

A method to deconvolve stellar rotational velocities II

The probability distribution function via Tikhonov regularization

Alejandra Christen¹, Pedro Escarate^{2,3}, Michel Curé⁴, Diego F. Rial⁵, and Julia Cassetti⁶

¹ Instituto de Estadística, Pontificia Universidad Católica de Valparaíso, 2950 Valparaíso, Chile
e-mail: alejandra.christen@pucv.cl

² Centro Avanzado de Ingeniería Eléctrica y Electrónica, Universidad Técnica Federico Santa María, 1680 Valparaíso, Chile

³ Large Binocular Telescope Observatory, Steward Observatory, Tucson, AZ 85546, USA

⁴ Instituto de Física y Astronomía, Universidad de Valparaíso, Chile

⁵ Departamento de Matemáticas, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, 1053 Buenos Aires, Argentina

⁶ Universidad Nacional de General Sarmiento, Buenos Aires, 1613 Buenos Aires, Argentina

Received 7 June 2016 / Accepted 12 September 2016

ABSTRACT

Aims. Knowing the distribution of stellar rotational velocities is essential for understanding stellar evolution. Because we measure the projected rotational speed $v \sin i$, we need to solve an ill-posed problem given by a Fredholm integral of the first kind to recover the “true” rotational velocity distribution.

Methods. After discretization of the Fredholm integral we apply the Tikhonov regularization method to obtain directly the probability distribution function for stellar rotational velocities. We propose a simple and straightforward procedure to determine the Tikhonov parameter. We applied Monte Carlo simulations to prove that the Tikhonov method is a consistent estimator and asymptotically unbiased.

Results. This method is applied to a sample of cluster stars. We obtain confidence intervals using a bootstrap method. Our results are in close agreement with those obtained using the Lucy method for recovering the probability density distribution of rotational velocities. Furthermore, Lucy estimation lies inside our confidence interval.

Conclusions. Tikhonov regularization is a highly robust method that deconvolves the rotational velocity probability density function from a sample of $v \sin i$ data directly without the need for any convergence criteria.

Key words. stars: rotation – methods: statistical – methods: numerical – methods: data analysis – stars: fundamental parameters

1. Introduction

Understanding how stars rotate is essential in describing and modelling many aspects of stellar evolution. From spectroscopy observations we can only get the projected velocity, $v \sin i$, where i is the inclination angle with respect to the line of sight. Furthermore, in order to deconvolve (disentangle or unfold) the rotational velocity distribution function, an assumption on the distribution of rotational axes is required. The standard practice is to assume that the distribution of stellar axes is uniformly (randomly) distributed over the sphere. Using this assumption Chandrasekhar & Münch (1950) studied the integral equation that describes the distribution of “true” (v) and apparent ($v \sin i$) rotational velocities, deriving a formal solution, which is proportional to a derivative of an Abel’s Integral. Chandrasekhar & Münch (1950) method is not usually applied, because the differentiation of the formal solution can lead to misleading results due to intrinsic numerical problems associated to the derivative of the Abel’s integral.

Curé et al. (2014) extended the work of Chandrasekhar & Münch (1950) integrating the formal solution, and obtained the cumulative distribution function (CDF) for the rotational velocities. This CDF is attained in only one step demonstrating the robustness to this method.

While the CDF identifies the distribution of the speed of rotation, it is sometimes useful to have the probability density

function (PDF) for easy handling and to directly appreciate certain properties of the distribution (e.g., the maximum, its symmetry and its variability, etc.). It is also known that the observed values of the projected rotational velocities are provided with measurement error. The aim of this work is to propose a methodology that directly provides the PDF, taking into account the measurement errors and avoiding numerical problems arising from the derivative of the CDF. Regularization methods are techniques widely used to deconvolve inverse problems. Image processing, geophysics and machine learning are some of the areas where they are usually applied (Bouhamidi 2007; Deng et al. 2013; Fomel 2007). Among the regularization methods we find: truncated singular value decomposition (TSVD), selective singular value decomposition (SSVD) and the Tikhonov regularization method (Hansen 2010).

In this article we obtain the estimated probability distribution function directly from the Fredholm integral by means of the Tikhonov regularization method.

After its introduction by Tikhonov (1943) to solve integral equation problems, this method (known as Ridge Regression in statistics) has been developed and extensively used ever since (see, e.g., Tikhonov 1963, 1995; Tikhonov & Arsenin 1977; Eggermont 1993; and Hansen 2010). It allows for an increase in the numerical stability and for errors of measurement to be taken into account.

This article is structured as follows: in Sect. 2 we briefly present the mathematical description of the method and describe a procedure for calculating the Tikhonov factor. In Sect. 3, we perform Monte Carlo simulations to show the robustness of this method. In Sect. 4, a real sample of cluster stars is deconvolved using Tikhonov regularization, confidence intervals are calculated using a bootstrap method, and a comparison is made between our PDF results and both those obtained with the Lucy (1974) method and CDF results from the work of Curé et al. (2014). Our conclusions and plans for future work are presented in the final section.

2. Tikhonov regularization method

Many inverse problems in physics and astronomy are given in terms of the Fredholm integral of the first kind (Lucy 1994; Hansen 2010). Namely:

$$f_Y(y) = \int p(y|x) f_X(x) dx. \quad (1)$$

Here f_Y is a function accessible to observation and f_X is the function of interest. The kernel $p(y|x)$ of this integral is related to the remoteness of the measurement process; in this case, the projection of the distribution of stellar axes.

Chandrasekhar & Münch (1950) were the first to consider the integral equation governing the distribution of “true” and apparent (projected) rotational velocities of stars, $y = x \sin i$, where $x = v$ is the rotational speed and i is the inclination angle with respect to the line of sight. Assuming a uniform distribution of stellar axes over the sphere (see Curé et al. 2014, for details), this integral equation (Eq. (1)) reads as follows:

$$f_Y(y) = \int_y^\infty \frac{y}{x \sqrt{x^2 - y^2}} f_X(x) dx. \quad (2)$$

Expressing Eq. (2) in matrix form (by a quadrature discretization of the problem), we get:

$$Y = A X. \quad (3)$$

Now, A is a matrix representing the kernel $p(y|x)$, Y is a vector representing the density of projected rotational velocities $f_Y(y)$ and X is the unknown vector representing the density of “true” rotational velocities $f_X(x)$.

