

LETTER TO THE EDITOR

Kolmogorov analysis of pulsar TOA

N. Galikyan^{1,2} , A. A. Kocharyan³ , and V. G. Gurzadyan^{1,4,*} 

¹ Center for Cosmology and Astrophysics, Alikhanian National Laboratory and Yerevan State University, Yerevan, Armenia

² Department of Physics, Sapienza University of Rome, Rome, Italy

³ School of Physics and Astronomy, Monash University, Clayton, Australia

⁴ SIA, Sapienza University of Rome, Rome, Italy

Received 12 January 2026 / Accepted 17 February 2026

ABSTRACT

The Kolmogorov stochasticity parameter (KSP) as a sensitive descriptor of the degree of randomness of signals was used to analyze the properties of the NANOGrav pulsar timing data associated with a stochastic gravitational wave background. The time of arrival (TOA) data of white noise for 68 pulsars were analyzed regarding their KSP properties. The analysis enabled us to obtain the degree of randomness of the white noise for various pulsars and to reveal its inhomogeneity, i.e., pulsars with low and high randomness of the white noise. The time dependence of the randomness in the white noise was also studied, indicating the existence of nonstationary physical processes influencing the pulsar timing. The KSP is thus acting as an indicator for the degree of the agreement between the observations and the timing models and as a test in revealing the contribution of various physical processes in the stochastic background signal.

Key words. methods: data analysis – methods: numerical – pulsars: general

1. Introduction

The data on the pulsar-timing monitoring obtained by the North American Nanohertz Observatory for Gravitational Waves (NANOGrav) (Agazie et al. 2023a, 2025) and other collaborations opened a new important window to trace the variety of ongoing processes in the Universe. The released pulsar timing array (PTA) data are linked to the stochastic gravitational-wave background (GWB) in the nanohertz band produced by binary supermassive black holes (Sesana et al. 2008; Agarwal et al. 2026; Sesana & Figueroa 2025), and they are also used to constrain several models such as those pertaining to the nature of dark matter and the evolution of early cosmology (e.g., Tiruvaskar & Gordon 2026; Ben-Dayan et al. 2026 and references therein).

The key issue in the pulsar timing is the analysis of the stochastic properties of the time of arrival (TOA) data of residuals for the studied 68 pulsars, for example the differences between the observed data and modeled parameters (Lam et al. 2025). Therefore, the study of tiny properties of the stochastic background by different methods and hence, of the residuals' association to white noise, can be crucial for both revealing the details from the pulsar astrometric parameters of pulsars up those of the supermassive black hole binaries (e.g. of their orbital eccentricities) or of the contribution of other effects (see Fiscella et al. 2025 and references therein). Moreover, white noise can contain contributions from instrumental noise, calibration of the radio telescopes, in addition to interference from terrestrial sources. Thus, while the comprehensive study of white noise is crucial for revealing the origin of the nanohertz gravitational wave background, the use of different descriptors can be of particular importance for the separation of contributions for the abovementioned effects in white noise among others.

For this study we applied the Kolmogorov stochasticity parameter (KSP) to analyze the properties of the white noise (cf. Barenboim et al. 2025) of TOA associated with both the TOA data and the pulsar timing models. The Kolmogorov stochasticity parameter is a sensitive quantitative descriptor of the degree of randomness of sequences (Kolmogorov 1933; Arnold 2008a). The KSP has been applied to the analysis of various signals (e.g., Arnold 2008a; Atto et al. 2013; Hoffman et al. 2023, including human genomic sequences (Gurzadyan et al. 2015), as well as those of an astrophysical origin (Gurzadyan & Kocharyan 2008; Gurzadyan et al. 2009, 2010; Gurzadyan & Durret 2011; Galikyan et al. 2025, see also Gurzadyan 1999). For example, the KSP enabled us to trace the properties of the cold spot as of a non-Gaussian region in the cosmic microwave background sky (Planck data) and to conclude on its void origin in the large-scale matter distribution (Gurzadyan et al. 2014); the void nature of the cold spot was soon after confirmed by an galaxy survey study (Szapudi et al. 2015).

