Identifying chemically identical stars using adverserial disentanglement

Damien de Mijolla
Department of Physics and Astronomy,
University College London,
Gower Street, WC1E 6BT, UK
ucapdde@ucl.ac.uk

Melissa Ness
Department of Astronomy,
Columbia University,
Pupin Physics Laboratories, New York, NY 10027, USA

Serena Viti
Department of Physics and Astronomy,
University College London,
Gower Street, WC1E 6BT, UK

Abstract

As the chemistry of stars is interlinked with their birthsites, identifying stars with identical chemical abundances is a promising avenue for unravelling the history of our galaxy. We present a method for precisely identifying chemically identical stars from spectroscopic observations. In our approach, a neural network is trained to learn a latent which is disentangled from non-chemical factors of variation. This allows for identifying chemically identical stars without any reliance on chemical models.

1 Introduction

Galactic archaeology is a rapidly growing field of astronomy in which observations of stars are used to reconstruct the history of our galaxy. Chemical tagging, the process of identifying chemically identical stars from their spectra, is a particularly promising technique that could help with estimating the number of star-forming clusters in the galactic disk (e.g. [12] and [15]) and how stars have moved over time (e.g. [1] and [7]). This is because chemically identical stars are expected to have the same or nearby initial birth sites.

Multiple spectroscopic surveys such as APOGEE [16] and GALAH [6] have given the astronomical community a wealth of high-resolution stellar spectra, enabling galactic archaeological studies. However, chemical tagging will require extremely precise chemical composition estimations for many practical use cases. The current methodology for estimating chemical compositions involve comparing observed spectra to libraries of simulated spectra obtained from running complex astrophysical simulations. In practice, because of imperfections in the modelling and poorly constrained parameters in the radiative-transfer modelling, there is a model mismatch between astrophysical simulations and reality. This forward-model mismatch leads to large systematic uncertainties in the estimated chemical compositions.

Second Workshop on Machine Learning and the Physical Sciences (NeurIPS 2019), Vancouver, Canada.

In this work, we apply supervised disentanglement learning using neural networks to address the problem of astrochemical tagging. Our primary contribution is to show how model misspecification in chemical estimation can be bypassed through supervised disentanglement. Our method, which bridges the gap between simulation driven science and deep learning, consists of training a neural network to auto-encode spectra using a latent disentangled from all factors of variations unrelated to the chemical compositions. To our knowledge, this is the first application of supervised disentanglement that makes use of the disentangled latent without further processing, allowing our task and associated dataset to be particularly useful for evaluating disentanglement algorithms.

2 Related Work

Disentangled representation learning:There exists a substantial body of literature on the subject of leveraging neural networks for learning disentangled representations. This includes fully unsupervised methods such as beta-vae [8] and infogan [5] but also supervised methods [10] [4]. In the case of unsupervised methods, disentanglement involves learning a latent representation in which different factors of variations, for example lighting conditions and object orientations in images, are encoded in distinct portions of the latent. Supervised methods extend this framework by specifying labels for factors of variations from which the latent learnt by the neural network should be independent after training. When this latent is combined with a conditional decoder conditioned on the labels, it becomes possible to selectively modify inputs to the neural network such that all aspects of the inputs are unchanged apart from those controlled by the factors of variation.

Astronomical chemical tagging: In recent year, there has been a breath of research around chemical tagging. This has involved research into assessing the feasibility of chemical tagging [21] [19]. Attempts at applying chemical tagging [9], at improving the accuracy of estimated parameters for chemical tagging [14], [24], [18] as well as novel methodologies for carrying out chemical tagging. Novel methodologies include those presented in [2] and [11] which use techniques from the field of phylogeny. The work presented in [22], similarly to us, attempt to identify stars without requiring detailed chemical modelling. They do this by removing the effects of parameters other than abundances through a polynomial fit. Contrary to our appraoch, this is a relatively inflexible model that does not fully consider cross-dependencies between chemical compositions and other parameters.

