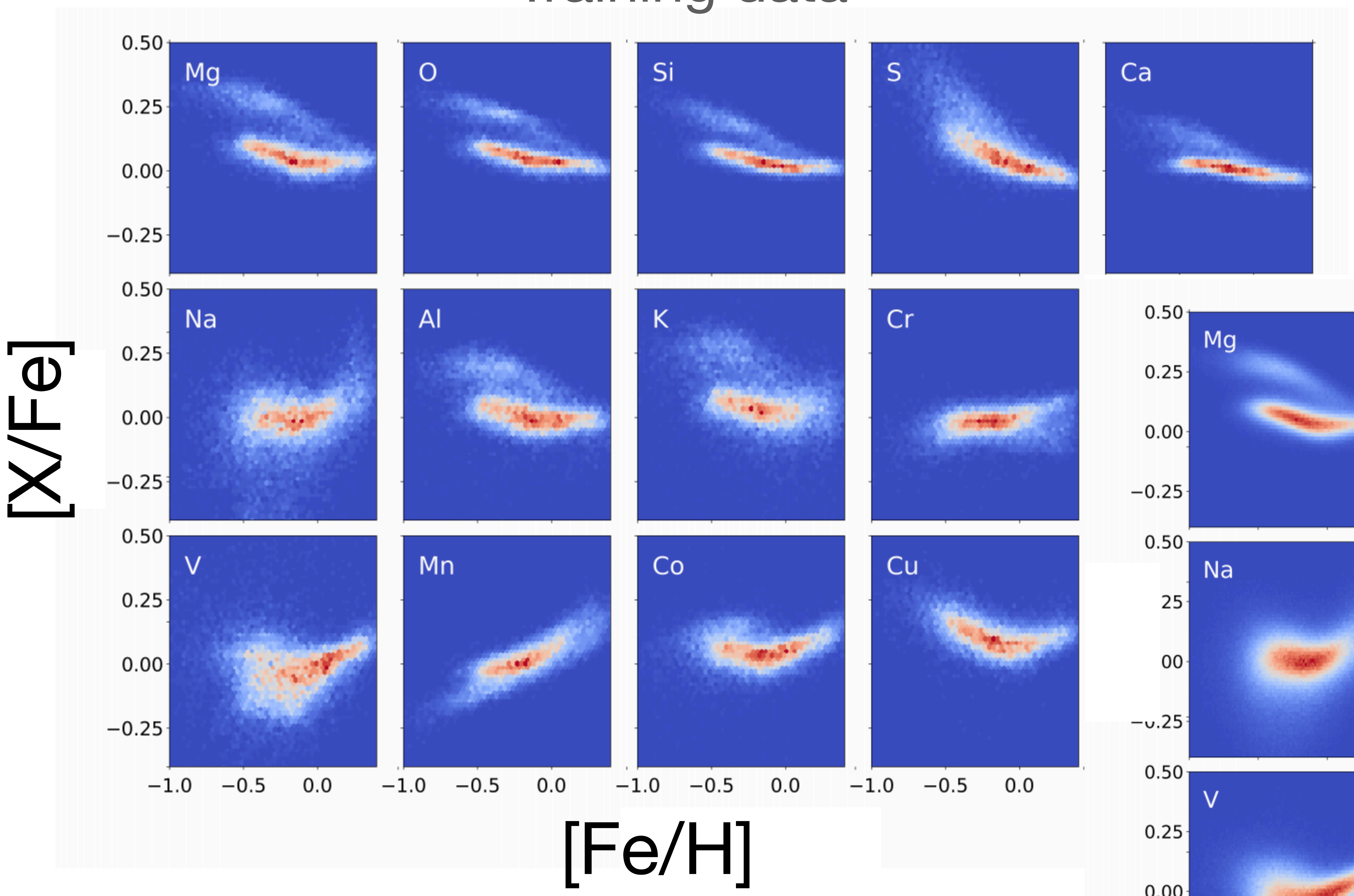
What is minimal subset of elements you need such that, upon conditioning, your dispersion is reduced to noise?

Contradictory results!

- In APOGEE, can reduce all the APOGEE abundances to noise level using only Fe & Mg (or Fe and age) ➡ dimension is 2 (e.g., Ness+19, Lu+21)
- Weinberg et al. 2019 & Griffith et al. 2020: [Mg/H] & [Mg/Fe] ➡ [X/Mg]
- YST et al. (2012), Andrews et al. (2017) & Price-Jones & Bovy (2018): more than 2 components matter

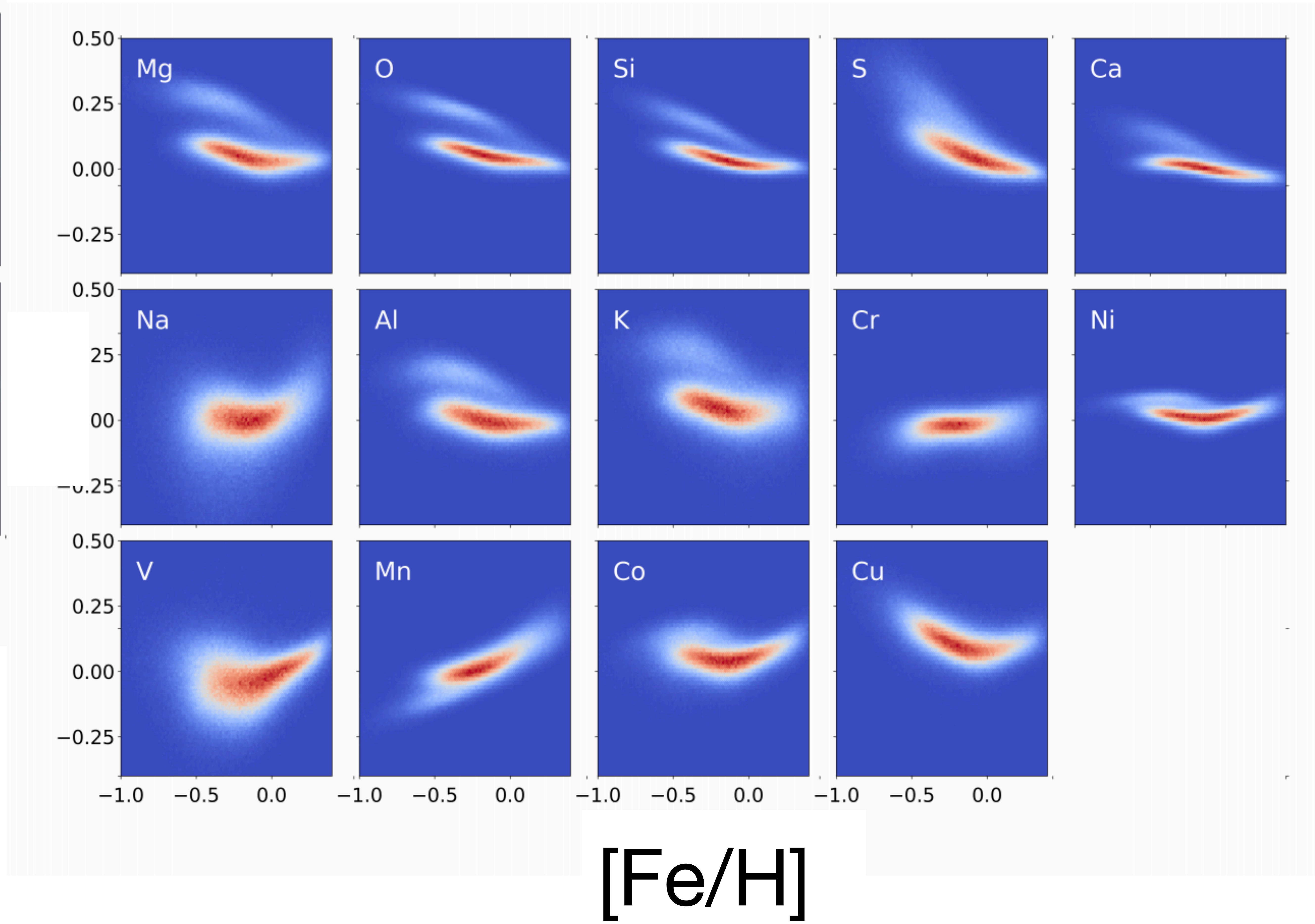
Credit: YST (ANU)