Since the observed data are measured with error, the last equation is an example of a discrete ill-posed problem, that is, small errors in the measured data can produce large variations in the recovered function which make the solution unstable (Ivanov et al. 2002, and references therein). Nevertheless, in the decades following the work of Chandrasekhar & Münch (1950), much mathematical theory on this kind of problem has been developed. Among others, one of the most common methods is the Tikhonov regularisation method (Tikhonov & Arsenin 1977; Tikhonov et al. 1995; Hansen 2010).

The standard method to solve Eq. (3) is to apply ordinary least squares (OLS), that is, $\min\{\|A X - Y\|^2\}$, where $\|\cdot\|$ represents the euclidean norm. However, for ill-posed problems this method fails in the sense that it can produce unstable estimators. In order to avoid this problem the Tikhonov regularization method imposes a regularization term to be included in the minimization process, namely:

$$\min\{\|A X - Y\|^2\} \rightarrow \min\{\|A X - Y\|^2 + \lambda^2 \|L(X - X_0)\|^2\}, \quad (4)$$

where λ is the Tikhonov factor. The standard definition for the L matrix is $L = I$, where I is the identity matrix and X_0 is an initial estimation, setting $X_0 = 0$, when there is no previous information. There exist different quantitative approaches to obtain the Tikhonov factor, for example, Generalized Cross-Validation (GCV), L-curve Method, Discrepancy Principle, and Restricted Maximum Likelihood. More details of these are explained in, Press et al. (2007), Hansen (2010) and Tikhonov & Arsenin (1977), for example. Once the λ -value is attained, the solution X_λ of the regularized problem by the Tikhonov method is given by:

$$X_\lambda = (A^T A + \lambda^2 I)^{-1} A^T Y. \quad (5)$$

In this article we use the Tikhonov regularization method using singular value decomposition (see Appendix A for details) to deconvolve the distribution of the rotational stellar velocities.

In the data analyzed in this article the L-curve method failed, that is, we do not obtain the “L” shape in the L-curve plot, but only the horizontal part of it (see details in Appendix B). For this reason we propose the method described below to select the Tikhonov factor based on the fact that, when $\lambda \rightarrow 0$, X_λ tends to the exact solution X , the difference between two regularized solutions tends to 0. In Monte Carlo runs (Sect. 3) the Tikhonov factor has been calculated with our proposed method (see below). We proved (Sect. 3) empirically that the Tikhonov estimator we obtained is unbiased and consistent; both desirable properties of any statistical estimator.

We determine the value of the Tikhonov factor, λ , using the following iterative procedure, which turned out to be fast and efficient in obtaining the regularization parameter in the case of smooth solutions:

- i) We start with an initial value of λ ($\lambda = \lambda_0$).
- ii) In each following iteration we reduce the value of λ by a factor f , ($\lambda_j = f^j \lambda_0$), we use typically $f = 0.99$.
- iii) At iteration step j we calculate the difference between the corresponding regularization solutions: $\phi = \|X_{\lambda_j} - X_{\lambda_{j-1}}\|$.
- iv) If ϕ is small enough, that is, $\phi < \epsilon$, we stop the iterative process and get the value of λ . Typically a value of $\epsilon = 10^{-7}$ has been used in this procedure.

In Appendix B, we show the criteria for selecting λ_0 and factor f .

3. Monte Carlo simulation

In this section we present the results of Monte Carlo numerical simulations to assess the performance of the Tikhonov regularization method in deconvolving rotational velocity distribution from the Fredholm integral. Our Monte Carlo runs consist of $n_{MC} = 1000$ independent replications for each chosen scenario, where we considered two specific distributions of rotational velocities. Therefore, we simulate 30 different cases, described as follows:

- a) Unimodal distribution: we chose a Maxwellian distribution

$$f_M(x) = \sqrt{\frac{2}{\pi}} \frac{1}{\sigma^3} x^2 e^{-\frac{x^2}{2\sigma^2}}, \quad x > 0, \quad (6)$$

with parameter $\sigma = 8$, which is the same distribution used in Curé et al. (2014). Furthermore, we considered three different cases, each one including an additive error from a uniform distribution $U[-\sigma_\epsilon, \sigma_\epsilon]$, with PDF given by $f_U(x) = 1/(2\sigma_\epsilon)$ for $-\sigma_\epsilon \leq x \leq \sigma_\epsilon$. The chosen values of σ_ϵ are $\sigma_\epsilon = 0.5, 1, 2$ (km s⁻¹).

