
Scalable Extreme Deconvolution

Anonymous Author(s)

Affiliation

Address

email

Abstract

1 The Extreme Deconvolution method is a mixture-model based approach originally
2 intended to perform density estimation on noisy astronomy datasets with known
3 noise covariances for each datapoint. Newer much larger datasets are available to
4 which the existing method for fitting the model cannot scale to. We propose two
5 extensions to the extreme deconvolution method based on an online variation of the
6 EM algorithm, and direct optimisation of the log-likelihood via SGD, both which
7 can make use of GPU computation. We demonstrate that these methods provide
8 faster fitting of the model, whilst being able to scale to much larger datasets.

9 1 Introduction

10 Many scientific applications require an estimate of some density $p(\mathbf{x})$ to be fitted using a set of
11 samples \mathbf{x}_i . In many cases the dataset is noisy, but we wish to fit our density estimate to the true
12 underlying noiseless data. Extreme Deconvolution is a method that allows Gaussian Mixture Models
13 (GMMs) to be fitted to data where we know the covariance of the Gaussian noise added to each
14 sample [2]. The method was originally developed to perform density estimation on the dataset
15 produced by the Hipparcos satellite [8]. The Hipparcos catalogue consists of astrometric solutions
16 (positions and velocities on the sky) and photometry (light intensity) for 118,218 stars, with associated
17 noise covariances provided for each entry.

18 The successor to the Hipparcos mission, Gaia, aims to produce an even larger catalogue, with
19 entries for an estimated 1 billion astronomical objects [5]. The existing method of fitting extreme
20 deconvolution models cannot cope with such a dataset, as it requires all of the data to fit in memory,
21 and must make a full pass over the dataset before it can update parameters. In this work, we propose
22 two new methods for fitting the extreme deconvolution model that can scale to such datasets, based
23 on an online variation of the Expectation-Maximisation (EM) algorithm, and stochastic gradient
24 descent (SGD) on the log-likelihood. We develop implementations that can leverage GPU-based
25 computation, and show that they provide comparable density estimates to the existing method, whilst
26 being much faster to train.

27 2 Background

28 The aim of extreme deconvolution is to perform density estimation on a noisy d -dimensional dataset
29 $\{\mathbf{x}_i\}_{i=0}^N$, where \mathbf{x}_i was generated by adding zero-mean Gaussian noise ϵ_i with known per-datapoint
30 covariance S_i to a true value \mathbf{v}_i ,

$$\mathbf{x}_i = \mathbf{v}_i + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(\mathbf{0}, S_i). \quad (1)$$

31 We assume that \mathbf{v}_i can be modelled by a mixture of Gaussians with K components,

$$p(\mathbf{v}_i | \theta) = \sum_j^K \alpha_j \mathcal{N}(\mathbf{v} | \mathbf{m}_j, V_j), \quad \theta = \{\alpha_j, \mathbf{m}_j, V_j\}_{j=1}^K \quad (2)$$

As both the noise model is Gaussian and the density estimation model is a mixture of Gaussians, the likelihood of \mathbf{x}_i is also a Gaussian mixture, and the total log-likelihood of the dataset is

$$\mathcal{L}(\theta) = \sum_i^N \log \sum_j^K \alpha_j \mathcal{N}(\mathbf{v}_i | \mathbf{m}_j, T_{ij}), \quad T_{ij} = V_j + S_i. \quad (3)$$

The original model also includes a method for handling the case where \mathbf{x}_i has been transformed from \mathbf{v}_i via a per-datapoint projection matrix R_i , which we omit here for simplicity. Missing data can be handled either by making R_i rank-deficient, or by setting elements of the covariance matrix S_i to very large values.

3 Methods

3.1 Online Expectation-Maximisation

The original method for fitting extreme deconvolution models used a modification of the Expectation-Maximisation (EM) algorithm for mixture models [6]. In the E-step, the expected sufficient statistics for each datapoint that need to be computed are $s_{ij} = \{q_{ij}, q_{ij}\mathbf{v}_i, \text{ and } q_{ij}\mathbf{v}_i\mathbf{v}_i^T\}$. The M-step comprises normalisation of the sums of expected sufficient statistics to get the mixture component parameters.