Training data



Fancy flow

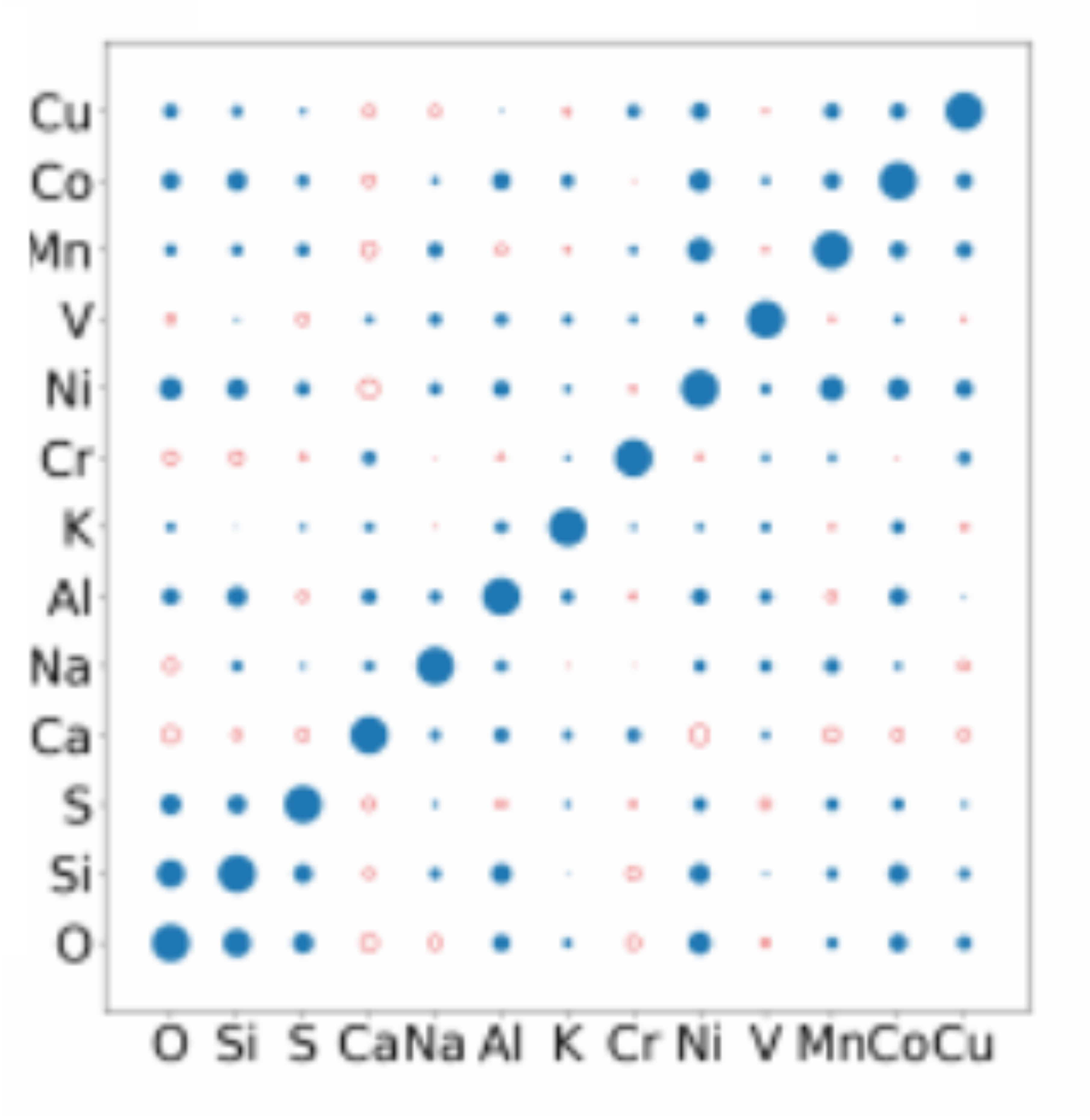
Generated data



YST & D. Weinberg 2021

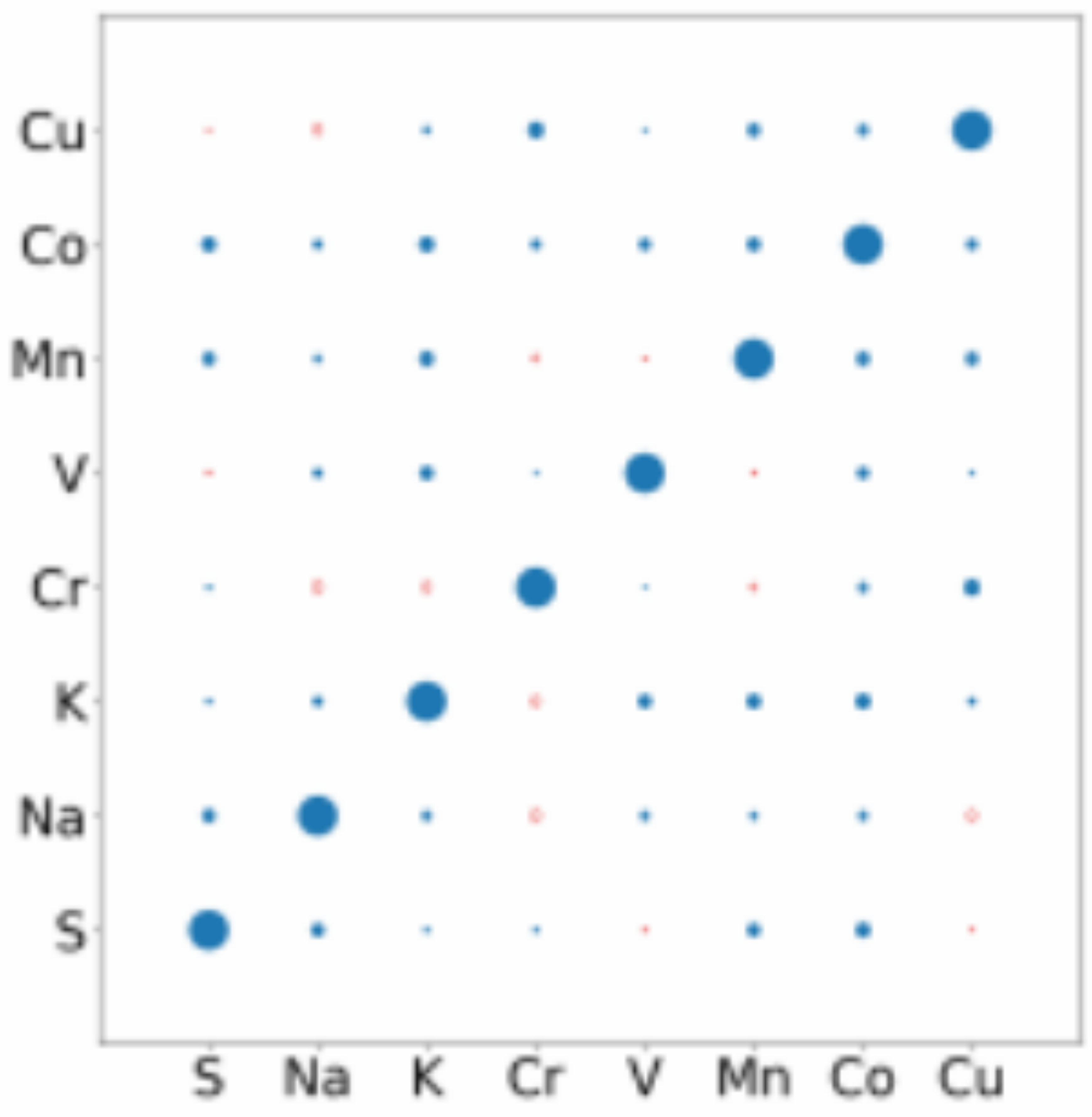
Conditioned on Fe & Mg

