

A statistical analysis of the nulling pulsar population

Sofia Z. Sheikh^{1,2} and Mariah G. MacDonald^{1,2}  ^{1,2}★

¹*Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA*

²*Center for Exoplanets and Habitable Worlds, The Pennsylvania State University, University Park, PA 16802, USA*

Accepted 2021 January 27. Received 2021 January 25; in original form 2020 March 26

ABSTRACT

Approximately 8 per cent of the ~ 2800 known pulsars exhibit ‘nulling,’ a temporary broad-band cessation of normal pulsar emission. Nulling behaviour can be coarsely quantified by the nulling fraction, which describes the percentage of time a given pulsar will be found in a null state. In this paper, we perform the most thorough statistical analysis thus far of the properties of 141 known nulling pulsars. We find weak, non-linear correlations between nulling fraction and pulse width, as well as nulling fraction and spin period which could be attributed to selection effects. We also further investigate a recently hypothesized gap at 40 per cent nulling fraction. While a local minimum does exist in the distribution, we cannot confirm a consistent and unique break in the distribution when we investigate with univariate and multivariate clustering methods, nor can we prove the existence of two statistically distinct populations about this minimum. Using the same methods, we find that nulling pulsars are a statistically different population from normal, radio, non-nulling pulsars, which has never been quantitatively verified. In addition, we summarize the findings of the prior nulling pulsar statistics literature, which are notoriously contradictory. This study, in context, furthers the idea that nulling fraction alone does not contain enough information to describe the behaviour of a nulling pulsar and that other parameters such as null lengths and null randomness, in addition to a better understanding of selection effects, are required to fully understand this phenomenon.

Key words: pulsars: general – stars: statistics – methods: statistical.

1 INTRODUCTION

Pulsar ‘nulling’ (first reported in Backer 1970) is a relatively common phenomenon where otherwise strong and steady radiation from a pulsar decreases by at least a factor of 100,¹ and possibly entirely, over all frequencies (Ritchings 1976; Wang, Manchester & Johnston 2007). These nulls occur for durations of a single pulse period all the way up to days, and affect all components of a pulse simultaneously (e.g. both core and conal components, Rankin 1986; interulses, Hankins 1984).

Nulling pulsars can be identified by looking for gaps in emission in phase-time diagrams (Wang et al. 2007). The nulling fractions (NFs) can then be measured with pulse energy distributions, which will have an overabundance of pulses near zero energy, sometimes leading to a separate well-resolved peak in pulsars with large NF. These overabundances can be measured and reported as a percentage of the total number of pulses that are nulled (Ritchings 1976).

There are some caveats to this method. First, nulling fraction is only one of the parameters that characterizes nulling behaviour. Other parameters, including the duration of a nulling episode and distribution of these durations, are needed to fully describe the nulling behaviour of a pulsar (Gajjar, Joshi & Kramer 2012). Secondly, sensitivity limits make it difficult to identify whether pulsars are entirely dormant during a null, or whether there is some residual low-level emission (Wang et al. 2007).

Some nulling pulsars are reported to show changes in their energetics before, after, or during a pulse (e.g. an increase in pulse energy after a null or a slow-down rate that drops during a null; Backer 1970; Kramer et al. 2006), suggesting that nulls, like other phenomena such as mode-changing,² occur due to changes in a pulsar’s entire magnetospheric environment (Lyne & Ashworth 1982; Lyne et al. 2010; Gajjar, Joshi & Wright 2014a). This, when considered with the simultaneous broad-band cessation during a null, implies a non-uniform flow of particles into the emission region or a global breakdown of the emission mechanism (in a ‘tipping point’ like instability; Wang et al. 2007; Gajjar et al. 2014b), and not changes in relative geometry.

Determining correlations between nulls and other pulsar properties and behaviour allows for characterization and investigations of the pulsar emission mechanism and its evolution over time. Pulsar characteristic age, pulse morphology, and pulse periods have been closely investigated in relation to NF, but there is currently no consensus about their impact on NF. Some studies have reported a positive correlation between characteristic age and nulling fraction (Ritchings 1976; Wang et al. 2007), interpreting this as a sign that pulsars null more and more as they age until they ‘die’ and stop emitting entirely (Ritchings 1976). Other studies find no correlation with age at all (Biggs 1992), or no correlation with age within a morphological group (Rankin 1986). Using classes from Rankin (1983, 1986) suggested that different pulse morphologies lead to different NFs (with triple and multiple profiles having the highest NFs). However,

* E-mail: mmacdonald@psu.edu

¹Factors up to 1420 have been found (e.g. B1706-16; Naidu et al. 2018)

²The likely instability-induced switching from one primary pulse profile to other secondary ones (Wang, Manchester & Johnston 2007)

even using the same classes, Wang et al. (2007a) did not find such a correlation and suggested that the known relationship between complex pulse morphology and older age (Huguenin, Manchester & Taylor 1971) could have caused the previously observed correlated behaviour. Suggestions of a correlation between NF and period (Ritchings 1976; Biggs 1992) are hard to entangle from the previous discussion of correlations with age and morphology, and are also debated in the literature (Rankin 1986).