Thus, we used the KSP for the following: (a) to quantify how independent the TOA residuals are; (b) to reveal how homogeneous the randomness is regarding individual pulsars, i.e., if there are pulsars attributed to low and high randomness signals; (c) to test if nonstationary effects can be traced in the white noise; and (d) to determine how closely the timing models for individual pulsars match their observations.

2. Kolmogorov stochasticity parameter

First, let us define the Kolmogorov stochasticity parameter as in (Kolmogorov 1933; Arnold 2008a,b,c, 2009a,b). Let X be a real-valued random variable. Its cumulative distribution function (CDF) is

$$F(x) \equiv \Pr(X \leq x), \quad (1)$$

* Corresponding author: gurzadyan@yerphi.am

where $\Pr(\cdot)$ denotes the probability under the assumed probabilistic model. Throughout this section, F refers to the theoretical (model/null) CDF.

Given an observed sample x_1, \dots, x_n (not necessarily distinct), let $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ denote the order statistics. The corresponding empirical CDF (ECDF) is defined by

$$F_n(x) \equiv \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{x_i \leq x\}, \quad (2)$$

which is a right-continuous step function,

$$F_n(x) = \begin{cases} 0, & x < x_{(1)}, \\ k/n, & x_{(k)} \leq x < x_{(k+1)}, \quad k = 1, \dots, n-1, \\ 1, & x_{(n)} \leq x. \end{cases} \quad (3)$$

Our basic goal is to quantify the discrepancy between the ECDF of the data and a reference (theoretical) CDF F . In applications, F represents a null/model hypothesis (e.g., a specified distribution, or a distribution implied by a simulated noise model). When F is fully specified and continuous, the resulting statistic is distribution-free under the null.

When the (one-sample) Kolmogorov distance is defined,

$$D_n \equiv \sup_{x \in \mathbb{R}} |F_n(x) - F(x)|, \quad (4)$$

and for its scaled form, the Kolmogorov stochasticity parameter,

$$\lambda_n \equiv \sqrt{n} D_n = \sqrt{n} \sup_{x \in \mathbb{R}} |F_n(x) - F(x)|. \quad (5)$$

Because the sample $\{X_i\}_{i=1}^n$ is random, the ECDF F_n and thus λ_n are random variables. Under the null hypothesis $H_0: X_1, \dots, X_n \sim F$ (with F continuous), λ_n has a distribution function

$$\Phi_n(\lambda) \equiv \Pr(\lambda_n \leq \lambda), \quad (6)$$

which converges to a universal limiting distribution $\Phi(\cdot)$ independent of F :

$$\Phi(\lambda) = \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2 \lambda^2}, \quad \Phi(0) = 0. \quad (7)$$

This is the statement of Kolmogorov's theorem (Kolmogorov 1933).

Note that, according to this theorem and the property of the limiting function Φ , the probable values of λ are within 0.3 and 2.2. In other words, at its lower values, $\Phi(\lambda)$ tends rapidly to 0 and to 1 at higher values.

An important feature of the KSP criterion is its applicability to relatively small sequences (orbits of dynamical systems) as is explicitly shown in Arnold (2008a), which compared two sequences of 15 two-digit numbers and concluded that the stochasticity probability of the one sequence is approximately 300 times higher than for other one. Then, for random sequences, the stochasticity parameter λ_n has a distribution close to the function $\Phi(\lambda)$, while the distribution is different from that function for nonrandom sequences.

The numerical experiments performed in Gurzadyan et al. (2011), using a single scaling of the ratio of stochastic to regular components of the generated sequences including a uniform distribution, revealed the efficiency of the KSP as an indicator of the degree of randomness of composite signals.

3. Pulsar timing

Pulsar timing is based on measuring TOAs of radio pulses with high precision and comparing them to the predictions of a timing model. The timing model encodes all deterministic contributions to the TOAs—such as the pulsar's spin evolution, astrometric parameters, binary motion (when present), and propagation effects. Any mismatch between the measured TOAs and the timing model – i.e., the timing residuals – contains information about additional physical or instrumental processes not captured by the deterministic model.