3 Method

Given a dataset of stellar spectra, our goal is to identify groups of stars sharing a common chemical composition. Chemical composition is traditionally estimated by comparing observational stellar spectra to synthetic spectra derived from stellar models and then identifying the best fitting synthetic spectra. However because atomic line strengths are dependent on complicated and poorly constrained radiative transfer physics, abundances are usually estimated from a fraction of the spectral lines present in the dataset using imperfect simulations. The reliance on a limited number of lines and model mismatch both hinder chemical abundance estimations.

Here, we take for chemical tagging an approach similar to that used for face recognition and person re-identification [23]. We associate to every stellar spectra a latent, and we seek to learn, using a neural network, a mapping such that stars with identical compositions map to identical or near-identical latent representations. We can then identify stars with identical compositions by finding spectra that are particularly close according to a distance metric of choice. However, unlike in [23], our problem is complicated by the fact that we cannot train a fully supervised approach since, for real stars, there does not exist a large, representative dataset for which we have perfect knowledge of the chemical composition. This makes fully supervised training impossible and drives us towards our proposed approach which consists in training a neural network to solve a disentanglement learning training task, which after training results in a latent with the properties discussed above.

The auxiliary task we use to train the neural network is a supervised disentanglement learning task. In this task, we rely on standard methods to estimate all of the physical parameters responsible for "non-chemical" variations in the spectra. For the case of red giant type stars, as considered in this paper, the non-chemical factors of variation will be the effective temperature T_{eff} and the surface gravity log_g. We then train a neural network with a supervised disentanglement loss term to

learn a latent disentangled from the non-chemical factors of variation but for which it is possible to reconstruct the spectra using the latent combined with the non chemical factors of variations. The intuition behind our method is that, by enforcing that the latent representation is disentangled from the non-chemical factors of variation, we ensure that our latent contains no information about the non-chemical factors of variation. At the same time, since we require it to be possible to reconstruct the spectra from the latents combined with the non-chemical factors of variations, we are forcing the latent to encode all of the information about chemical factors of variations. This leads to a latent "containing" all the chemical factors of variations but nothing else.

With this method, the neural network can learn to identify chemically identical stars without ever explicitly being taught about chemical composition. In other words, our approach bypasses the need for radiative-transfer modelling of spectral lines and replaces it with a disentanglement learning task. Using the neural network offers two significant benefits compared to traditional methods for estimating chemical composition. Most importantly, our method is able to precisely identify chemically identical stars without any chemical modelling. Secondly, since chemical models are not perfectly known, unlike for simulation-driven estimation, the neural network can leverage the whole spectral range for estimating abundances, leading to more precise estimates.

Our method still comes with restrictions. The representation learnt by the neural network can only be a good proxy for the chemical abundance, if the chemical abundances are themselves actually fully disentangled from the temperature and logG. This assumption is likely to be a very good approximation for red-giant type stars (which are some of the most observed type of stars in current astronomical surveys) but may not be as good for other stellar types. However, by adding the metallicity to the disentangled parameters it may still be possible to apply this method to other types of stars. Additionally, we must account for all of the non-chemical factors of variations for our method to work.

4 Experiments

Because of the lack of reliable labels associated with real spectra and because we wanted to evaluate our method in a controlled environment where the independence between the abundances and non-chemical parameters was guaranteed and we were confident that there were no unaccounted parameters (as might for example have been the case if telluric lines were not correctly removed or if there was persistance in the detectors) we decided to test our method on synthetic spectra. As this is work in progress, at this point in time, the parameter ranges used for generating our simulations do not exactly match those expected from observations. As such, what we present in this extended abstract is an assessment of the validity of the method.

4.1 Dataset

We created a synthetic dataset of spectra (we give an example full synthetic spectra in Appendix A). Our training dataset contained N=25000 pairs of spectra each composed of approximately 7000 wavelength bins covering the same wavelength range as the spectra in the APOGEE survey. Both spectra in pairs have a shared chemical composition but differing non-chemical parameters (effective temperature and gravity). Our validation dataset consisted of N=12500 pairs of stars. We refer the reader to the appendix B for more detail on the procedure used for creating the synthetic spectra.