The statistical methods used in the majority of this previous literature have been preliminary at best. Most studies graphed relevant variables against each other and looked for trends via visual inspection, though they were working with much less data and often admitted the limitations to claiming a correlation. Rankin (1986) calculated means and standard deviations of different pulsar populations to claim that they were distinct. Some studies, specifically Biggs (1992) and Li & Wang (1995), attempted to account for the upper limits and uncertainties on the nulling fraction measurements by using survival analysis methods such as Cox’s proportion hazard model (Cox 1972) and Kendall’s rank correlations (Kendall 1938). Most recently, Konar & Deka (2019) quantified correlations with nulling fraction using Pearson’s R.

Konar & Deka (2019) provided the results of a thorough literature survey for NF values and the most recent statistical survey of nulling pulsar properties. The authors tentatively propose a gap in NF values around 40 per cent and claim that a Kolmogorov–Smirnov test on the spin period of nulling pulsars with NF values greater than and less than 40 per cent imply that there are two distinct populations of pulsars (Konar & Deka 2019). The authors also did not find correlations between NF and any intrinsic pulsar parameters, in contrast with previous studies.

In this paper, we will expand upon the table of pulsar parameters provided by Konar & Deka (2019), investigate correlations between intrinsic pulsar parameters and null fractions, compare the nulling and non-nulling pulsar populations, and statistically characterize and attempt to confirm the two populations around the suggested 40 per cent NF gap from Konar & Deka (2019). In Section 2, we explain how we compile and prepare our data and define all of the parameters that will be used in the statistical analysis described in Section 3. We then explore the existence of the two distinct populations in null fraction suggested by Konar & Deka (2019) in Section 4, and look at trends between nulling fraction and morphological class in Section 5. We provide the results from the correlation analyses between the nulling fraction and between other parameters in the nulling pulsar population in Section 6. In Section 7, we look at the differences between the nulling and non-nulling pulsar populations before discussing our results in Section 8 and concluding in Section 9.

2 DATA

The data used in the analyses in Sections 4–7 are comprised of 15 pulsar parameters and derived quantities. We show 14 of the 15 parameters used in the study in Table 1. We discuss the remaining parameter, morphology, below. We gathered these parameters for a set of 141 pulsars contained in tables 1–5 of Konar & Deka (2019), excluding pulsars in their sample without a measured nulling fraction. Pulsars with 2 or 3 measurements of nulling fraction in the literature have 2 or 3 rows in the table, respectively, for a total of 162 rows.³

³There are actually 163 individual estimates of nulling fractions, but the estimate for B1713-40 from Wang et al. (2007) was accidentally omitted in our analyses.

In our analyses, we use all 162 measurements of nulling fraction, as if each measurement were from a separate pulsar. Table 1 includes a description of each parameter and an explanation of how it was collected.

In addition to the parameters shown in Table 1, we compile Rankin (1983) pulse morphology classes for 141 of the 162 pulsar entries in our sample. 120 of these have been previously classified, and we assigned classifications to 22 additional pulsar entries based on published pulse profiles and analyses. These classifications are: S_d (conal-component single-peaked profile), S_i (core-component single-peaked profile), D (double-peaked profile), T (triple-peaked profile), and M (multiple profile containing more than three peaks in structure). We added an additional letter U (‘undetermined’) for the remaining 20 pulsars and did not consider them in the morphology analysis in Section 5.

3 METHODS

3.1 Cluster analyses

Cluster analyses are a set of methods which can determine how many distinct subsets lie within a single data set, which points belong to each of the subsets, and whether the subsets themselves are statistically distinct. We employ multiple different clustering techniques, described below.

3.1.1 Kernel density estimation

A kernel density estimation (KDE) is a non-parametric estimation of probability density functions, defined by

$$\hat{f}(x; H) = n^{-1} \sum_{i=1}^n K_H(x - X_i), \quad (1)$$

where $x = (x_1, x_2)^T$ and $X_i = (X_{i1}, X_{i2})^T$ for $i = 1, 2, \dots, n$ a random sample drawn from a density f , $K(x)$ is the kernel which is a symmetric probability density function, and H is the positive-definite and symmetric bandwidth matrix. We use an axis-aligned normal kernel. When estimating the location of the gap in nulling fraction, we use a variety of bandwidths from $U[1.0, 100.0]$ (see Section 4 for full analysis).