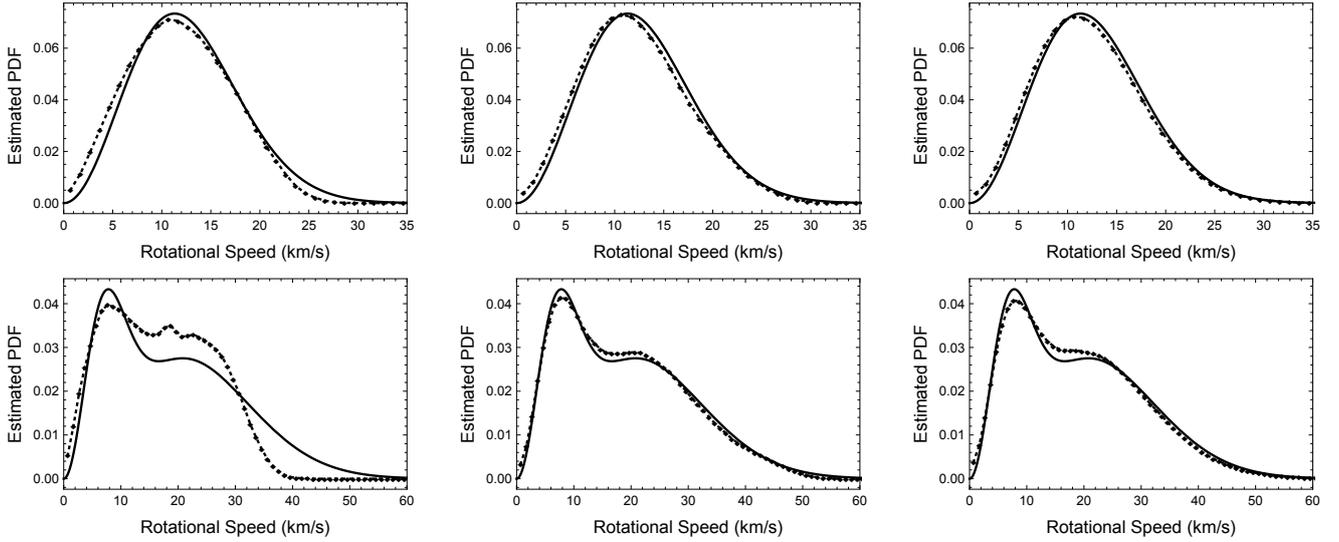


Fig. 1. *Upper panels:* univariate Maxwellian distribution, with parameter $\sigma = 8$, is shown with a solid line in all *upper panels*; black squares connected by dashed lines represent the mean of the $n_{\text{MC}} = 1000$ samples of Tikhonov regularization. Results are for: $n_s = 30$ with $\sigma_\epsilon = 0.5$ (*upper left*), $n_s = 100$ with $\sigma_\epsilon = 1$ (*upper center*) and $n_s = 1000$ with $\sigma_\epsilon = 2$ (*upper right*). *Lower panels:* bivariate Maxwellian distributions are shown by a solid line in all *lower panels*, with parameters $\sigma_1 = 5$ and $\sigma_2 = 15$ and amplitudes $A = 0.7$ and $B = 0.3$. Black squares connected by a dashed line show the estimated PDFs obtained by Tikhonov regularization. Results are for: $n_s = 30$ with $\sigma_\epsilon = 0.5$ (*lower left*), $n_s = 300$ with $\sigma_\epsilon = 1$ (*lower center*) and $n_s = 1000$ with $\sigma_\epsilon = 2$ (*lower right*).

- b) Bimodal distribution: for a mix of two Maxwellian distributions

$$f_{2M}(x) = \sqrt{\frac{2}{\pi}} \frac{x^2}{A+B} \left(\frac{A}{\sigma_1^3} e^{-\frac{x^2}{2\sigma_1^2}} + \frac{B}{\sigma_2^3} e^{-\frac{x^2}{2\sigma_2^2}} \right), \quad x > 0, \quad (7)$$

dispersion parameters are: $\sigma_1 = 5$ and $\sigma_2 = 15$, and amplitudes: $A = 0.3$ and $B = 0.7$. We consider the same additive error cases as for the unimodal distribution.

Furthermore, for both uni- and bimodal cases, we consider five sample lengths n_s : $n_s = 30, 100, 300, 1000, 10000$.

For each independent Monte Carlo sample two samples needed to be simulated, one from the distribution of the rotational velocities (uni- or bimodal) and another for the kernel, $p(y|x)$, representing the distribution of the inclination angles in the Fredholm integral (Eq. (2)). Then, we multiplied each element of the first sample with the corresponding element of the second sample and added the error term. This gave the final sample of $v \sin i$ for each scenario. The following step was to estimate the PDF of the projected rotational velocities with a Kernel Density Estimator (KDE, Silverman 1986). Using a grid of n_g points we discretized the Fredholm integral obtaining the linear system (Eq. (4)). With this data we calculated the Tikhonov factor λ using the procedure described above and obtained the Tikhonov regularization solution, X_λ , which is the estimated PDF of rotational speeds.

Figure 1 upper panels show, using a solid line, the original Maxwellian distribution (Eq. (6)) together with the mean estimated PDF of all Monte Carlo simulations (black squares connected by a dashed line) for different values of σ_ϵ and n_s . It is clearly shown that sample lengths of order $n_s \sim 30$ give acceptable results when compared with the original sample. For larger sample lengths, $n_s \gtrsim 100$, the agreement between the original distribution and the mean of the estimated PDFs is almost exact. Although the mean estimated distribution is slightly shifted to lower velocities. Lower panels of Fig. 1 show the original bimodal mixed Maxwellian distributions (in a solid line) together

with the mean estimated PDF (black squares connected by a dashed line). When a sample length is of the order $n_s \sim 30$, a difference between the estimated PDF and the original PDF is observed. Nevertheless, the Tikhonov regularized solution retrieves the bimodality and delivers the approximate position of the maximum of both components, but gives a false estimate of the tail of the original distribution.

In the other cases ($n_s \gtrsim 100$) the mean of the Tikhonov regularization solutions is very close to the original mixture of Maxwellian distributions, although the estimated value of the amplitudes is slightly lower (first distribution) and slightly higher (second distribution) than the original one.

In order to quantify the error of the estimated PDF, we calculated (following Curé et al. 2014) the mean integrated square error (MISE), that is:

$$\text{MISE} = \frac{1}{n_{\text{MC}}} \sum_{j=1}^{n_{\text{MC}}} \left(\frac{1}{n_g} \sum_{i=1}^{n_g} (\hat{f}_j(x_i) - f(x_i))^2 \right), \quad (8)$$

where $f(x)$ represents the original distribution function of rotational speeds and $\hat{f}_j(x)$ represents the estimated Tikhonov regularization density of the j -run in Monte Carlo simulations.

In the left panel of Fig. 2 we plotted the MISE as a function of sample length for $\sigma_\epsilon = 0.5$. In the other cases ($\sigma_\epsilon = 1, 2$) the MISE value is very similar to values $\text{MISE} \lesssim 10^{-4}$. Also it can be seen that as the sample size increases, MISE tends to zero, that is, $\text{MISE}(\hat{f}) \rightarrow 0$ when $n_s \rightarrow \infty$.