4.2 Network Architecture

Our architecture consists of a conditional encoder and a conditional decoder. The conditional encoder denoted $E(X, U)$ takes as inputs a spectra x and its associated vector of non-chemical factors of variation u. The output of the conditonal encoder z is then combined with the non-chemical vector and feed to the conditional decoder which attempts to reconstruct the input spectra x.

We enforce that the neural network learns a disentangled representation through a fader like neural network architecture. In this approach, as presented in [13], we discretize the non-chemical parameter spaces, yielding a one-hot-encoding u_n. An adversary network denoted $A(Z)$ takes as input the learnt latents and is trained to predict the discretized non-chemical parameter values through a cross-entropy loss. The autoencoder is then jointly trained through an adverserial loss term designed to encourage the learned latent to be uninformative for the adversary. In this scenario the training objective for the

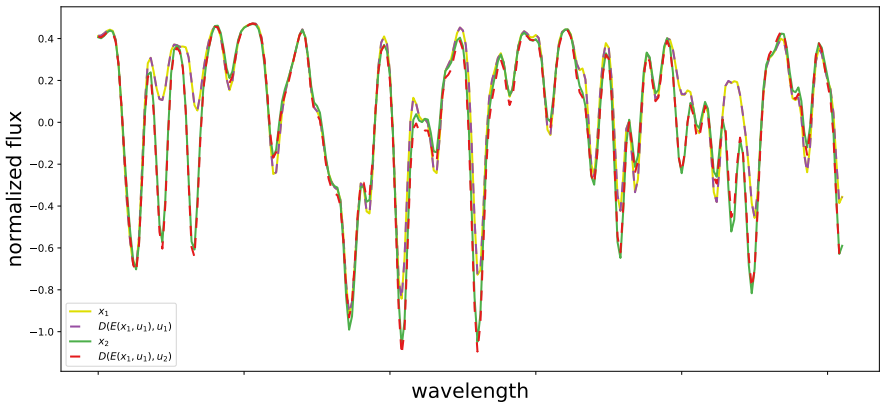

Figure 1: Visualization of the spectra learned by the neural network as compared to real spectra with and without swapping. For clarity, we are only showing the first 256 wavelength bins.

encoder-decoder is

$$L_{AE}(\theta_E, \theta_D \mid \theta_A) = E_{(x,u)\sim p(x,u)}[\|D(E(x,u),U) - x\|_2^2] - \lambda_1 E_{(x,u,u_n)\sim p(x,u)}[-u_n log(A(E(x,u)))] \tag{1}$$

while the training objective for the adversary is simply the cross entropy loss, ie

$$L_A = E_{(x,u,u_n)\sim p(x,u)}[-u_n log(A(E(x,u)))] \tag{2}$$

4.3 Results

We present in Table 1 quantitative evaluations of our method. We evaluated the reconstruction loss with and without swapping the latents of spectra pairs. As both stars in a pair share an identical chemical composition, under perfect disentanglement, swapping the latent should have no effect on the reconstruction loss. The reconstruction loss after swapping latents thus offers a metric for quantifying the quality of the disentanglement. We see, both visually from Figure 1 and quantitatively from Table 1 that the reconstruction loss remains very small even after swapping. Interestingly, all existing metrics for supervised neural network disentanglement in the literature require visual inspection or rely on evaluating how easy it is for a neural network to estimate the disentangled parameters. We believe that using a secondary network to evaluate the disentanglement is to be avoided as the training objective and the evaluation metric become identical. In addition, the retrieved values may be affected by the network architecture and its associated inaccuracies. We thus believe that our task and dataset offer an easier and more natural method for evaluating disentanglement.

We also conducted experiment towards evaluating our method's ability at identifying chemically identical stars. To do this, we obtained latents for every spectra in the validation dataset using our encoder and identified the most similar latents in the validation dataset according to an L1 distance metric. We report here the fraction of stars for which the true stellar twins latent appeared as the most similar latent (0% doppelgangers), amongst the 25 most similar latents (0.1% doppelgangers) and amongst the 250 most similar latents (1% doppelgangers) in our validation dataset of 25000 stars.