Cappé and Moulines [4] proposed an online variation of EM for latent data models. Their method involves computing a stochastic approximation to the sums of expected sufficient statistics. At each iteration t , the sums of expected sufficient statistics $s_t = \sum_i s_{ij}$ are computed over a minibatch using the current estimate of the parameters, and the stochastic estimate \hat{s}_t updated as

$$\hat{s}_t = (1 - \lambda)\hat{s}_{t-1} + \lambda s_t \quad (4)$$

where λ is a sufficiently small step-size. λ should theoretically decrease at each iteration according to the Robbins-Monro conditions, but in practice we found that using a constant step-size was sufficient. After each update, the parameters of the model are updated by normalising \hat{s}_t appropriately.

TODO: More complete explanation if space.

A direct implementation suggests that we should update V_j as

$$V_j = E(q_{ij}\mathbf{v}_i j \mathbf{v}_i j^T) - \mathbf{m}_j \mathbf{m}_j^T \quad (5)$$

This is usually acceptable when using double-precision floats, but using GPU based computation restricts us to single precision. Performing the update of V_j like this with single-precision floats will result in catastrophic cancellation if the means are sufficiently large compared to the variance. Instead, we compute smoothed versions of the sum of squared deviation products which avoids this issue.

(6)

TODO: Better justification of why this works, or alternate method.

3.2 Stochastic Gradient Descent

An alternative to EM-based methods is to optimise the log-likelihood directly. The log-likelihood has some constraints on its parameters, namely that the weight α_j sum to 1 and that the covariances V_j are positive definite. Directly fitting it with standard gradient-based optimisers requires a transformation of the parameters to remove the constraints. The mixture weights α_j can be parameterised by taking the softmax of an unconstrained vector \mathbf{z} ,

$$\alpha_j = \frac{e^{z_j}}{\sum_k^K e^{z_k}}. \quad (7)$$

We parametrise the covariances V_j via the Cholesky decomposition,

$$V_j = L_j L_j^T \quad (8)$$

where L_j is a lower-triangular matrix. We also constrain the diagonal elements of L_j to be positive by taking the log of them. Note that for a standard GMM we could bypass forming the covariance

Table 1: Total validation and test log-likelihoods along with average training time for the Gaia data subset. Average over 5 runs with standard deviation.

Method	Validation	Test	Training Time (s)
Existing EM	-	-	
Online EM	-	-	
SGD	$-5.468 \times 10^6 \pm 2.46 \times 10^3$	-	1469 ± 11

altogether and evaluate the log PDF of the Normal distribution directly with the Cholesky factor. For the extreme deconvolution model, this cannot be done, as we need to form the covariance T_{ij} for every datapoint. Having removed the constraints, we can then take the gradient of the log-likelihood with respect to the parameters and optimise it using a standard minibatch-based optimiser. Pre-multiplying the gradient with the inverse Fisher information matrix may be necessary to keep the covariances stable during optimisation with standard stochastic gradient descent. In practice we found that using the Adam optimisation algorithm eliminates the need to do this [7].

4 Experiments

We implemented both of our methods in PyTorch. The use of PyTorch allows us to take advantage of its automatic differentiation framework and in-built optimisers for the SGD method, as well as allowing us to write code targeting both CPU and GPU-based computation.

To evaluate our methods, we used a subset of rows from the Gaia DR2 source table [3]. We selected the 5 primary astrometric features, along with the BP-RP colour and mean flux in the G-band. In total there were approximately 2 million rows. Where data were missing, we set the field to zero and the noise variance to a large value. This is only a small fraction of the full dataset size, but this allows us to fit the training data into memory, a requirement for use with the original implementation of extreme deconvolution. We used the same number of components $K = 128$ for each model, following [1]. In practice we would want to select a value of K by cross-validation.

Initialisation of all models was performed by first running a minibatch version of the K-means clustering algorithm on the dataset to get estimates of the weights and means [9]. Initial covariances were set to the identity matrix. Early stopping was done based on a held-out validation set comprising 10% of the rows. Final model performance was evaluated on another held-out test set comprising another 10% of the rows. Table 1 reports the validation and test log-likelihoods for each method.

TODO: Add plot of log-likelihood results against dataset size.

5 Discussion

Our results have shown that both of our proposed methods perform comparably to the existing method of fitting extreme deconvolution models, whilst being able to work with datasets that cannot fit in memory. The results also show that using GPU-based computation speeds up training considerably.

Further improvements to our approaches are possible. The original paper presents a method of getting out of local maxima by splitting and merging clusters, with the split-merge criteria evaluated on the whole dataset. It should be possible to replace the criteria with stochastic estimates, which would permit them to be used with both the SGD and online EM methods. Our approaches also add more free parameters to be selected, including learning rate and batch size. This adds scope for hyperparameter optimisation to improve the log-likelihood, potentially making use of Bayesian optimisation.