The right panel in Fig. 2 shows Tikhonov factor as a function of sample size. These factors are of the same order of magnitude for both types of distribution (uni- and bimodal). Our simulations confirm for all sample lengths and different σ_ϵ values that the Tikhonov factor (λ) is almost independent of the magnitude of the error σ_ϵ . Furthermore, since the Tikhonov parameter changes only slightly as a function of sample size, we can consider the Tikhonov factor as being almost independent of the sample length, n_s .

To confirm this result, we performed MC simulations with a fixed value of λ . The range of λ was from 0.002 to 0.01 with a

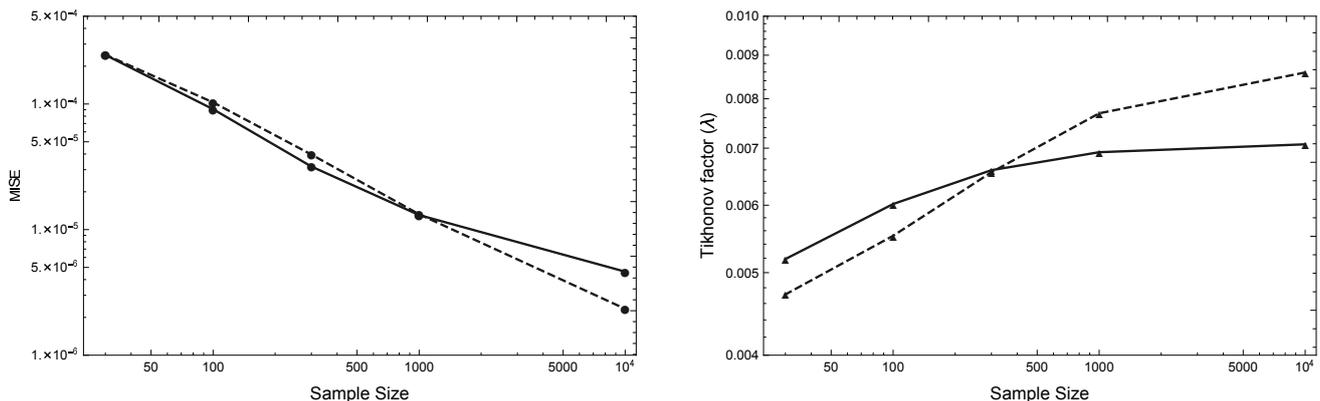


Fig. 2. Left panel: value of MISE (black dots) from the estimated PDF for univariate distributions (solid line) and bivariate distributions (dashed line) both using $\sigma_\epsilon = 0.5$. Right panel: value of Tikhonov factor λ as a function of sample size for the cases where $\sigma_\epsilon = 2$. See text for details.

step of $\Delta_\lambda = 0.001$. We calculated the MISE from $n_{MC} = 1000$ samples, each with a size of $n_s = 1000$ for each value of λ . For the unimodal Maxwellian distribution the values of the MISE vary, increasing from 7.319×10^{-5} to 7.320×10^{-5} for this range of λ , an almost negligible difference. In the case of a bimodal Maxwellian distribution the scenario is very similar. Using the same range of λ , the MISE values vary, increasing from 4.717×10^{-5} to 4.718×10^{-5} . Similar behaviour is found when $n_s = 30, 100, 300, 10000$, supporting our claim that the Tikhonov factor (λ) is almost independent of the sample size.

By means of the average of the estimated PDFs we can estimate the expected value for the Tikhonov regularization solution. In all cases, the mean of the estimated PDFs is very close to the original unimodal or bimodal distributions, and this mean probability density function is closer to the true PDF when increasing the sample size.

This fact shows, empirically, that the studied estimator is asymptotically unbiased. Therefore, since MISE tends to zero when n_s tends to infinity, it is implied that the variance of the Tikhonov regularization estimator tends to zero as well and hence is a consistent estimator.

4. Deconvolving a real sample

In this section, we perform the following steps: (i) apply the Tikhonov regularization method to a sample of measured $v \sin i$ data of cluster stars in order to estimate the rotational velocity probability density distribution; and (ii) compare the application of different methods to deconvolve the velocity distribution together with previous non-parametric results from the literature.

4.1. Tarantula sample

We select the Tarantula sample for single O-type stars from the VLT Flames Tarantula Survey, where Ramírez-Agudelo et al. (2013) deconvolved the rotational velocity distribution using the Lucy (1974) method (see also Richardson 1972). This sample contains 216 stars with $v \sin i$ data from 40 km s^{-1} up to 610 km s^{-1} . Following Ramírez-Agudelo et al. (2013) for comparison purposes, we also omitted the two largest values (outliers) of the sample. To build the Y vector, we used the KDE method with the following bandwidths (Silverman 1986, pages 45 and 47):

$$h_1 = 0.79 IQR n_s^{-1/5} \quad (9)$$

$$h_2 = 0.9 \min\{\Sigma, IQR/1.34\} n_s^{-1/5}, \quad (10)$$

here, IQR is the interquartile range, Σ is the standard deviation of the sample and n_s is the sample length.

Figure 3 shows, by a solid line, the rotational velocity distribution after Tikhonov regularization. Our procedure for Tikhonov factor determination gives a value of $\lambda = 0.0174$ for a step of $\Delta x = 2 \text{ km s}^{-1}$. Figure 3 also shows the confidence intervals calculated using a bootstrap method ($n_{BS} = 3000$, Efron & Tibshirani 1993). The lower is the bandwidth, the wider is the confidence interval. The “bump” around 400 to 450 km s^{-1} is wider in our case ranging from 340 km s^{-1} to 480 km s^{-1} . This discrepancy is probably due to the use of the KDE method with a Gaussian kernel in Y .

4.2. Comparing results

For the Tarantula sample, we calculate the CDF by direct integration of the PDF obtained via the Tikhonov regularization method and compare this to the CDF calculated using the method described by Curé et al. (2014). Figure 4 shows both CDFs. Their agreement is remarkable. In addition to our results for the PDF, Fig. 3 also shows the PDF (dashed lines) obtained from Ramírez-Agudelo et al. (2013, see their Fig. 17) calculated using the Lucy (1974) method. It can clearly be seen that Lucy-PDF lies inside our confidence interval.