$[Fe/H] = 0, [\alpha/Fe] = 0$



Conditioned on Fe, Mg, O, Si, Ca and Al

$[Fe/H] = 0, [\alpha/Fe] = 0$



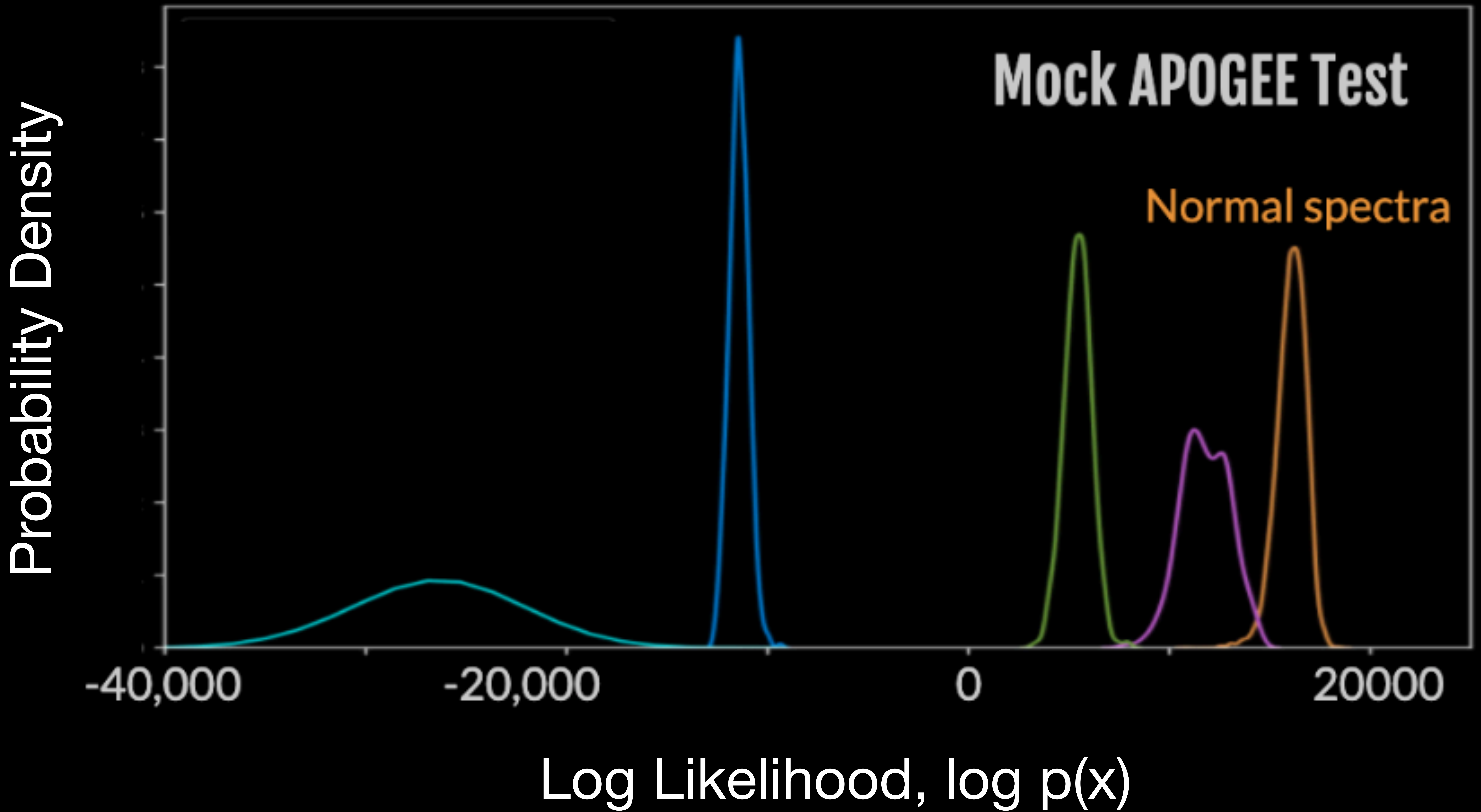
YST & D. Weinberg 2021