3.1.2 Jenks’ natural breaks optimization

Jenks’ natural breaks optimization (JNBO) is a univariate class interval method originally developed for the classifying geographic data for cartography (Jenks 1977). Now it is used in many fields beyond geography to find a user-specified number of homogeneous classification groups in 1D data (e.g. Rahadiano, Fariza & Hasim 2015). Once the number of classes has been defined, the JNBO algorithm tries to iteratively minimize the sum of the variances from the class means (Coulson 1987). Jenks–Fisher is a more runtime-efficient version of this same algorithm, incorporating the method put forth in Fisher (1958) (Bivand 2019).

3.1.3 Hierarchical clustering

Hierarchical clustering is a clustering method which does not immediately produce a user-specified number of clusters. Agglomerative algorithms such as `hclust` in R instead look for the two ‘closest’ points (by some linkage metric) and then combine them into a cluster, iteratively repeating until all points are incorporated into a single

Table 1. 14 of the 15 pulsar parameters used in the following statistical analyses. Equations are only given for derived parameters. References marked with * give the equations used for derived parameters.

Name	Symbol	Units	Equation	Data reference	Description
Pulse period	P	s	N/A	Konar & Deka (2019), from Manchester et al. (2005)	Barycentric spin period
Spindown rate	\dot{P}	s/s	N/A	Manchester et al. (2005)	First time derivative of pulse period
Magnetic flux	B_s	G	$B_s \propto (P\dot{P})^{1/2}$	Konar & Deka (2019), from Manchester et al. (2005), Ostriker & Gunn (1969)*	Surface magnetic flux density
Dispersion measure	DM	$\text{cm}^3 \text{pc}^{-1}$	N/A	Manchester et al. (2005)	Indirect measure of pulsar distance from frequency-dependent arrival time delay (e.g. Jokiipii & Lerche 1969)
Characteristic age	τ_c	yr	$\tau_c \sim \frac{P}{2\dot{P}}$	Existing table values	Measure of pulsar age from period and spin-down rate
Kinetic age	τ_k	yr	$\tau_k = \frac{z}{v_z}$	Derived with z and v_z from Manchester et al. (2005), using <code>ASTROPY</code> (Robitaille et al. 2013)	Period-independent pulsar age derived from pulsar's distance from the galactic plane and z -velocity away from the plane [†]
Radio luminosity	L_{R400}	mJy kpc^{-2}	$L_{R400} \propto S_{400} d^2$	Derived with d and S_{400} from Manchester et al. (2005)	Radio luminosity at 400 MHz, quantifies intrinsic pulsar power
Spindown Luminosity	\dot{E}	mJy/kpc^2	$\dot{E} \propto \dot{P} P^{-3}$	Manchester et al. (2005), Gold (1969)*	Spindown energy loss rate
Maximum magnetic flux density	B_{lc}	G	$B_{lc} \propto P\dot{P}^{-5}$	Existing table values, Goldreich (1969), Lyne, Ritchings & Smith (1975)*	Magnetic field in the 'light cylinder' region of the pulsar magnetosphere, where particle velocities are near c
Plasma flow parameter	Q	None	$Q \propto P^{1.1} \dot{P}^{-0.4}$	Existing table values, Beskin, Gurevich & Istomin (1984)*	Derived parameter empirically related to pulsar morphology, with $Q < 1$ generally indicating a single-pulse morphology (Beskin 2009)
Pulse width	W	ms	N/A	Manchester et al. (2005)	Width of the average pulse profile measured at 50 and 10 per cent of the peak brightness
Magnetic inclination angle	α	°	N/A	Malov (1990), Malov & Nikitina (2011), Nikitina & Malov (2017) [‡]	Angle between spin axis of the pulsar and magnetic moment
Nulling fraction	NF	per cent	N/A	tables 1–5 of Konar & Deka (2019)	Percentage of time that a pulsar is in a nulled state

Notes. [†]These data were only available for 96 pulsars in the sample. It should be noted that pulse widths change significantly when measured at different frequencies, or with different time resolutions, and that any use of these data must take into account the large additional uncertainties introduced to a potential correlation.