These stochastic contributions are commonly separated into low-frequency, temporally correlated “red noise” (e.g., intrinsic spin noise or a stochastic gravitational-wave background), which are the type of signals that researchers usually look for. It is important to note that not all pulsars exhibit a red-noise component, and the contribution to the red noise can come from various effects (Fiscella et al. 2025), which can introduce characteristic structures in timing residuals. High-frequency, uncorrelated “white noise” mainly arises from radiometer noise and measurement uncertainties, representing random fluctuations in the time series. The white-noise component can be calculated as $WhiteNoise = TOA - TimingModel - RedNoise$ (if present).

For our analysis, we used PINT (Luo et al. 2019, 2021; Susobhanan et al. 2024), a high-precision open-source package for pulsar timing data analysis written in Python. It is developed in collaboration with NANOGRAV and is largely used for analysis (Agazie et al. 2023a,b; Smith et al. 2023), among other things. It allows one to read TOA data and load timing models with their parameters, for example, provided by NANOGRAV (Agazie et al. 2025). Moreover, in PINT, one can simulate the white and red noise components of pulsar signals separately and estimate the white-noise contribution. The simulated white noise and the observed white noise should have the same statistical properties, i.e., the signal that we refer to as white noise in the observed data is indeed a white-noise signal. To test this, we used the KSP as a statistical characteristic of a signal and compared the observed signal with the simulated one as described in Sect. 4.1.

4. Analysis

Consider a time series of length N that represents the residual component of the TOAs. To obtain the KSP characteristic histogram for the series, we calculated λ inside a sliding window of length $10^{-1} \cdot N$ with a sliding step of $10^{-3} \cdot N$. Inside the window, we randomly selected 80% of the points and calculated λ using the generalized normal distribution as the theoretical distribution. Its probability density is given via Eq. (8). The final λ for that window is the mean of the 80% which was repeated five times. Namely, the algorithm estimated the distribution of λ , and its running allowed the λ distribution's stability to be estimated four more times, as 80% of the subsamplings each yielded different distributions. This enabled us to ensure the stability of the procedure, to define the errors bars, and to decrease the influence of outliers, if any. The specific choice of parameters – the 80% subsampling and the 5×5 repetitions – were made in order to adequately represent the distributions for comparison, while avoiding both over-averaging and undersampling:

$$f_{\beta}(x) = \frac{1}{2\alpha\Gamma(1/\beta)} e^{\left(\frac{|x-\mu|}{\alpha}\right)^{\beta}} \quad (8)$$

$$\alpha = \sqrt{\frac{\Gamma(1/\beta)}{\Gamma(3/\beta)}}; \quad \mu = 0$$

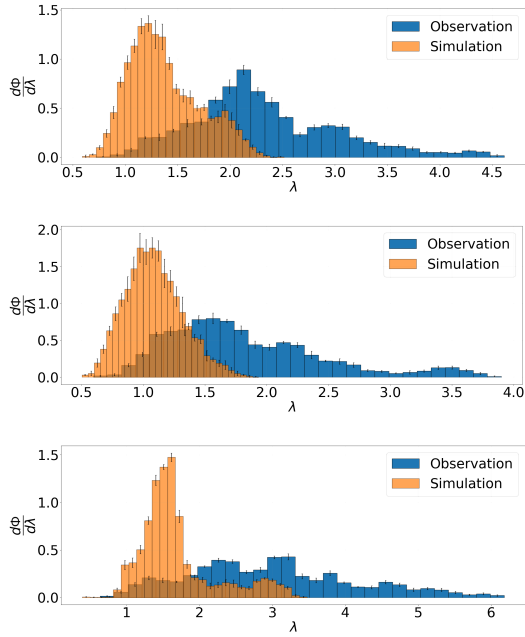


Fig. 1. KSP distributions for the simulated/observed data of pulsars’ (J1022+1001, J1125+7819, J1903+0327) white noise, when the histograms do differ. Orange and blue histograms correspond to simulations and observations, respectively.