In order to evaluate whether or not both estimated PDFs correspond to the same distribution, we obtained the q-q plot and calculated the respective quantiles. Figure 5 shows the q-q plot of these densities, confirming that both come from the same probability distribution.

5. Conclusions

In this work we have obtained the estimated probability distribution function of “true” rotational velocities using the Tikhonov regularization method. Furthermore, this estimated PDF uses a Tikhonov parameter λ obtained by means of an iterative method with a specific stopping criterion in comparison to the widely used iterative method of Lucy (1974).

Through Monte Carlo numerical simulations we assess the proposed method in two cases: when the rotational velocity distribution is described by a Maxwell distribution and for a mixture of two Maxwell distributions. For each situation different scenarios were evaluated obtaining satisfactory results for all of them, except for $n_s = 30$, when the velocities are described by a mixture of two Maxwellian distributions.

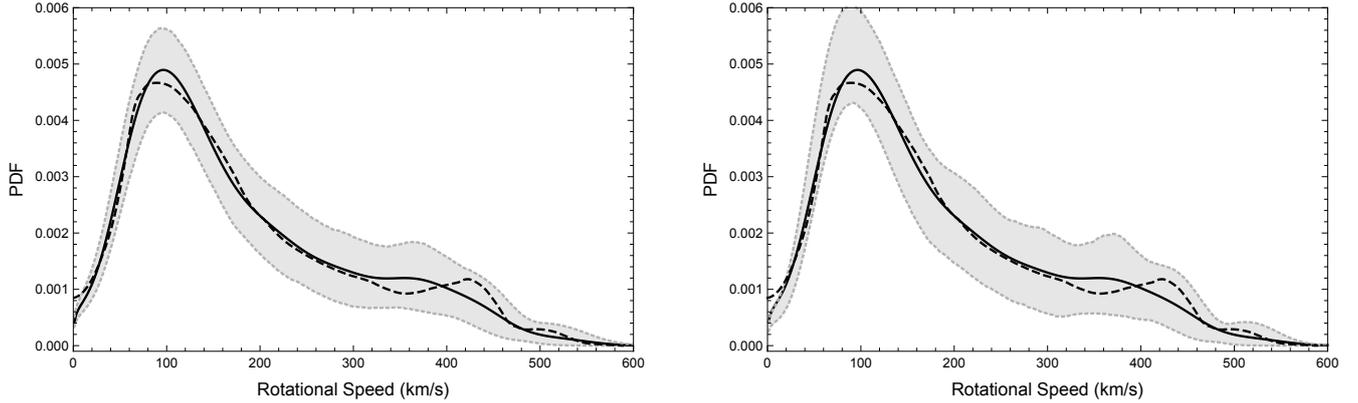


Fig. 3. Estimated PDF from the Tarantula sample (solid lines). Both panels with $\lambda = 0.0174$ and $\Delta x = 2 \text{ km s}^{-1}$, *Left panel* with bandwidth $h_1 = 35.676$ and *right panel* with bandwidth $h_2 = 30.313$. Gray-shaded regions represent the 2.5% (lower) and 97.5% (upper) confidence intervals calculated using a bootstrap method. Dashed lines show the PDF (from Ramírez-Agudelo et al. 2013) obtained using the Lucy (1974) method.

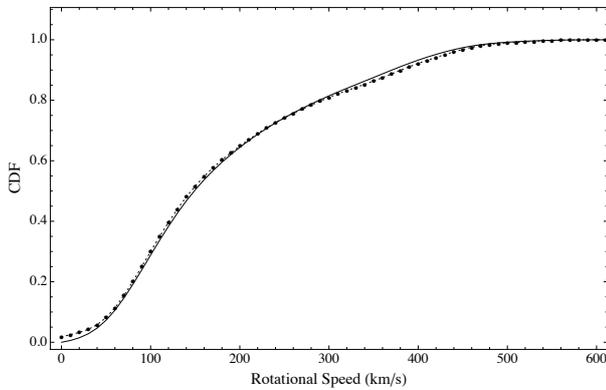


Fig. 4. Estimated cumulative rotational velocity distribution function for the Tarantula sample (solid line) obtained using Tikhonov regularization using a spacing of $\Delta x = 2 \text{ km s}^{-1}$ for the velocities. Dots connected by dashed lines show the CDF calculated using the Curé et al. (2014) method with a spacing of $\Delta x = 10 \text{ km s}^{-1}$.

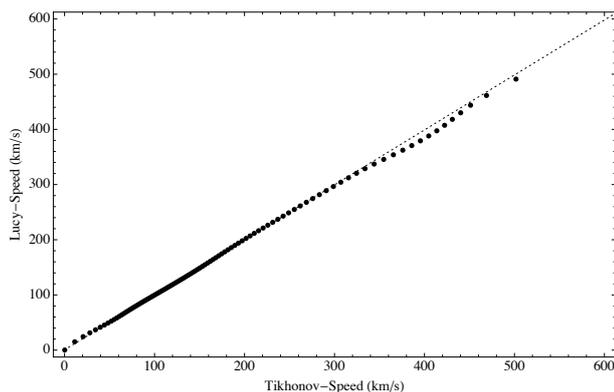


Fig. 5. q-q plot from the Tarantula sample, black dots represent the quantiles of each distribution, one calculated using the Tikhonov regularization method and the other calculated using the Lucy method (data from Ramírez-Agudelo et al. 2013).

This method retrieves the typical rotational velocity distribution for uni- and bimodal distributions. We showed, empirically, that the studied estimator is asymptotically unbiased and its variance tends to zero. Furthermore, we use the MISE as a measure of goodness of fit of the estimated PDF, which is approximately

equal to or less than 10^{-4} for all sample sizes and tends to zero when n_s tends to infinity.

We apply this method to a set of observed data from the Tarantula cluster (Ramírez-Agudelo et al. 2013). The estimated PDF from the Tikhonov regularization method agreed very well with the PDF obtained using the Lucy method, as shown by the q-q plot, demonstrating the method’s ability to deconvolve rotational velocities distribution.