5 Conclusion

We present a method for identifying chemically identical stars without ever needing to explicitly model their chemistry. We have demonstrated and validated our method on a synthetic dataset of stellar spectra. We anticipate that our method may find applications in other scientific domains where it is only possible to precisely estimate a portion of the factors of variations required for explaining a set of observations.

Experimental evaluation

Task	Error	Error Baseline
Reconstruction (MSE)	2.2E-6	**4.1E-7**
Reconstruction with swapped latents (MSE)	**3.7E-4**	9.7E-3
Fraction with 0% doppelgangers	**60.7**%	8.1%
Fraction with less then 0.1% doppelgangers	**82.3**%	14.4%
Fraction with less then 1% doppelgangers	**93.7**%	24.8%

Table 1: Quantitative results evaluated using a validation dataset. The baseline is a conditionoal autoencoder **without** disentanglement but otherwise identical. Best performant model is bolded.

References

[1] Angus Beane, Melissa K. Ness, and Megan Bedell. Actions Are Weak Stellar Age Indicators in the Milky Way Disk. *The Astrophysical Journal*, 867(1):31, Nov 2018.

[2] Sergi Blanco-Cuaresma and Didier Fraix-Burnet. A phylogenetic approach to chemical tagging. *Astronomy & Astrophysics*, 618:A65, oct 2018.

[3] Jo Bovy. THE CHEMICAL HOMOGENEITY OF OPEN CLUSTERS. *The Astrophysical Journal*, 817(1):49, jan 2016.

[4] Jiawei Chen, Janusz Konrad, and Prakash Ishwar. A Cyclically-Trained Adversarial Network for Invariant Representation Learning. *arXiv e-prints*, page arXiv:1906.09313, Jun 2019.

[5] Xi Chen, Yan Duan, Rein Houthooft, John Schulman, Ilya Sutskever, and Pieter Abbeel. Infogan: Interpretable representation learning by information maximizing generative adversarial nets. In *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain*, pages 2172–2180, 2016.

[6] G. M. De Silva, K. C. Freeman, J. Bland-Hawthorn, S. Martell, E. Wylie de Boer, M. Asplund, S. Keller, S. Sharma, D. B. Zucker, T. Zwitter, B. Anguiano, C. Bacigalupo, D. Bayliss, M. A. Beavis, M. Bergemann, S. Campbell, R. Cannon, D. Carollo, L. Casagrande, A. R. Casey, G. Da Costa, V. D'Orazi, A. Dotter, L. Duong, A. Heger, M. J. Ireland, P. R. Kafle, J. Kos, J. Lattanzio, G. F. Lewis, J. Lin, K. Lind, U. Munari, D. M. Nataf, S. O'Toole, Q. Parker, W. Reid, K. J. Schlesinger, A. Sheinis, J. D. Simpson, D. Stello, Y. S. Ting, G. Traven, F. Watson, R. Wittenmyer, D. Yong, and M. Žerjal. The GALAH survey: scientific motivation. , 449(3):2604–2617, May 2015.

[7] Neige Frankel, Hans-Walter Rix, Yuan-Sen Ting, Melissa Ness, and David W. Hogg. Measuring Radial Orbit Migration in the Galactic Disk. *The Astrophysical Journal*, 865(2):96, Oct 2018.

[8] Irina Higgins, Loïc Matthey, Arka Pal, Christopher Burgess, Xavier Glorot, Matthew Botvinick, Shakir Mohamed, and Alexander Lerchner. beta-vae: Learning basic visual concepts with a constrained variational framework. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*, 2017.

[9] David W. Hogg, Andrew R. Casey, Melissa Ness, Hans-Walter Rix, Daniel Foreman-Mackey, Sten Hasselquist, Anna Y. Q. Ho, Jon A. Holtzman, Steven R. Majewski, Sarah L. Martell, Szabolcs Mészáros, David L. Nidever, and Matthew Shetrone. CHEMICAL TAGGING CAN WORK: IDENTIFICATION OF STELLAR PHASE-SPACE STRUCTURES PURELY BY CHEMICAL-ABUNDANCE SIMILARITY. *The Astrophysical Journal*, 833(2):262, dec 2016.