The question remains of which of our two methods should be used for any particular problem? As the results show, the SGD method is faster than the online EM method, at a cost of having a slightly worse log-likelihood. The online EM method performed better, but during training would occasionally terminate early as a result of an update forming an invalid covariance matrix, necessitating a restart. (This issue also exists with the implementation of the existing method). The answer will therefore depend on exactly what dataset we wish to perform density estimation on, and how accurate an estimation we require. In practice, we would suggest using the online EM if the size of the dataset

was small enough to permit several restarts, whilst an extremely large dataset would probably be better fitted with the SGD method.

TODO: Add an example plot of online-EM going wrong if space?

References

- [1] L. Anderson, D. W. Hogg, B. Leistedt, A. M. Price-Whelan, and J. Bovy. Improving Gaia Parallax Precision with a Data-driven Model of Stars. 156(4):145. ISSN 1538-3881. doi: 10.3847/1538-3881/aad7bf. URL <https://doi.org/10.3847/1538-3881/aad7bf>.
- [2] J. Bovy, D. W. Hogg, and S. T. Roweis. Extreme deconvolution: Inferring complete distribution functions from noisy, heterogeneous and incomplete observations. 5:1657–1677. ISSN 1932-6157, 1941-7330. doi: 10.1214/10-AOAS439. URL <https://projecteuclid.org/euclid.aoas/1310562737>.
- [3] A. G. A. Brown, A. Vallenari, T. Prusti, J. H. J. de Bruijne, C. Babusiaux, C. a. L. Bailer-Jones, M. Biermann, D. W. Evans, L. Eyer, F. Jansen, C. Jordi, S. A. Klioner, U. Lammers, L. Lindgren, X. Luri, F. Mignard, C. Panem, D. Pourbaix, S. Randich, P. Sartoretti, H. I. Siddiqui, C. Soubiran, F. van Leeuwen, N. A. Walton, F. Arenou, U. Bastian, M. Cropper, R. Drimmel, D. Katz, M. G. Lattanzi, J. Bakker, C. Cacciari, J. Castañeda, L. Chaoul, N. Cheek, F. D. Angeli, C. Fabricius, R. Guerra, B. Holl, E. Masana, R. Messineo, N. Mowlavi, K. Nienartowicz, P. Panuzzo, J. Portell, M. Riello, G. M. Seabroke, P. Tanga, F. Thévenin, G. Gracia-Abril, G. Comoretto, M. Garcia-Reinaldos, D. Teyssier, M. Altmann, R. Andrae, M. Audard, I. Bellas-Velidis, K. Benson, J. Berthier, R. Blomme, P. Burgess, G. Busso, B. Carry, A. Cellino, G. Clementini, M. Clotet, O. Creevey, M. Davidson, J. D. Ridder, L. Delchambre, A. Dell’Oro, C. Ducourant, J. Fernández-Hernández, M. Fouesneau, Y. Frémat, L. Galluccio, M. García-Torres, J. González-Núñez, J. J. González-Vidal, E. Gosset, L. P. Guy, J.-L. Halbwachs, N. C. Hambly, D. L. Harrison, J. Hernández, D. Hestroffer, S. T. Hodgkin, A. Hutton, G. Jasiewicz, A. Jean-Antoine-Piccolo, S. Jordan, A. J. Korn, A. Krone-Martins, A. C. Lanzafame, T. Lebzelter, W. Löffler, M. Manteiga, P. M. Marrese, J. M. Martín-Fleitas, A. Moitinho, A. Mora, K. Muinonen, J. Osinde, E. Pancino, T. Pauwels, J.-M. Petit, A. Recio-Blanco, P. J. Richards, L. Rimoldini, A. C. Robin, L. M. Sarro, C. Siopis, M. Smith, A. Sozzetti, M. Süveges, J. Torra, W. van Reeve, U. Abbas, A. A. Aramburu, S. Accart, C. Aerts, G. Altavilla, M. A. Álvarez, R. Alvarez, J. Alves, R. I. Anderson, A. H. Andrei, E. A. Varela, E. Antiche, T. Antoja, B. Arcay, T. L. Astraatmadja, N. Bach, S. G. Baker, L. Balaguer-Núñez, P. Balm, C. Barache, C. Barata, D. Barbato, F. Barblan, P. S. Barklem, D. Barrado, M. Barros, M. A. Barstow, S. B. Muñoz, J.-L. Bassilana, U. Becciani, M. Bellazzini, A. Berihuete, S. Bertone, L. Bianchi, O. Bienaymé, S. Blanco-Cuaresma, T. Boch, C. Boeche, A. Bombrun, R. Borrachero, D. Bossini, S. Bouquillon, G. Bourda, A. Bragaglia, L. Bramante, M. A. Breddels, A. Bressan, N. Brouillet, T. Brüsemeister, E. Brugaletta, B. Bucciarelli, A. Burlacu, D. Busonero, A. G. Butkevich, R. Buzzi, E. Caffau, R. Cancelliere, G. Cannizzaro, T. Cantat-Gaudin, R. Carballo, T. Carlucci, J. M. Carrasco, L. Casamiquela, M. Castellani, A. Castro-Ginard, P. Charlot, L. Chemin, A. Chiavassa, G. Cocozza, G. Costigan, S. Cowell, F. Crifo, M. Crosta, C. Crowley, J. Cuypers†, C. Dafonte, Y. Damerdj, A. Dapergolas, P. David, M. David, P. de Laverny, F. D. Luise, R. D. March, D. de Martino, R. de Souza, A. de Torres, J. Debosscher, E. del Pozo, M. Delbo, A. Delgado, H. E. Delgado, P. D. Matteo, S. Diakite, C. Diener, E. Distefano, C. Dolding, P. Drazinos, J. Durán, B. Edvardsson, H. Enke, K. Eriksson, P. Esquej, G. E. Bontemps, C. Fabre, M. Fabrizio, S. Faigler, A. J. Falcão, M. F. Casas, L. Federici, G. Fedorets, P. Fernique, F. Figueras, F. Filippi, K. Findeisen, A. Fonti, E. Fraile, M. Fraser, B. Frézouls, M. Gai, S. Galleti, D. Garabato, F. García-Sedano, A. Garofalo, N. Garralda, A. Gavel, P. Gavras, J. Gerssen, R. Geyer, P. Giacobbe, G. Gilmore, S. Girona, G. Giuffrida, F. Glass, M. Gomes, M. Granvik, A. Gueguen, A. Guerrier, J. Guiraud, R. Gutiérrez-Sánchez, R. Haigron, D. Hatzidimitriou, M. Hauser, M. Haywood, U. Heiter, A. Helmi, J. Heu, T. Hilger, D. Hobbs, W. Hofmann, G. Holland, H. E. Huckle, A. Hypki, V. Icardi, K. Janßen, G. J. de Fombelle, P. G. Jonker, L. Juhász, F. Julbe, A. Karampelas, A. Kewley, J. Klar, A. Kochoska, R. Kohley, K. Kolenberg, M. Kontizas, E. Kontizas, S. E. Koposov, G. Kordopatis, Z. Kostrzewa-Rutkowska, P. Koubsky, S. Lambert, A. F. Lanza, Y. Lasne, J.-B. Lavigne, Y. L. Fustec, C. L. Poncin-Lafitte, Y. Lebreton, S. Leccia, N. Leclerc, I. Lecoœur-Taibi, H. Lenhardt, F. Leroux, S. Liao, E. Licata, H. E. P. Lindstrøm, T. A. Lister, E. Livanou, A. Lobel, M. López, S. Managau, R. G. Mann, G. Mantetlet, O. Marchal, J. M. Marchant, M. Marconi, S. Marinoni, G. Marschallkó, D. J. Marshall, M. Martino, G. Marton, N. Mary, D. Massari, G. Matijević, T. Mazeh, P. J. McMillan, S. Messina,