In comparison with the method that delivers the CDF described by Curé et al. (2014), Tikhonov regularization solution gives, by direct integration of the PDF, almost the same non-parametric estimation of the true underlying cumulative distribution function of rotational velocities.

To summarize, we previously developed a method to obtain the CDF of “true” rotational velocities (Curé et al. 2014). This current publication presents the Tikhonov regularization method for obtaining the corresponding PDF directly from the Fredholm integral. Both methods calculate in a simple and straightforward way (PDF, CDF), without any assumptions of the underlying distribution.

Future work: we would like to develop a general function of the kernel of the Fredholm integral, $p(y|x)$, in order to describe an arbitrary orientation of rotational axes. Thus, we can study the distribution of rotational speeds relaxing the standard assumption of uniformity of stellar axes.

Acknowledgements. A.C. thanks the support from the Instituto de Estadística, Pontificia Universidad Católica de Valparaíso. P.E. thanks the support from the Advanced Center for Electrical and Electronic Engineering, AC3E, Basal Fund Conicyt FB0008. M.C. thanks the support from the Centro de Astrofísica de Valparaíso and the Centro Interdisciplinario de Estudios Atmosféricos y Astroestadística. J.C. and D.R. thanks project “Métodos de deconvolución aplicados a la astronomía”, Ministerio de Educación – Secretaría de Políticas and Universitarias, the Universidad Nacional de General Sarmiento and the Universidad de Buenos Aires for their financial support. D.R. also acknowledges the support of the project PIP11420090100165, CONICET.

References

- Bouhamidi, A., & Jbilou, K. 2007, *J. Comput. Appl. Math.*, **206**, 86
- Burger, M. 2007, *Inverse Problems*, Lecture Notes, Winter 2007/08 (University Muenster)
- Chandrasekhar, S., & Münch, G. 1950, *ApJ*, **111**, 142
- Curé, M., Rial, D. F., Christen, A., & Cassetti, J. 2014, *A&A*, **565**, A85
- Deng, L.-J., Huang, T.-Z., Zhao, L., & Wang, S. 2013, *J. Opt. Soc. Am. A*, **30**, 5

- Efron, B., & Tibshirani, R. J. 1993, *An Introduction to the Bootstrap* (Chapman & Hall CRC)
- Eggermont, P. P. B. 1993, *SIAM J. Math. Anal.*, 24, 6
- Fomel, S. 2007, *Geophysics*, 72, 29
- Hansen, P. C. 2010, *Discrete Inverse Problems: Insight and Algorithms* (SIAM)
- Ivanov, V., Vasin, V., & Tanana, V. 2002, *Theory of linear ill-posed problems and its applications* (Utrecht, Boston: VSP)
- Lucy, L. B. 1974, *AJ*, 79, 745
- Lucy, L. B. 1994, *Rev. Mod. Astron.*, 7, 31
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2007, *Numerical recipes* (Cambridge University Press)
- Ramírez-Agudelo, O. H., Simón-Díaz, S., Sana, H., et al. 2013, *A&A*, 560, A29
- Richardson, W. H. 1972, *J. Opt. Soc. America*, 62, 55
- Silverman, B. W. 1986, *Density estimation for statistics and data analysis. Monographs on Statistics and Applied Probability No. 26* (London: Chapman and Hall)
- Tikhonov, A. N. 1943, *C. R. (Doklady) Acad. Sci. URSS (N. S.)*, 39, 176
- Tikhonov, A. N. 1963, *Soviet Math Dokl* 4, 1035, English translation of Dokl Akad Nauk SSSR, 151, 501
- Tikhonov, A. N., & Arsenin, V. Y. 1977, *Solution of Ill-posed Problems* (Washington: Winston & Sons)
- Tikhonov, A. N., Goncharsky, A. V., Stepanov, V. V., & Yagola, A. G. 1995, *Numerical Methods for the Solution of Ill-Posed Problems* (Kluwer Academic Publishers)

Appendix A: The Tikhonov regularization method

In this Appendix we give a brief description of the Tikhonov regularization method closely following Burger (2007) and Eggermont (1993). Suppose that we have a linear system of the form

$$\mathbf{A} \mathbf{X} = \mathbf{Y}, \quad (\text{A.1})$$

with a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, and vectors $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^n$. Suppose additionally that \mathbf{A} is a symmetrical, positive definite matrix. In this case, from spectral theory for symmetrical matrices there exist eigenvalues, $0 < \mu_1 \leq \dots \leq \mu_n$ and corresponding eigenvectors $\mathbf{u}_i \in \mathbb{R}^n$, with the euclidean norm $\|\mathbf{u}_i\| = 1$, such that

$$\mathbf{A} = \sum_{i=1}^n \mu_i \mathbf{u}_i \mathbf{u}_i^T, \quad (\text{A.2})$$

where we consider $\mathbf{u}_i \in \mathbb{R}^{n \times 1}$.

Since the solution of (A.1) is given by:

$$\mathbf{X} = \sum_{i=1}^n \mu_i^{-1} \mathbf{u}_i \mathbf{u}_i^T \mathbf{Y}, \quad (\text{A.3})$$

small eigenvalues of \mathbf{A} can cause numerical difficulties when they are arbitrarily close to zero and the problem is ill-posed. The condition number $\kappa := \mu_n/\mu_1$, is a measure of the stability of the system. For simplicity we shall assume that $\mu_n = 1$ then $\kappa = 1/\mu_1$. When we have data with error \mathbf{Y}_δ instead of \mathbf{Y} , satisfying $\|\mathbf{Y}_\delta - \mathbf{Y}\| < \delta$, we obtain a solution \mathbf{X}_δ and the error in the solution is:

$$\|\mathbf{X}_\delta - \mathbf{X}\|^2 = \sum_{i=1}^n \mu_i^{-2} |\mathbf{u}_i^T (\mathbf{Y}_\delta - \mathbf{Y})|^2 \leq \mu_1^{-2} \|\mathbf{Y}_\delta - \mathbf{Y}\|^2, \quad (\text{A.4})$$

then $\|\mathbf{X}_\delta - \mathbf{X}\|^2 \leq \kappa \delta$.