High-dimensionality applications

I. Find chemical outliers

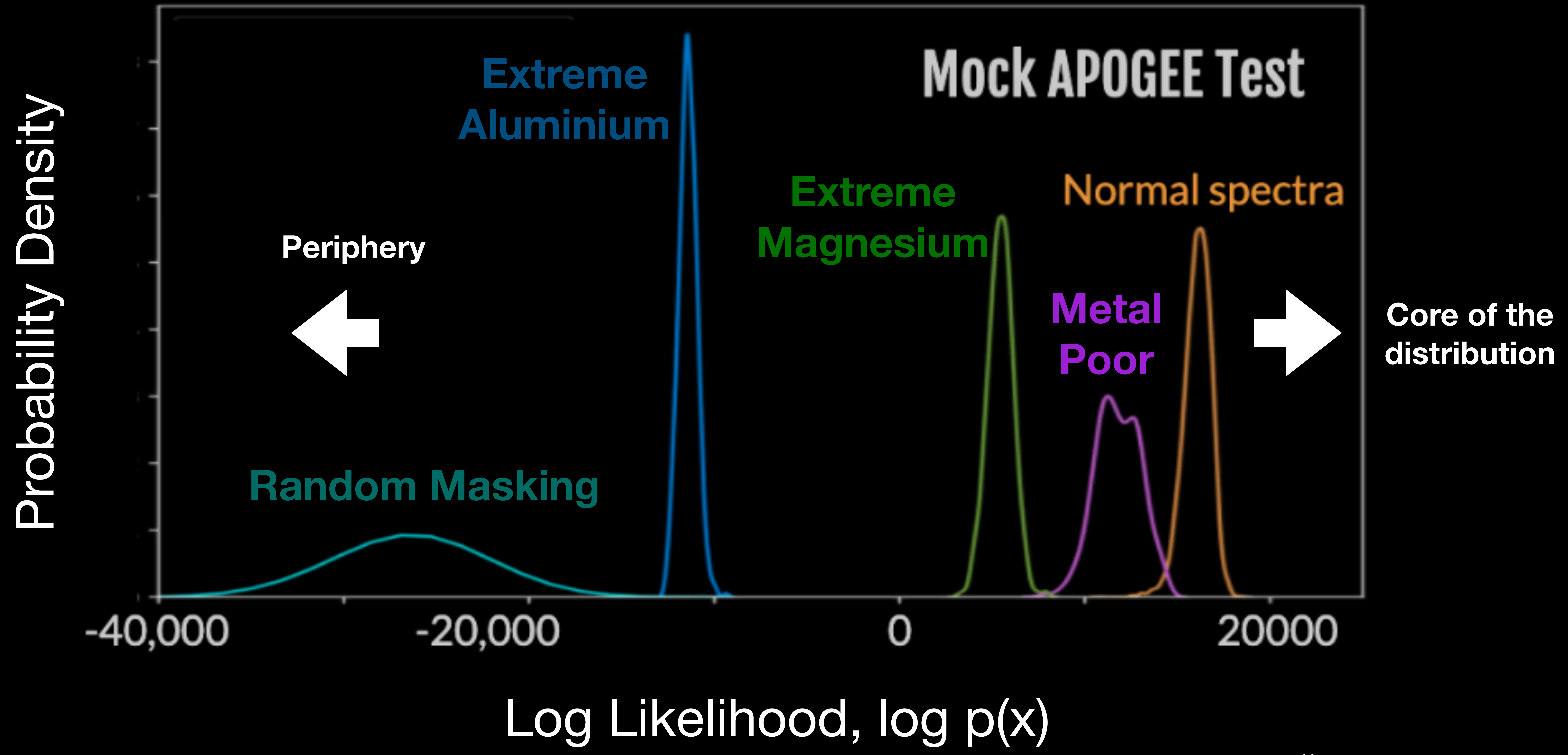
II. Identify hidden atomic transitions

I: one in a million



Ciucă & Ting, In Prep.