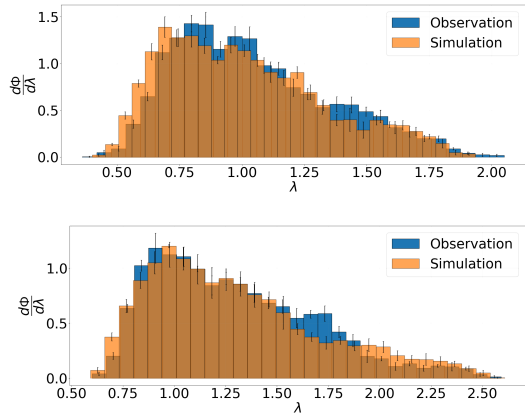


Fig. 2. Same as in Figure 1, but for sample pulsars (J1745+1017, J1910+1256) when the KSP distributions for the simulated/observed data are similar.

In Eq. (8) we have only a single free parameter β that describes the sharpness of the distribution. The scale parameter α was chosen such that the theoretical distribution had a unit variance. We considered two scenarios of data normalization and a choice of β . First, when β was a global parameter of a time series as described in Sect. 4.1, and, second, where β was local and had a time dependence (Sect. 4.2).

4.1. $\beta = \text{const}$

Consider a time series of residuals of a pulsar, $\{R_i\}_{i=1}^N$, where N is the length of the series. We assumed that they are identically distributed. In this scenario, we first normalized our series by subtracting the mean of the series \bar{R} and dividing by its standard deviation ΔR . Hence, we obtain normalized series $r_i = \frac{R_i - \bar{R}}{\Delta R}$.

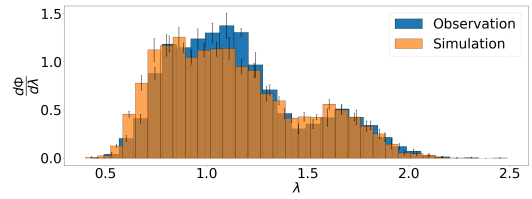


Fig. 3. Case of pulsar J1125+7819, when the KSP distributions became similar after introducing $\beta(t)$.

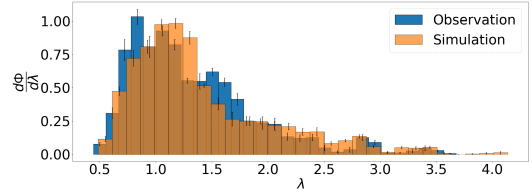


Fig. 4. KSP distributions for pulsar J1910+1256. Introducing the time dependence in β does not significantly affect χ^2 .

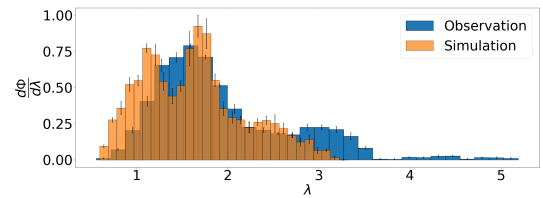


Fig. 5. KSP distributions for the pulsar J1022+1001 after introducing the time dependence in β . Although $\langle \chi^2 \rangle$ is smaller in the $\beta(t)$ approach (Table 2) than in the $\beta = \text{const}$ case (Table 1), they are of the same order.

We performed this procedure on observed and simulated data using PINT (Luo et al. 2019, 2021; Susobhanan et al. 2024). Simulations used the same models, and the parameters within them were used to extract residuals from the TOA. We observed that some of the pulsars’ KSP distributions are not similar, as shown in Fig. 1, while for others they are, as shown in Fig. 2.

For a quantitative picture, we created five epochs of simulations for each pulsar presented, and calculated the reduced χ^2 of the KSP histograms between observed and simulated data. To calculate χ^2 , we linearly interpolated histogram values and their corresponding errors into the same bins. Table 1 shows the minimum, maximum, mean, median, and standard deviation of χ^2 for each pulsar. Thus, the key point of our analysis is that the same numerical scheme was performed both for the data of observations and simulations.