One observes that with increasing condition number the error amplification increases simultaneously. Often the nature of the error is unknown, then it is necessary used a method to solve the linear system that deals with error effects. The regularization methods face this problem efficiently. If matrix \mathbf{A} is positive semi-definite, its eigenvalues are non-negative, but it can have a zero eigenvalue. In this case, let μ_m be the smallest positive eigenvalue, then the solution of (A.1) becomes:

$$\mathbf{X} = \sum_{i=m}^n \mu_i^{-1} \mathbf{u}_i \mathbf{u}_i^T \mathbf{Y}, \quad (\text{A.5})$$

and the problem is solvable if and only if $\mathbf{u}_i^T \mathbf{Y} = 0$ for $i < m$.

For data with error, we can use the projection $\mathbf{P}\mathbf{Y}_\delta$ onto the range of \mathbf{A} . This analysis can be extended to a general matrix $\mathbf{A} \in \mathbb{R}^{n \times m}$ by considering the associated system $\mathbf{A}^T \mathbf{A} \mathbf{X} = \mathbf{A}^T \mathbf{Y}$, being that the matrix $\mathbf{A}^T \mathbf{A}$ is always symmetrical, positive and semi-definite.

In order to shift away from zero the smallest eigenvalues of this matrix $\mathbf{A}^T \mathbf{A}$, it seems natural to approximate $\mathbf{A}^T \mathbf{A}$ for a family of matrices $\mathbf{A}_\lambda := \mathbf{A}^T \mathbf{A} + \lambda \mathbf{I}$, whose eigenvalues are $\mu_i + \lambda$, where μ_i 's are the eigenvalues of $\mathbf{A}^T \mathbf{A}$.

We obtain an approximated solution $\mathbf{X}_\lambda = \mathbf{A}_\lambda^{-1} \mathbf{A}^T \mathbf{Y}$ and for data with error we have $\mathbf{X}_{\lambda,\delta} = \mathbf{A}_\lambda^{-1} \mathbf{A}^T \mathbf{Y}_\delta$. The error of the estimation is then

$$\|\mathbf{X} - \mathbf{X}_{\lambda,\delta}\|^2 \leq \|\mathbf{X} - \mathbf{X}_\lambda\|^2 + \|\mathbf{X}_\lambda - \mathbf{X}_{\lambda,\delta}\|^2. \quad (\text{A.6})$$

The first term on the right side corresponds to the approximation error and the second term corresponds to the error in data. Using spectral theory (Burger 2007), we obtain:

$$\|\mathbf{X} - \mathbf{X}_{\lambda,\delta}\|^2 \leq \frac{\lambda}{\mu_1(\mu_1 + \lambda)} (\|\mathbf{Y}_\delta\| + \delta) + \frac{\delta}{(\mu_1 + \lambda)}. \quad (\text{A.7})$$

The first term on the right side decreases when λ tends to zero while the second term on the right side increases when λ tends to zero. Thus, we have to find an estimation of λ that is a compromise between the error of the approximation and the error from measurements.

The solution of the Tikhonov regularization can be obtained also from the Singular Value Decomposition (SVD) of matrix \mathbf{A} . In this case, we write a general matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ with rank n in the form:

$$\mathbf{A}^T \mathbf{A} = \sum_{i=1}^n \sigma_i^2 \mathbf{v}_i \mathbf{v}_i^T, \quad (\text{A.8})$$

where \mathbf{u}_i and \mathbf{v}_i are orthonormal vectors of dimensions m and n respectively, and $\sigma_i \geq 0$ are the singular values of $\mathbf{A}^T \mathbf{A}$ such that $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$. Under this decomposition the Tikhonov solution is given by:

$$\mathbf{X}_\lambda = \sum_{i=1}^n f_i \frac{\mathbf{u}_i^T \mathbf{Y}}{\sigma_i} \mathbf{v}_i, \quad (\text{A.9})$$

where $f_i, i = 1, \dots, n$ are defined by $f_i = \sigma_i / (\sigma_i + \lambda^2)$.

As we mentioned in Sect. 2 there are several methods to estimate λ . The most frequently used are the L-curve Criterion, the Discrepancy Principle and Generalized Cross Validation.

The L-curve is a plot of $\log(\|\mathbf{A} \mathbf{X}_\lambda - \mathbf{Y}\|_2)$ versus $\log(\|\mathbf{X}_\lambda\|_2)$, the logarithm of two square euclidean norm, for different values of the Tikhonov factor λ . This plot has the characteristic ‘‘L’’ shape (see Fig. B.1). According to Hansen (2010) the Tikhonov solution \mathbf{X}_λ can be decomposed as $\mathbf{X}_\lambda = \bar{\mathbf{X}}_\lambda + \mathbf{X}_{\lambda,e}$, where $\bar{\mathbf{X}}_\lambda = (\mathbf{A}^T \mathbf{A} + \lambda^2 \mathbf{I})^{-1} \mathbf{A}^T \mathbf{Y}$ is the regularized version of the exact solution \mathbf{X} , and $\mathbf{X}_{\lambda,e} = (\mathbf{A}^T \mathbf{A} + \lambda^2 \mathbf{I})^{-1} \mathbf{A}^T e$ is the solution obtained by applying Tikhonov regularization to the error component e . For small values of λ , the error dominates the L-curve because the regularized solution \mathbf{X}_λ is dominated by $\mathbf{X}_{\lambda,e}$ and for large values of λ , \mathbf{X}_λ is dominated by $\bar{\mathbf{X}}_\lambda$, the unperturbed term. The λ chosen achieves a compromise between the two parts, allocated in the corner of the L-curve. The L-curve criterion for choosing the regularization factor is one of the most frequently used methods. The advantages are robustness and ability to manage observations with correlated errors. The limitations of the L-curve are the reconstruction of very smooth, exact solutions and treatment with a large amount of data (Hansen 2010).