- 167 D. Michalik, N. R. Millar, D. Molina, R. Molinaro, L. Molnár, P. Montegriffo, R. Mor, R. Mor-
168 bidelli, T. Morel, D. Morris, A. F. Mulone, T. Muraveva, I. Musella, G. Nelemans, L. Nicastro,
169 L. Noval, W. O'Mullane, C. Ordénovic, D. Ordóñez-Blanco, P. Osborne, C. Pagani, I. Pagano,
170 F. Pailler, H. Palacin, L. Palaversa, A. Panahi, M. Pawlak, A. M. Piersimoni, F.-X. Pineau,
171 E. Plachy, G. Plum, E. Poggio, E. Poujoulet, A. Prša, L. Pulone, E. Racero, S. Ragaini, N. Ram-
172 baux, M. Ramos-Lerate, S. Regibo, C. Reylé, F. Riclet, V. Ripepi, A. Riva, A. Rivard, G. Rixon,
173 T. Roegiers, M. Roelens, M. Romero-Gómez, N. Rowell, F. Royer, L. Ruiz-Dern, G. Sadowski,
174 T. S. Sellés, J. Sahlmann, J. Salgado, E. Salguero, N. Sanna, T. Santana-Ros, M. Sarasso, H. Savi-
175 etto, M. Schultheis, E. Sciacca, M. Segol, J. C. Segovia, D. Ségransan, I.-C. Shih, L. Siltala,
176 A. F. Silva, R. L. Smart, K. W. Smith, E. Solano, F. Solitro, R. Sordo, S. S. Nieto, J. Souchay,
177 A. Spagna, F. Spoto, U. Stampa, I. A. Steele, H. Steidelmüller, C. A. Stephenson, H. Stoev, F. F.
178 Suess, J. Surdej, L. Szabados, E. Szegedi-Elek, D. Tapiador, F. Taris, G. Tauran, M. B. Taylor,
179 R. Teixeira, D. Terrett, P. Teyssandier, W. Thuillot, A. Titarenko, F. T. Clotet, C. Turon, A. Ulla,
180 E. Utrilla, S. Uzzi, M. Vaillant, G. Valentini, V. Valette, A. van Elteren, E. V. Hemelryck, M. van
181 Leeuwen, M. Vaschetto, A. Vecchiato, J. Veljanoski, Y. Viala, D. Vicente, S. Vogt, C. von Essen,
182 H. Voss, V. Votruba, S. Voutsinas, G. Walmsley, M. Weiler, O. Wertz, T. Wevers, Wyrzykowski,
183 A. Yoldas, M. Žerjal, H. Ziaeeppour, J. Zorec, S. Zschocke, S. Zucker, C. Zurbach, T. Zwitter,
184 and Gaia Collaboration. Gaia Data Release 2 - Summary of the contents and survey properties.
185 616:A1. ISSN 0004-6361, 1432-0746. doi: 10.1051/0004-6361/201833051. URL <https://www.aanda.org/articles/aa/abs/2018/08/aa33051-18/aa33051-18.html>.
186
- 187 [4] O. Cappé and E. Moulines. On-line expectation-maximization algorithm for latent data models.
188 71(3):593–613. ISSN 1467-9868. doi: 10.1111/j.1467-9868.2009.00698.x. URL <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9868.2009.00698.x>.
189
- 190 [5] G. Collaboration, T. Prusti, J. H. J. de Bruijne, A. G. A. Brown, A. Vallenari, C. Babusiaux,
191 C. a. L. Bailer-Jones, U. Bastian, M. Biermann, D. W. Evans, L. Eyer, F. Jansen, C. Jordi, S. A.
192 Klioner, U. Lammers, L. Lindegren, X. Luri, F. Mignard, D. J. Milligan, C. Panem, V. Poinsignon,
193 D. Pourbaix, S. Randich, G. Sarri, P. Sartoretti, H. I. Siddiqui, C. Soubiran, V. Valette, F. van
194 Leeuwen, N. A. Walton, C. Aerts, F. Arenou, M. Cropper, R. Drimmel, E. Høg, D. Katz, M. G.
195 Lattanzi, W. O'Mullane, E. K. Grebel, A. D. Holland, C. Huc, X. Passot, L. Bramante, C. Cac-
196 ciari, J. Castañeda, L. Chaoul, N. Cheek, F. De Angeli, C. Fabricius, R. Guerra, J. Hernández,