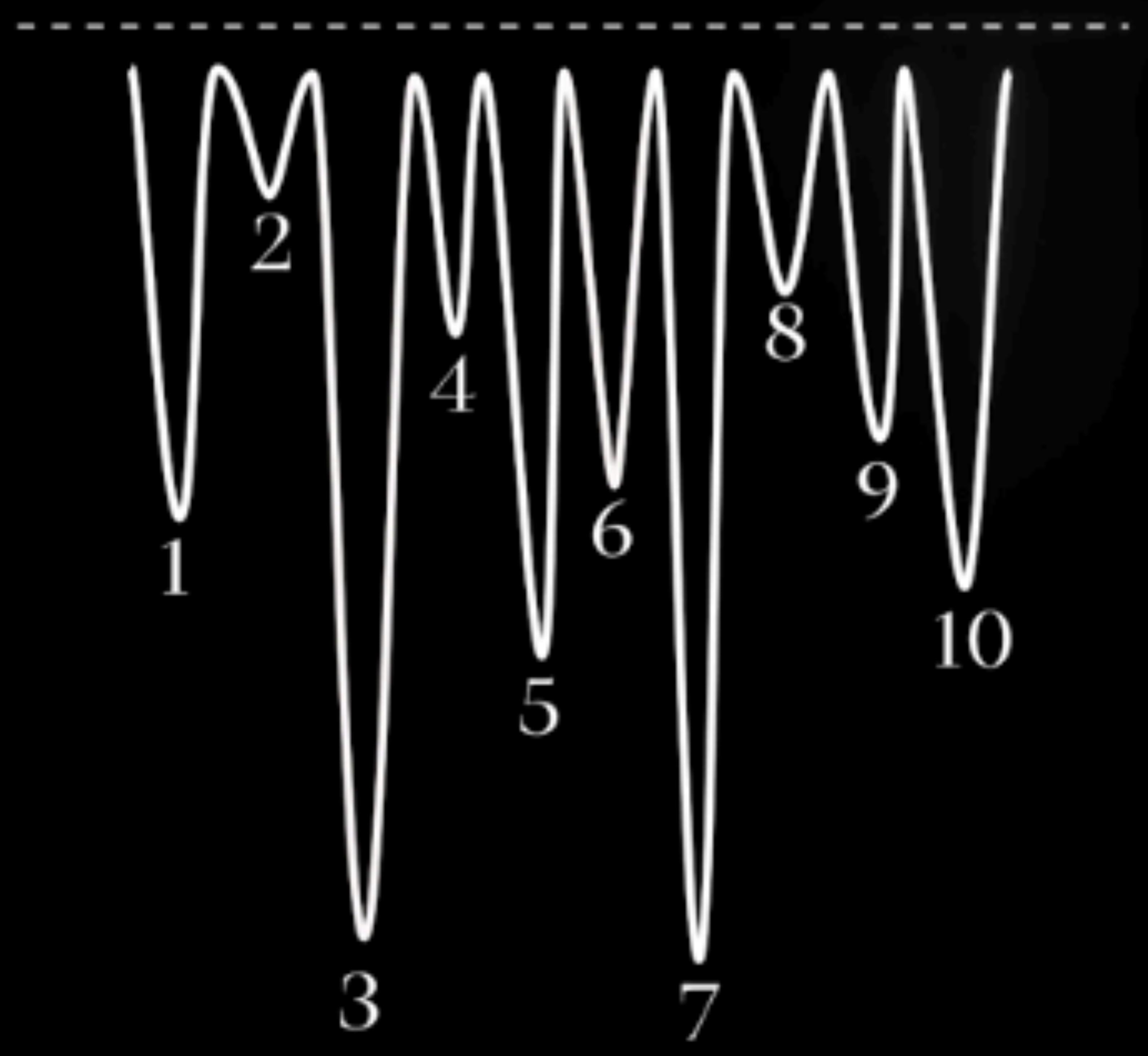
I: one in a million

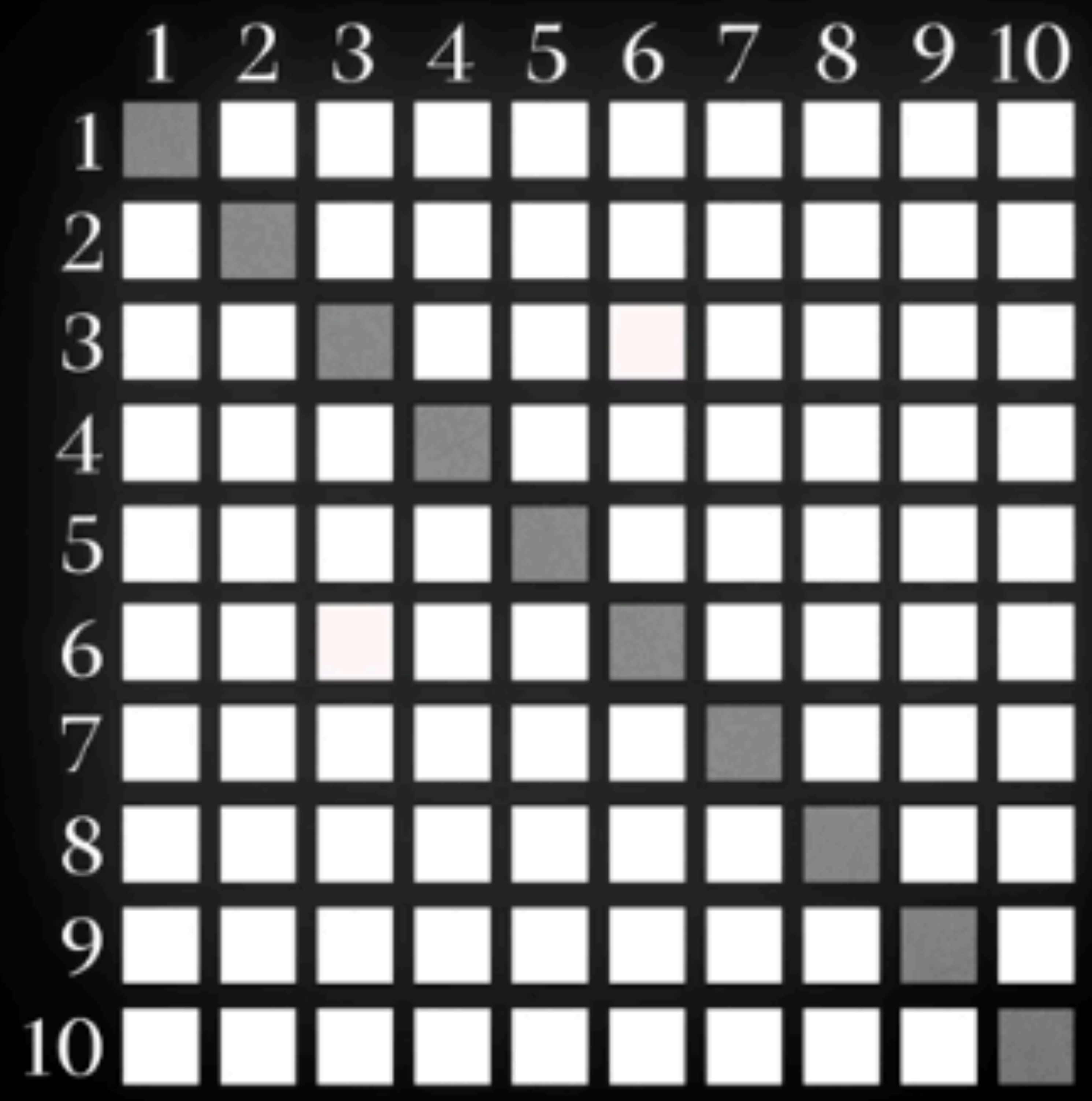


Ciucă & Ting, In Prep.