4.2. $\beta = \beta(t)$

Now we assume that the properties of the distribution may change over time, hence β has a time dependence. We calculated β around each residual using maximum likelihood estimation. The algorithm for finding β_i for a residual time series R_i is the following:

1. We weighed each residual value by weight,

$$w_{ik} \propto \exp\left[-\frac{1}{2}\left(\frac{t_k - t_i}{h}\right)^2\right], \quad k = \overline{1; N},$$

Table 1. Reduced- χ^2 data for each pulsar for $\beta = const.$

	min(χ^2)	max(χ^2)	med(χ^2)	$\langle \chi^2 \rangle$	$\sigma(\chi^2)$
J1022+1001	64	110	102	92	18
J1125+7819	40	150	70	80	40
J1745+1017	2.7	7.8	4.4	4.5	1.8
J1802-2124	52	86	72	71	12
J1832-0836	2	8	5	5	2
J1903+0327	68	112	78	84	15
J1910+1256	2	20	5	8	7
J1918-0642	5.3	7.6	6.4	6.4	0.8
J2043+1711	4	17	6	8	5
J2317+1439	36	73	48	51	14

Table 2. Same as in Table 1, but for $\beta = \beta(t).$

	min(χ^2)	max(χ^2)	med(χ^2)	$\langle \chi^2 \rangle$	$\sigma(\chi^2)$
J1022+1001	21	60	32	36	13
J1125+7819	2	14	8	7	4
J1745+1017	5	30	11	13	8
J1802-2124	30	90	60	60	20
J1832-0836	2	9	5	5	2
J1903+0327	18	54	43	41	12
J1910+1256	5	19	7	10	5
J1918-0642	6	23	16	15	6
J2043+1711	5	15	11	10	3
J2317+1439	12	26	14	18	6

where t_i, t_k are the time moments for R_i and R_k , respectively, and h is a hyperparameter that was chosen via cross-validation.

2. We calculated the weighted mean and standard deviation and obtained normalized time series for each time moment t_i :

$$r_{ik} = \frac{R_k - \bar{R}_i}{\Delta R_i}.$$

3. We estimated β_i as

$$\beta_i = \operatorname{argmin}_{\beta} \left[- \sum_k w_{ik} \log f_{\beta}(r_{ik}) \right].$$

The following procedure of calculating the KSP histogram is the same, with the exception that we normalized the distribution inside the sliding window similarly to Section 4.1, i.e., without the weights. Table 2 shows the χ^2 values for this scenario.

Figures 3, 4, 5 illustrate examples of three differentiable cases: non-similar distributions became similar; and similar and non-similar KSP distributions remained as they were.

5. Conclusions

We analyzed NANOGrav pulsar timing data using Kolmogorov's method. Namely, we concentrated on the white-noise component of the observations, as it is the remainder after subtracting the known physical contributions from the TOAs. Comparing the observed white noise with simulated white noise and using the Kolmogorov stochasticity parameter of the signal as a metric enabled us to probe the possible physical components in the TOAs. These physical components may be deterministic, and thus they might be included in the timing models, or stochastic,

as part of the red noise (e.g., the GWB). The analysis was performed blindly with the same algorithm applied to all pulsars considered.

Assuming a generalized normal distribution – with a single free β parameter describing the sharpness of the distribution – as the underlying distribution for the white noise, we find both cases, when the white noise stochasticity agrees with the observations and simulations and cases where it does not. The introduction of a time-dependent $\beta(t)$ improved the similarity; however, instances remain in which the observational and simulated stochasticity still disagree. This can indicate the existence of nonstationary physical processes influencing the pulsar timing.

The KSP method can be used to test different models affecting pulsar timing, including for example modified theories of gravity or models for the dark sector of the universe (e.g., Gurzadyan et al. 2023, 2025, 2026). Namely, then the task will be to reproduce TOAs with the observed KSP properties, especially depending upon the availability of higher precision and more observational data.

Acknowledgements. We are thankful to the referee for many valuable comments. We acknowledge the use of The NANOGrav 15-Year Data Set; <https://zenodo.org/records/16051178> and ASCL Code Record of Astrophysics Source Code Library ASCL.net; ascl:1902.007.