197 A. Jean-Antoine-Piccolo, E. Masana, R. Messineo, N. Mowlavi, K. Nienartowicz, D. Ordóñez-
198 Blanco, P. Panuzzo, J. Portell, P. J. Richards, M. Riello, G. M. Seabroke, P. Tanga, F. Thévenin,
199 J. Torra, S. G. Els, G. Gracia-Abril, G. Comoretto, M. Garcia-Reinaldos, T. Lock, E. Mercier,
200 M. Altmann, R. Andrae, T. L. Astraatmadja, I. Bellas-Velidis, K. Benson, J. Berthier, R. Blomme,
201 G. Busso, B. Carry, A. Cellino, G. Clementini, S. Cowell, O. Creevey, J. Cuypers, M. Davidson,
202 J. De Ridder, A. de Torres, L. Delchambre, A. Dell'Oro, C. Ducourant, Y. Frémat, M. García-
203 Torres, E. Gosset, J.-L. Halbwachs, N. C. Hambly, D. L. Harrison, M. Hauser, D. Hestroffer,
204 S. T. Hodgkin, H. E. Huckle, A. Hutton, G. Jasiewicz, S. Jordan, M. Kontizas, A. J. Korn,
205 A. C. Lanzafame, M. Manteiga, A. Moitinho, K. Muinonen, J. Osinde, E. Pancino, T. Pauwels,
206 J.-M. Petit, A. Recio-Blanco, A. C. Robin, L. M. Sarro, C. Siopis, M. Smith, K. W. Smith,
207 A. Sozzetti, W. Thuillot, W. van Reeve, Y. Viala, U. Abbas, A. Abreu Aramburu, S. Accart,
208 J. J. Aguado, P. M. Allan, W. Allasia, G. Altavilla, M. A. Álvarez, J. Alves, R. I. Anderson,
209 A. H. Andrei, E. Anglada Varela, E. Antiche, T. Antoja, S. Antón, B. Arcay, A. Atzei, L. Ay-
210 ache, N. Bach, S. G. Baker, L. Balaguer-Núñez, C. Barache, C. Barata, A. Barbier, F. Barblan,
211 M. Baroni, D. Barrado y Navascués, M. Barros, M. A. Barstow, U. Becciani, M. Bellazzini,
212 G. Bellei, A. Bello García, V. Belokurov, P. Bendjoya, A. Berihuete, L. Bianchi, O. Bienaymé,
213 F. Billebaud, N. Blagorodnova, S. Blanco-Cuaresma, T. Boch, A. Bombrun, R. Borrachero,
214 S. Bouquillon, G. Bourda, H. Bouy, A. Bragaglia, M. A. Breddels, N. Brouillet, T. Brüsemeister,
215 B. Bucciarelli, F. Budnik, P. Burgess, R. Burgon, A. Burlacu, D. Busonero, R. Buzzi, E. Caf-
216 fau, J. Cambras, H. Campbell, R. Cancelliere, T. Cantat-Gaudin, T. Carlucci, J. M. Carrasco,
217 M. Castellani, P. Charlot, J. Charnas, P. Charvet, F. Chassat, A. Chiavassa, M. Clotet, G. Coccozza,
218 R. S. Collins, P. Collins, G. Costigan, F. Crifo, N. J. G. Cross, M. Crosta, C. Crowley, C. Da-
219 fonte, Y. Damerdj, A. Dapergolas, P. David, M. David, P. De Cat, F. de Felice, P. de Laverny,
220 F. De Luise, R. De March, D. de Martino, R. de Souza, J. Debosscher, E. del Pozo, M. Delbo,
221 A. Delgado, H. E. Delgado, F. di Marco, P. Di Matteo, S. Diakite, E. Distefano, C. Dolding,
222 S. Dos Anjos, P. Drazinos, J. Durán, Y. Dzigian, E. Ecale, B. Edvardsson, H. Enke, M. Erdmann,
223 D. Escolar, M. Espina, N. W. Evans, G. Eynard Bontemps, C. Fabre, M. Fabrizio, S. Faigler, A. J.
224 Falcão, M. Farràs Casas, F. Faye, L. Federici, G. Fedorets, J. Fernández-Hernández, P. Fernique,
225 A. Fienga, F. Figueras, F. Filippi, K. Findeisen, A. Fonti, M. Fouesneau, E. Fraile, M. Fraser,