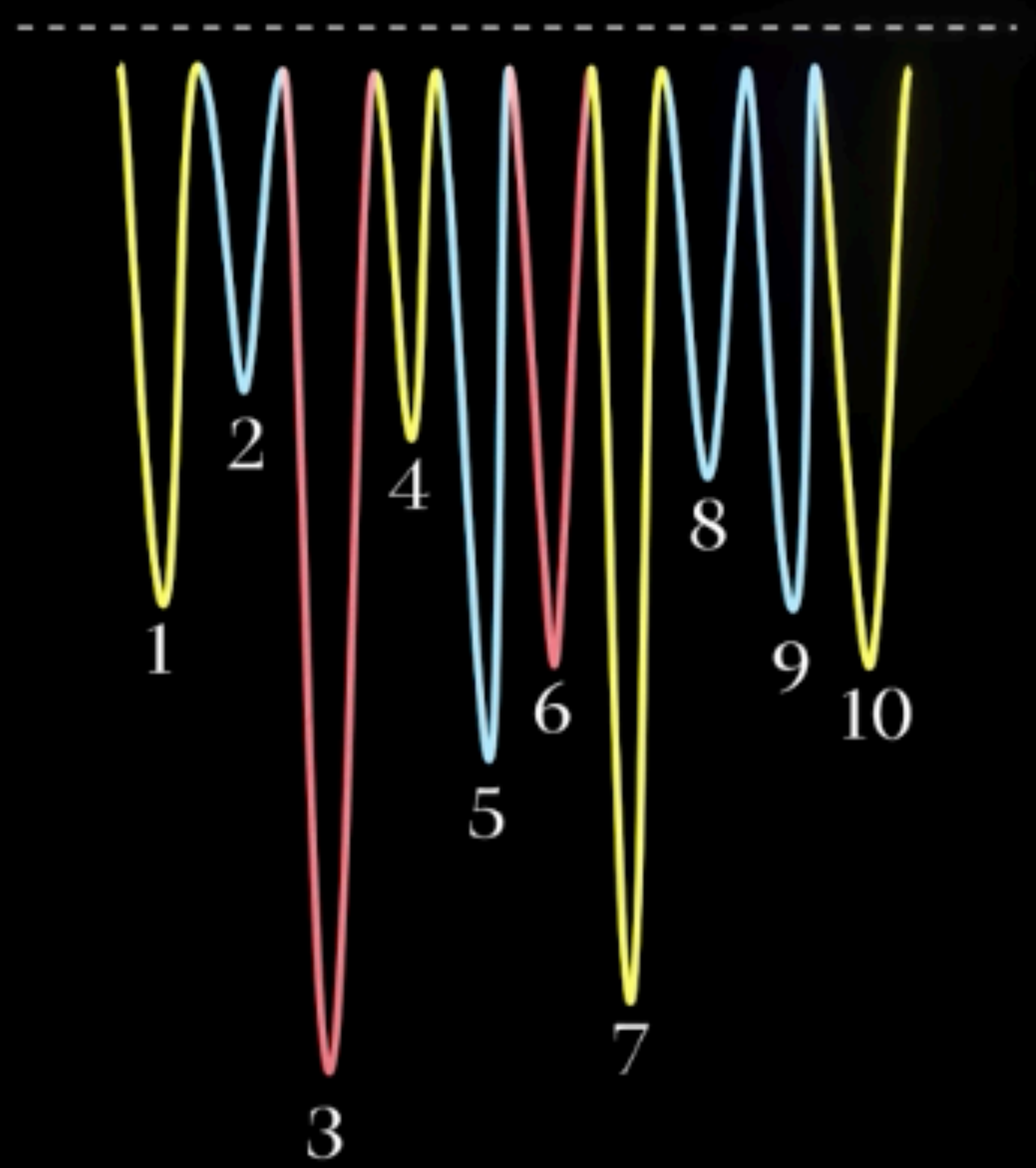
II: what's lurking in starlight?

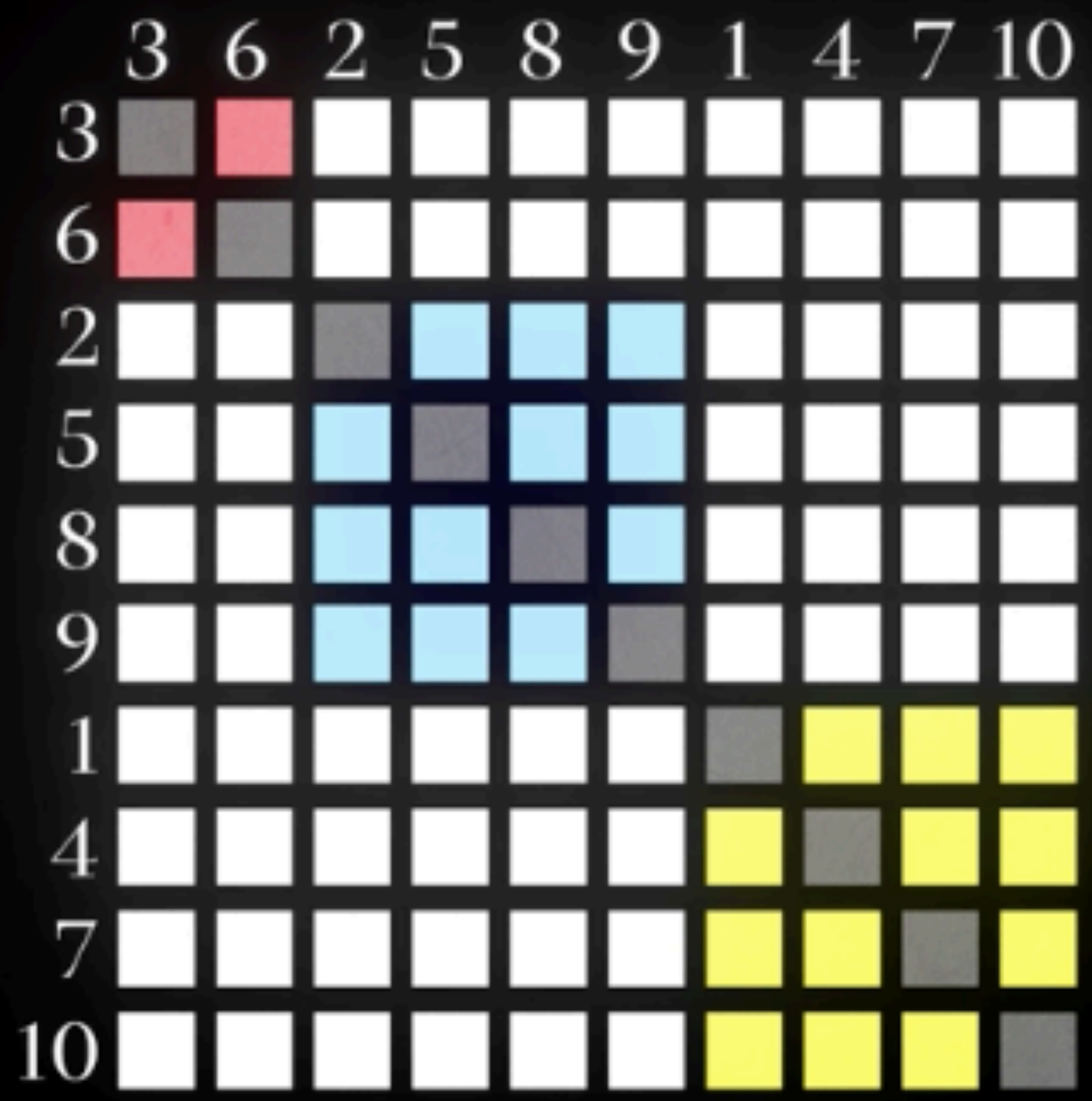
$$p\,(x_i|\mathrm{Teff},\mathrm{logg}) \;\Rightarrow\; \mathrm{Corr}(x_i, x_j)$$