[10] Ananya Harsh Jha, Saket Anand, Maneesh Singh, and V. S. R. Veeravasarapu. Disentangling factors of variation with cycle-consistent variational auto-encoders. In *Computer Vision - ECCV 2018 - 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part III*, pages 829–845, 2018.

[11] Paula Jofré, Payel Das, Jaume Bertranpetit, and Robert Foley. Cosmic phylogeny: reconstructing the chemical history of the solar neighbourhood with an evolutionary tree. *Monthly Notices of the Royal Astronomical Society*, 467(1):1140–1153, may 2017.

[12] Harshil Kamdar, Charlie Conroy, Yuan-Sen Ting, Ana Bonaca, Benjamin Johnson, and Phillip Cargile. A Dynamical Model for Clustered Star Formation in the Galactic Disk. *arXiv e-prints*, page arXiv:1902.10719, Feb 2019.

[13] Guillaume Lample, Neil Zeghidour, Nicolas Usunier, Antoine Bordes, Ludovic Denoyer, and Marc'Aurelio Ranzato. Fader networks: Manipulating images by sliding attributes. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA*, pages 5969–5978, 2017.

[14] Henry W Leung and Jo Bovy. Deep learning of multi-element abundances from high-resolution spectroscopic data. *Monthly Notices of the Royal Astronomical Society*, 483(3):3255–3277, nov 2018.

[15] Ting S. Li, Allyson A. Sheffield, Kathryn V. Johnston, Jennifer L. Marshall, Steven R. Majewski, Adrian M. Price-Whelan, Guillermo J. Damke, Rachael L. Beaton, Edouard J. Bernard, Whitney Richardson, Sanjib Sharma, and Branimir Sesar. Exploring Halo Substructure with Giant Stars. XV. Discovery of a Connection between the Monoceros Ring and the Triangulum-Andromeda Overdensity? *The Astrophysical Journal*, 844(1):74, Jul 2017.

[16] Steven R. Majewski, Ricardo P. Schiavon, Peter M. Frinchaboy, Carlos Allende Prieto, Robert Barkhouser, Dmitry Bizyaev, Basil Blank, Sophia Brunner, Adam Burton, Ricardo Carrera, S. Drew Chojnowski, Kátia Cunha, Courtney Epstein, Greg Fitzgerald, Ana E. García Pérez, Fred R. Hearty, Chuck Henderson, Jon A. Holtzman, Jennifer A. Johnson, Charles R. Lam, James E. Lawler, Paul Maseman, Szabolcs Mészáros, Matthew Nelson, Duy Coung Nguyen, David L. Nidever, Marc Pinsonneault, Matthew Shetrone, Stephen Smee, Verne V. Smith, Todd Stolberg, Michael F. Skrutskie, Eric Walker, John C. Wilson, Gail Zasowski, Friedrich Anders, Sarbani Basu, Stephane Beland, Michael R. Blanton, Jo Bovy, Joel R. Brownstein, Joleen Carlberg, William Chaplin, Cristina Chiappini, Daniel J. Eisenstein, Yvonne Elsworth, Diane Feuillet, Scott W. Fleming, Jessica Galbraith-Frew, Rafael A. García, D. Aníbal García-Hernández, Bruce A. Gillespie, Léo Girardi, James E. Gunn, Sten Hasselquist, Michael R. Hayden, Saskia Hekker, Inese Ivans, Karen Kinemuchi, Mark Klaene, Suvrath Mahadevan, Savita Mathur, Benoît Mosser, Demitri Muna, Jeffrey A. Munn, Robert C. Nichol, Robert W. O'Connell, John K. Parejko, A. C. Robin, Helio Rocha-Pinto, Matthias Schultheis, Aldo M. Serenelli, Neville Shane, Victor Silva Aguirre, Jennifer S. Sobeck, Benjamin Thompson, Nicholas W. Troup, David H. Weinberg, and Olga Zamora. The Apache Point Observatory Galactic Evolution Experiment (APOGEE). , 154(3):94, Sep 2017.