226 J. Fuchs, R. Furnell, M. Gai, S. Galleti, L. Galluccio, D. Garabato, F. García-Sedano, P. Garé,
 227 A. Garofalo, N. Garralda, P. Gavras, J. Gerssen, R. Geyer, G. Gilmore, S. Girona, G. Giuf-
 228 frida, M. Gomes, A. González-Marcos, J. González-Núñez, J. J. González-Vidal, M. Granvik,
 229 A. Guerrier, P. Guillout, J. Guiraud, A. Gürpide, R. Gutiérrez-Sánchez, L. P. Guy, R. Haigron,
 230 D. Hatzidimitriou, M. Haywood, U. Heiter, A. Helmi, D. Hobbs, W. Hofmann, B. Holl, G. Hol-
 231 land, J. a. S. Hunt, A. Hypki, V. Icardi, M. Irwin, G. Jevardat de Fombelle, P. Jofré, P. G. Jonker,
 232 A. Jorissen, F. Julbe, A. Karampelas, A. Kochoska, R. Kohley, K. Kolenberg, E. Kontizas, S. E.
 233 Koposov, G. Kordopatis, P. Koubsky, A. Kowalczyk, A. Krone-Martins, M. Kudryashova, I. Kull,
 234 R. K. Bachchan, F. Lacoste-Seris, A. F. Lanza, J.-B. Lavigne, C. Le Poncin-Lafitte, Y. Lebre-
 235 ton, T. Lebzelter, S. Leccia, N. Leclerc, I. Lecoeur-Taïbi, V. Lemaitre, H. Lenhardt, F. Leroux,
 236 S. Liao, E. Licata, H. E. P. Lindstrøm, T. A. Lister, E. Livanou, A. Lobel, W. Löffler, M. López,
 237 A. Lopez-Lozano, D. Lorenz, T. Loureiro, I. MacDonald, T. Magalhães Fernandes, S. Managau,
 238 R. G. Mann, G. Mantelet, O. Marchal, J. M. Marchant, M. Marconi, J. Marie, S. Marinoni, P. M.
 239 Marrese, G. Marschalkó, D. J. Marshall, J. M. Martín-Fleitas, M. Martino, N. Mary, G. Matije-
 240 T. Mazeh, P. J. McMillan, S. Messina, A. Mestre, D. Michalik, N. R. Millar, B. M. H. Miranda,
 241 D. Molina, R. Molinaro, M. Molinaro, L. Molnár, M. Moniez, P. Montegriffo, D. Monteiro,
 242 R. Mor, A. Mora, R. Morbidelli, T. Morel, S. Morgenthaler, T. Morley, D. Morris, A. F. Mu-
 243 lone, T. Muraveva, I. Musella, J. Narbonne, G. Nelemans, L. Nicastro, L. Noval, C. Ordénovic,
 244 J. Ordieres-Meré, P. Osborne, C. Pagani, I. Pagano, F. Pailler, H. Palacin, L. Palaversa, P. Parsons,
 245 T. Paulsen, M. Pecoraro, R. Pedrosa, H. Pentikäinen, J. Pereira, B. Pichon, A. M. Piersimoni,
 246 F.-X. Pineau, E. Plachy, G. Plum, E. Poujoulet, A. Prša, L. Pulone, S. Ragaini, S. Rago, N. Ram-
 247 baux, M. Ramos-Lerate, P. Ranalli, G. Rauw, A. Read, S. Regibo, F. Renk, C. Reylé, R. A.
 248 Ribeiro, L. Rimoldini, V. Ripepi, A. Riva, G. Rixon, M. Roelens, M. Romero-Gómez, N. Rowell,
 249 F. Royer, A. Rudolph, L. Ruiz-Dern, G. Sadowski, T. Sagristà Sellés, J. Sahlmann, J. Salgado,
 250 E. Salguero, M. Sarasso, H. Savietto, A. Schnorhk, M. Schultheis, E. Sciacca, M. Segol, J. C.
 251 Segovia, D. Segransan, E. Serpell, I.-C. Shih, R. Smareglia, R. L. Smart, C. Smith, E. Solano,