References

- Agarwal, N., Agazie, G., Anumarlapudi, A., et al. 2026, *ApJ*, 988, L11
- Agazie, G., Anumarlapudi, A., Archibald, A. M., et al. 2023a, *ApJ*, 951, L8
- Agazie, G., Alam, M. F., Anumarlapudi, A., et al. 2023b, *ApJ*, 951, L9
- Agazie, G., Kaplan, D. L., Susobhanan, A., et al. 2025, arXiv e-prints [arXiv:2510.16668]
- Arnold, V. I. 2008a, *Nonlinearity*, 21, T109
- Arnold, V. I. 2008b, *Uspekhi Mat. Nauk*, 63, 5
- Arnold, V. I. 2008c, IC/2008/001, ICTP, Trieste
- Arnold, V. I. 2009a, *Trans. Moscow Math. Soc.*, 70, 31
- Arnold, V. I. 2009b, *Funct. An. Other Math.*, 2, 139
- Atto, A. M., Berthoumieu, Y., & Megret, R. 2013, *Entropy*, 15, 4782
- Barenboim, G., Ireland, A., & Stebbins, A. 2025, arXiv e-prints [arXiv:2511.13866]
- Ben-Dayan, I., Kumar, U., & Verma, A. 2026, *J. High Energy Astrophys.*, 50, 100510
- Fiscella, S. V. S., Lam, M. T., Agazie, G., et al. 2025, arXiv e-prints [arXiv:2509.21203]
- Galikyan, N., Kocharyan, A. A., & Gurzadyan, V. G. 2025, *A&A*, 696, L21
- Gurzadyan, V. G. 1999, *EPL*, 46, 114
- Gurzadyan, V. G., & Durret, F. 2011, *EPL*, 95, 69001
- Gurzadyan, V. G., & Kocharyan, A. A. 2008, *A&A*, 492, L33
- Gurzadyan, V. G., Allahverdyan, A. E., Ghahramanyan, T., et al. 2009, *A&A*, 497, 343
- Gurzadyan, V. G., Kashin, A. L., Khachatryan, H. G., et al. 2010, *EPL*, 91, 19001
- Gurzadyan, V. G., Ghahramanyan, T., & Sargsyan, S. 2011, *EPL*, 95, 19001
- Gurzadyan, V. G., Kashin, A. L., Khachatryan, H., et al. 2014, *A&A*, 566, A135
- Gurzadyan, V. G., Yan, H., Vlahovic, G., et al. 2015, *Roy. Soc. Open Sci.*, 2, 150143
- Gurzadyan, V. G., Fimin, N. N., & Chechetkin, V. M. 2023, *A&A*, 677, A161
- Gurzadyan, V. G., Fimin, N. N., & Chechetkin, V. M. 2025, *A&A*, 694, A252
- Gurzadyan, V. G., Fimin, N. N., & Chechetkin, V. M. 2026, *A&A*, 705, A70
- Hoffman, C., Cheng, J., Ji, D., & Dabaghian, Y. 2023, *PNAS*, 120, e2218245120
- Kolmogorov, A. N. 1933, *G. Ist. Ital. Attuari*, 4, 83
- Lam, M. T., Kaplan, D. L., Agazie, G., et al. 2025, arXiv e-prints [arXiv:2506.03597]
- Luo, J., Ransom, S., Demorest, P., et al. 2019, Astrophysics Source Code Library [record ascl:1902.007]
- Luo, J., Ransom, S., Demorest, P., et al. 2021, *ApJ*, 911, 45
- Sesana, A., & Figuerao, D. G. 2025, arXiv e-prints [arXiv:2512.18822]
- Sesana, A., Vecchio, A., & Colacino, C. N. 2008, *MNRAS*, 390, 192
- Smith, D. A., Abdollahi, S., Ajello, M., et al. 2023, *ApJ*, 958, 191
- Susobhanan, A., Kaplan, D. L., Archibald, A. M., et al. 2024, *ApJ*, 971, 150
- Szapudi, I., Kovacs, A., Granet, B. R., et al. 2015, *MNRAS*, 450, 11
- Tiruvaskar, S., & Gordon, C. 2026, *Phys. Rev.*, 113, 043501