[17] Sz. Mészáros, C. Allende Prieto, B. Edvardsson, F. Castelli, A. E. García Pérez, B. Gustafsson, S. R. Majewski, B. Plez, R. Schiavon, M. Shetrone, and A. de Vicente. New ATLAS9 and MARCS Model Atmosphere Grids for the Apache Point Observatory Galactic Evolution Experiment (APOGEE). *The Astronomical Journal*, 144(4):120, Oct 2012.

[18] M. Ness, David W. Hogg, H.-W. Rix, Anna. Y. Q. Ho, and G. Zasowski. THE CANNON : A DATA-DRIVEN APPROACH TO STELLAR LABEL DETERMINATION. *The Astrophysical Journal*, 808(1):16, jul 2015.

[19] M. Ness, H-W. Rix, David W. Hogg, A. R. Casey, J. Holtzman, M. Fouesneau, G. Zasowski, D. Geisler, M. Shetrone, D. Minniti, Peter M. Frinchaboy, and Alexandre Roman-Lopes. Galactic Doppelgängers: The Chemical Similarity Among Field Stars and Among Stars with a Common Birth Origin. *The Astrophysical Journal*, 853(2):198, feb 2018.

[20] B. Plez. Turbospectrum: Code for spectral synthesis, May 2012.

[21] Natalie Price-Jones and Jo Bovy. The dimensionality of stellar chemical space using spectra from the Apache Point Observatory Galactic Evolution Experiment. *Monthly Notices of the Royal Astronomical Society*, 475(1):1410–1425, 12 2017.

[22] Natalie Price-Jones and Jo Bovy. Blind chemical tagging with DBSCAN: prospects for spectroscopic surveys. *Monthly Notices of the Royal Astronomical Society*, 487(1):871–886, jul 2019.

[23] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, Boston, MA, USA, June 7-12, 2015*, pages 815–823, 2015.

[24] Yuan-Sen Ting, Charlie Conroy, Hans-Walter Rix, and Phillip Cargile. The Payne : Self-consistent ab initio Fitting of Stellar Spectra. *The Astrophysical Journal*, 879(2):69, jul 2019.

A Spectra example

An example spectra

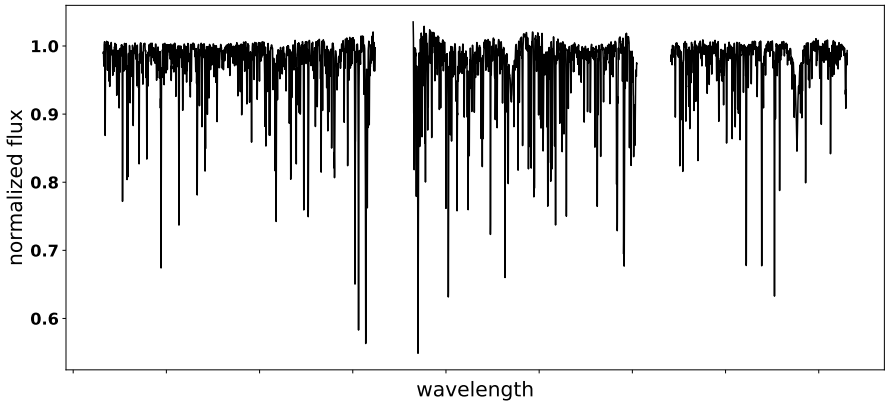

Figure 2: Example continuum-normalized spectra. The wavelength range reflect those found in spectra measured by the APOGEE survey. The spectra consists of many different spectral lines.

B Dataset

We created synthetic spectra using the APOGEEE python package introduced in [3] . This package is a wrapper that makes use of grids of ATLAS9 atmospheres [17] and of the spectral synthesis code Turbospectrum [20] to generate mock spectra with the same wavelength coverage as the APOGEE survey.

Our spectra were generated and continuum-normalized using the true continuum. Each pair of spectra were created independently. To create each pair, we sampled a shared chemical composition and distinct T and logG.