 252 F. Solitro, R. Sordo, S. Soria Nieto, J. Souchay, A. Spagna, F. Spoto, U. Stampa, I. A. Steele,
 253 H. Steidelmüller, C. A. Stephenson, H. Stoev, F. F. Suess, M. Süveges, J. Surdej, L. Szabados,
 254 E. Szegedi-Elek, D. Tapiador, F. Taris, G. Tauran, M. B. Taylor, R. Teixeira, D. Terrett, B. Tingley,
 255 S. C. Trager, C. Turon, A. Ulla, E. Utrilla, G. Valentini, A. van Elteren, E. Van Hemelryck,
 256 M. van Leeuwen, M. Varadi, A. Vecchiato, J. Veljanoski, T. Via, D. Vicente, S. Vogt, H. Voss,
 257 V. Votruba, S. Voutsinas, G. Walmsley, M. Weiler, K. Weingrill, D. Werner, T. Wevers, G. White-
 258 head, Wyrzykowski, A. Yoldas, M. Žerjal, S. Zucker, C. Zurbach, T. Zwitter, A. Alecu, M. Allen,
 259 C. Allende Prieto, A. Amorim, G. Anglada-Escudé, V. Arsenijevic, S. Azaz, P. Balm, M. Beck,
 260 H.-H. Bernstein, L. Bigot, A. Bijaoui, C. Blasco, M. Bonfigli, G. Bono, S. Boudreault, A. Bressan,
 261 S. Brown, P.-M. Brunet, P. Bunclark, R. Buonanno, A. G. Butkevich, C. Carret, C. Carrion,
 262 L. Chemin, F. Chéreau, L. Corcione, E. Darmigny, K. S. de Boer, P. de Teodoro, P. T. de Zeeuw,
 263 C. Delle Luche, C. D. Domingues, P. Dubath, F. Fodor, B. Frézouls, A. Fries, D. Fustes, D. Fyfe,
 264 E. Gallardo, J. Gallegos, D. Gardiol, M. Gebran, A. Gomboc, A. Gómez, E. Grux, A. Gueguen,
 265 A. Heyrovsky, J. Hoar, G. Iannicola, Y. Isasi Parache, A.-M. Janotto, E. Joliet, A. Jonckheere,
 266 R. Keil, D.-W. Kim, P. Klagyivik, J. Klar, J. Knude, O. Kochukhov, I. Kolka, J. Kos, A. Kutka,
 267 V. Lainey, D. LeBouquin, C. Liu, D. Loreggia, V. V. Makarov, M. G. Marseille, C. Martayan,
 268 O. Martinez-Rubi, B. Massart, F. Meynadier, S. Mignot, U. Munari, A.-T. Nguyen, T. Nordlander,
 269 P. Ocvirk, K. S. O’Flaherty, A. Olias Sanz, P. Ortiz, J. Osorio, D. Oszkiewicz, A. Ouzounis,
 270 M. Palmer, P. Park, E. Pasquato, C. Peltzer, J. Peralta, F. Péturaud, T. Pieniluoma, E. Pigozzi,
 271 J. Poels, G. Prat, T. Prod’homme, F. Raison, J. M. Rebordao, D. Riskey, B. Rocca-Volmerange,
 272 S. Rosen, M. I. Ruiz-Fuertes, F. Russo, S. Sembay, I. Serraller Vizcaino, A. Short, A. Siebert,
 273 H. Silva, D. Sinachopoulos, E. Slezak, M. Soffel, D. Sosnowska, V. Straižys, M. ter Linden,
 274 D. Terrell, S. Theil, C. Tiede, L. Troisi, P. Tsalmantza, D. Tur, M. Vaccari, F. Vachier, P. Valles,
 275 W. Van Hamme, L. Veltz, J. Virtanen, J.-M. Wallut, R. Wichmann, M. I. Wilkinson, H. Ziaepour,
 276 and S. Zschocke. The Gaia mission. 595:A1. doi: 10.1051/0004-6361/201629272. URL
 277 <https://ui.adsabs.harvard.edu/abs/2016A%26A...595A...1G/abstract>.