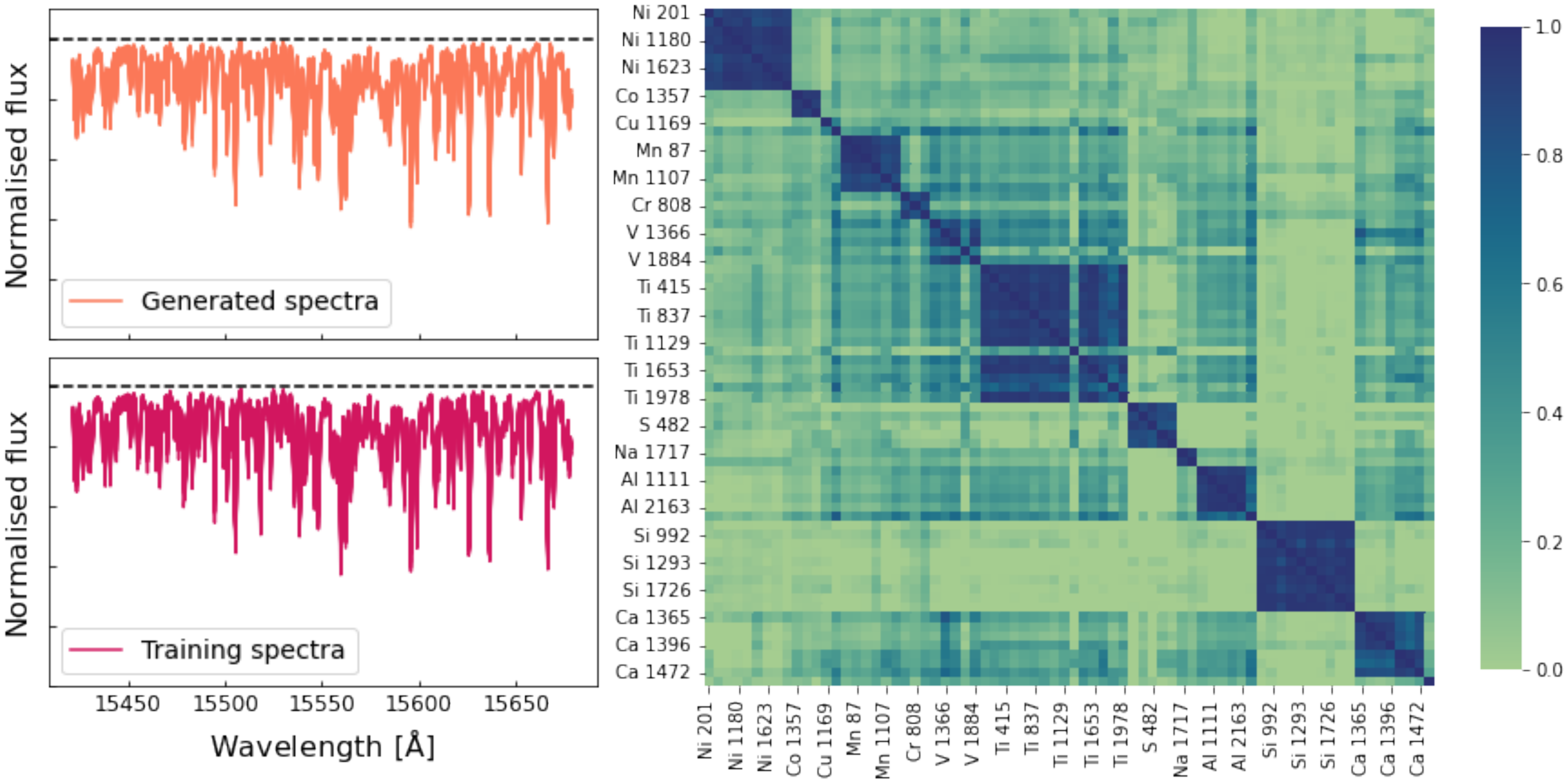
II: what's lurking in starlight?