Appendix B: Determination of regularization parameters

When we apply the L-curve method to different $v \sin i$ samples, the obtained values of the Tikhonov factor (λ) are ‘‘large’’. The large size of these values is due to the small values of the coefficients in singular value decomposition with almost constant singular values around 1, requiring the addition of too many terms to increase the norm of \mathbf{X} (the vertical part of the ‘‘L’’ shape, see Fig. B.1). We suspect that the reason for this is the smoothness of the solution (Hansen 2010). For the Tarantula sample (Sect. 4), the Tikhonov parameter delivered by the L-Curve, and the GCV methods, are the same, $\lambda = 0.2956$.

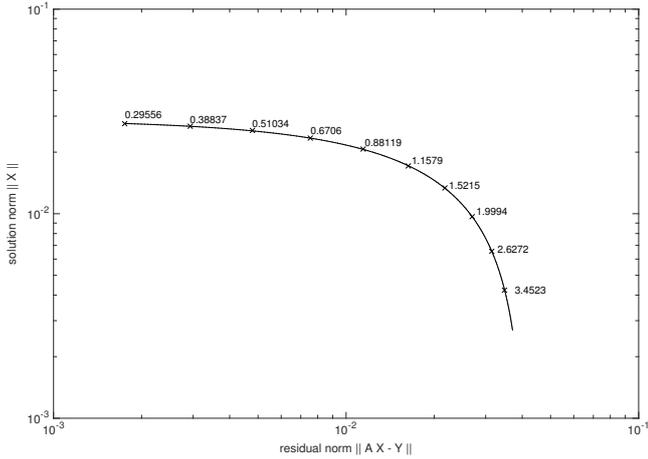


Fig. B.1. L-curve plot for the data obtained by the Monte Carlo sample for a unimodal distribution. Horizontal axis shows $\log(\|A X_\lambda - Y\|_2^2)$, i.e., the residuals of the regularization. Vertical axis shows $\log(\|X_\lambda\|_2^2)$, i.e., the norm of the regularization. Tikhonov parameter values (λ) are overplotted to the corresponding data points. Only the horizontal part of the typical “L” shape is shown. This situation occurs with very smooth exact solutions. See text for details.

Here we show how to determine the value of λ_0 and the choice of factor (f) to select the parameters λ of the Tikhonov method.

As we stated at the end of Sect. 2, we start with an initial value of λ_0 and calculate the Tikonov method to obtain the PDF, ($X_\lambda(1)$), then we multiply λ_0 by a factor f and obtain a new value of $\lambda = \lambda(1) = \lambda_0 \times f$, and another PDF ($X_\lambda(2)$), after applying Tikhonov method. After “ m ” iterations we have a set of $\{\lambda(1), \lambda(2), \dots, \lambda(m)\}$.

Defining $\phi(j)$ as:

$$\phi(j) = \|X_\lambda(j) - X_\lambda(j-1)\| \quad (\text{B.1})$$

where $\|\cdot\|$ represents the euclidian norm. After these “ m ” iterations we also have a set of $\{\phi(2), \phi(3), \dots, \phi(m)\}$. The iteration stops when the value of $\phi(m)$ is less than a certain value of ϵ . In our case we chose $\epsilon = 10^{-7}$.

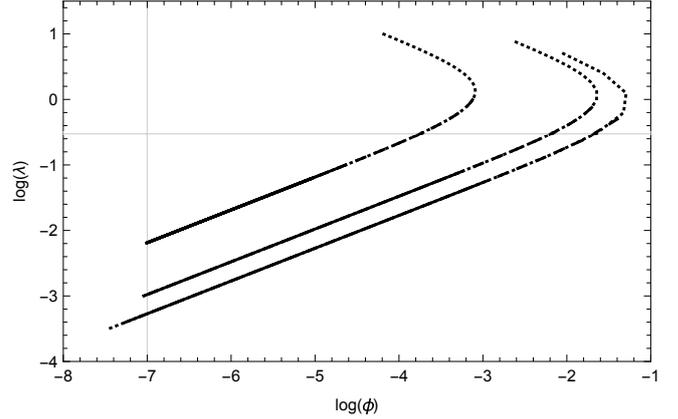


Fig. B.2. $\log(\lambda)$ versus $\log(\phi)$. Factor f varies from $f = 0.99$ (left), $f = 0.75$ (center) to $f = 0.5$ (right). Each of these curves starts with three initial values of λ_0 , dotted line ($\lambda_0 = 10$), dashed line ($\lambda_0 = 1$) and solid line ($\lambda_0 = 0.1$). See text for details. The vertical gray solid line shows the selected value of $\epsilon = 10^{-7}$ as a criterion to finish the iteration process. The horizontal gray solid line shows the value of $\lambda = 0.2956$ ($\log(\lambda) = -0.53$) obtained using the L–Curve or GCV methods.

Figure B.2 shows, $\log(\lambda)$ versus $\log(\phi)$ for different values of λ_0 and f , for the Tarantula sample. The initial values of λ_0 are: $\lambda_0 = 10$, shown as a dotted line in all three curves; $\lambda_0 = 1$, shown as dashed lines and $\lambda_0 = 0.1$, as solid lines.

For a given value of f , all three curves are superimposed, showing that the final value of λ is independent of the starting value λ_0 . Therefore we chose to start our calculations with $\lambda_0 = 0.1$. On the other hand, the critical parameter here is f ; the lower this value, the lower the final value of λ , when $\phi(m) \leq \epsilon$. Considering that λ is of the order λ^2 in Eq. (5), a very small parameter λ should be avoided in order to have a non-zero regularization term. Thus we select $f = 0.99$ as our default value to obtain the Tikhonov parameter λ .

It is clearly seen in Fig. B.2, that for $\log(\lambda) = -0.53$, that is, the value obtained by the L–Curve or GCV method (horizontal gray line) corresponds to a very “high” value of ϵ . If $f = 0.99$ then $\epsilon = 1.8 \times 10^{-4}$, value much larger than $\epsilon = 10^{-7}$, which is our criterion to stop this iteration process.