278 [6] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum Likelihood from Incomplete Data via
 279 the EM Algorithm. 39(1):1–38. ISSN 0035-9246. URL [https://www.jstor.org/stable/](https://www.jstor.org/stable/2984875)
 280 2984875.

281 [7] D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization. URL [http://arxiv.](http://arxiv.org/abs/1412.6980)
 282 [org/abs/1412.6980](http://arxiv.org/abs/1412.6980).

- 283 [8] M. A. C. Perryman, L. Lindegren, J. Kovalevsky, E. Høg, U. Bastian, P. L. Bernacca, M. Crézé,
284 F. Donati, M. Grenon, M. Grewing, F. Van Leeuwen, H. Van Der Marel, F. Mignard, C. A. Murray,
285 R. S. Le Poole, H. Schrijver, C. Turon, F. Arenou, M. Froeschlé, and C. S. Petersen. The Hipparcos
286 Catalogue. 323(1):49–52. URL <https://hal.archives-ouvertes.fr/hal-02053811>.
- 287 [9] D. Sculley. Web-scale k-means clustering. In *Proceedings of the 19th International Conference*
288 *on World Wide Web - WWW '10*, page 1177. ACM Press. ISBN 978-1-60558-799-8. doi:
289 10.1145/1772690.1772862. URL [http://portal.acm.org/citation.cfm?doid=1772690.](http://portal.acm.org/citation.cfm?doid=1772690.1772862)
290 1772862.