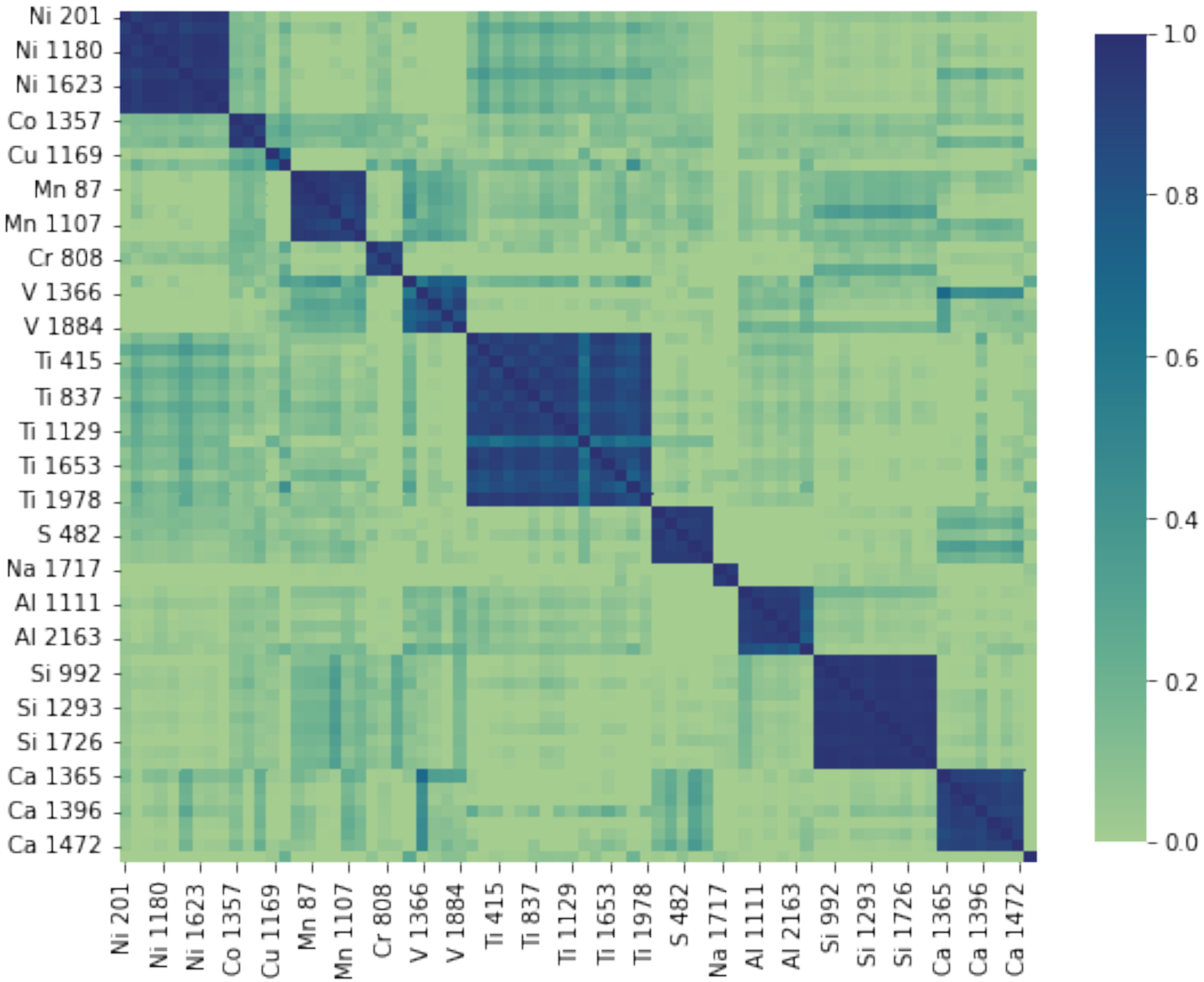
Modeling 2,000 dimensions

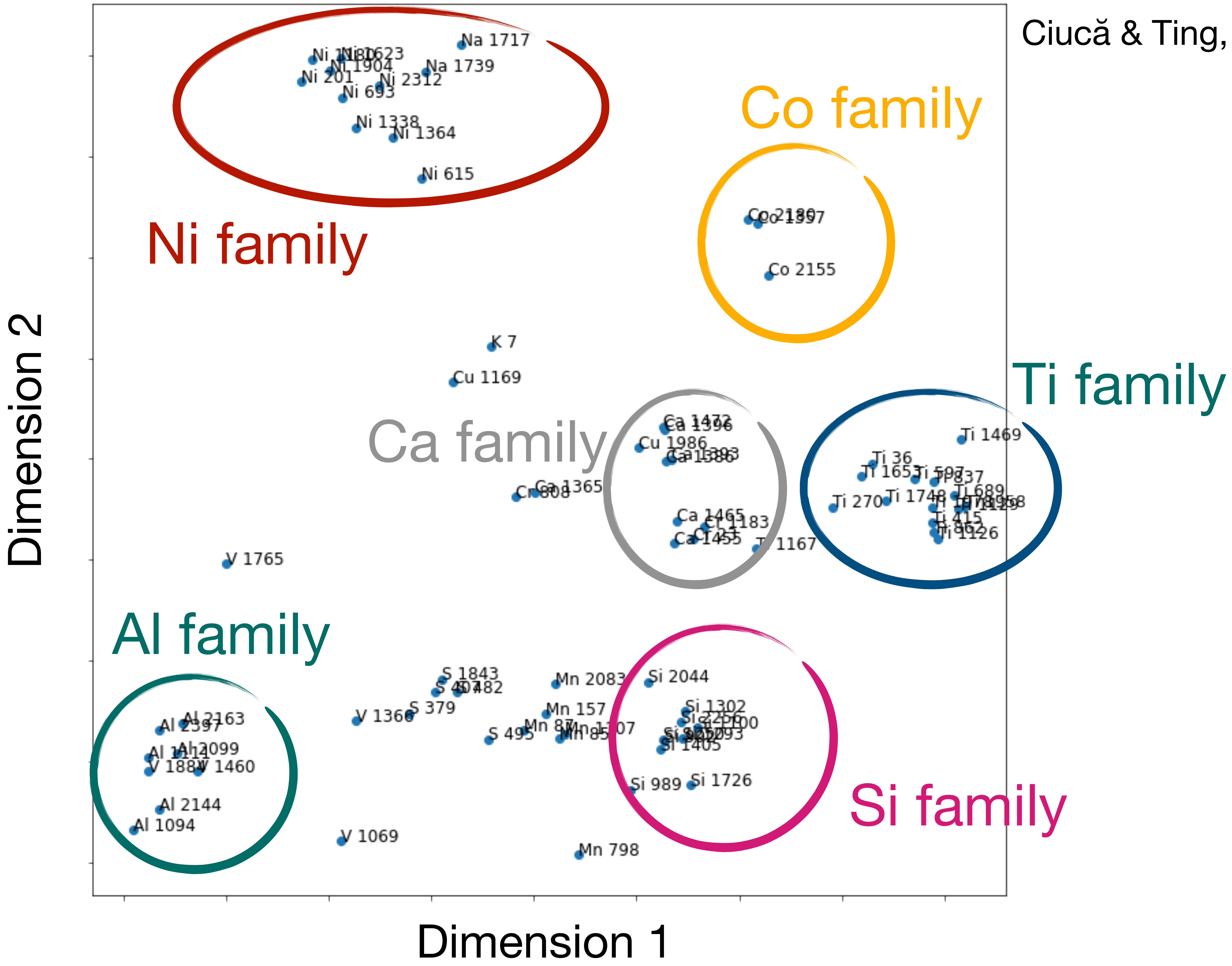
Ciucă & Ting, In Prep.



Generating samples and conditioning

Ciucă & Ting, In Prep.





Ciucă & Ting, In Prep.

Summary

- NF: flexible NN-based mapping & provide exact likelihood /w great potential for Galactic Archeology
- Stellar spectra: high-dimensional, up to 10,000 dimensions can be modelled via Normalising Flows
- 2 main applications: outlier detection and examine empirical line correlations
- Challenges ahead: noise, proper conditioning on labels

Kiitos paljon! Any questions?

ML Resources

- Andrew Ng: Intro to Machine Learning course on Coursera, notes here http://www.holehouse.org/mlclass/ (old), (new) http://cs229.stanford.edu/syllabus.html

- UCL & DeepMind deep learning course: https://deepmind.com/learning-resources/deep-learning-lecture-series-2020 (youtube also available)

- Deep Generative Modelling: https://deepgenerativemodels.github.io/notes/index.html

- Cool stuff on normalising flows https://github.com/janosh/awesome-normalizing-flows