[‡]We compiled α values (reported as β in some references) from the literature for 107 pulsars in our sample. As noted by Nikitina & Malov (2017), α can be calculated from pulse widths (assuming that the line of sight passes through the centre of the radiation cone as in Rankin 1990), polarization information from the main cone, or position angle and mean profile shape, as in Lyne & Manchester (1988). The reported α values used here are averages from the different methods, where possible. It can be difficult to measure α values with the pulse width method for pulsars with extremely narrow pulse widths (e.g. J1717-4054; Kerr et al. 2014).

cluster (Leisch 1999; Müllner et al. 2013). The results for any number of clusters are thus pre-calculated. The main drawback of hierarchical clustering is its large computational cost, but its ability to use different linkage metrics makes it a particularly flexible solution for clustering (Leisch 1999).

3.1.4 Bagged clustering

Bagged ('bootstrap aggregated') clustering is a clustering method that is both a partitioning method, like k-means clustering (MacQueen et al. 1967), and a hierarchical method (as described in Section 3.1.3) (Leisch 1999). Partitioning methods such as k-means are less effective at recovering structure if the clusters are not convex and, as mentioned in the previous section, hierarchical clustering can

be computationally intensive; bagged clustering algorithms such as `bclust` in R avoid both drawbacks (Leisch 1999).

3.1.5 Geometrical interval classification

Geometrical interval classification classifies data into intervals that have a geometric series. The algorithm forms these intervals by minimizing the sum of the squares of the number of points in each class. Interval classifications also aim for each class range to be approximately the same size and for the change between intervals to be consistent (Jianhua 2002). This algorithm is designed to accommodate continuous data and looks to highlight changes in the data at extreme values as well as at moderate values, making it ideal for heavily skewed and non-normal distributions.

3.2 Kolmogorov–Smirnov and Anderson–Darling tests

The Kolmogorov–Smirnov (K–S) test is a non-parametric test of 1D probability distributions that can be used to compare two samples. The K–S statistic quantifies the distance between the empirical distribution functions of the two samples and is

$$D_{n,m} = \sup |F_{1,n}(x) - F_{2,m}(x)|, \quad (2)$$

where $F_{1,n}$ and $F_{2,m}$ are the empirical distribution functions of the two samples, and \sup is the supremum function. The null hypothesis that the two samples are from the same underlying population is rejected at a level α if

$$D_{n,m} > c(\alpha) \sqrt{\frac{n+m}{nm}},$$

where

$$c(\alpha) = \sqrt{-\frac{1}{2} \ln \left(\frac{\alpha}{2} \right)}.$$

Astronomers have been using the K–S test, for two-samples as well as for goodness of fit, for decades now (e.g. Peacock 1983). This is mainly because it is easy to compute and quite simple to understand, but the test is also distribution-free and can be universally applied without restrictions to sample size. Unfortunately, the K–S test is insensitive when the differences between the empirical distribution functions are not in the middle, but rather at the beginning and the end.

The Anderson–Darling (A–D) test oftentimes tests for normality, or compares a sample’s empirical distribution function to that of another known form (e.g. lognormal, exponential). Using the A–D measure of agreement between distributions, though, it is possible to non-parametrically test a sample, or multiple samples, to see if they come from the same distribution.

The test is based on the distance between the two empirical distribution functions:

$$A^2 = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F(x))^2}{F(x)(1 - F(x))} dF(x). \quad (3)$$

Like the K–S test, the A–D test is both non-parametric and distribution free, but it does not suffer from the same insensitivity issues as the K–S test since it compares the two functions over the entire interval. Also like the K–S test, the null hypothesis states that the two tested samples are from the same underlying population. In comparing two populations, we employ both K–S and A–D tests.

3.3 Pearson’s R and the maximal information coefficient

The Pearson correlation coefficient, also known as ‘Pearson’s R ’ or bivariate correlation, measures the linear correlation between two different variables (Pearson 1895). The coefficient ranges between -1 and 1 , where 0 indicates no correlation, 0.3 – 0.5 indicates a weak correlation, 0.5 – 0.7 indicates a moderate correlation, and greater than 0.7 indicates a strong correlation; negative values of the same ranges indicate anticorrelations. The Pearson correlation coefficient is used widely across scientific fields when analysing data. The coefficient is the covariance of the two variables being tested, divided by the product of their standard deviations. Since we often do not know the entire population and merely study samples of the population, the coefficient is of the sample and calculated by

$$r = r_{xy} = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{(n-1)s_x s_y}, \quad (4)$$

where n is the number of data points in the sample, x_i and y_i are each data point, \bar{x} and \bar{y} are the sample means for x and y , and s_x and s_y are the sample standard deviations for x and y .

Although Pearson’s correlation coefficient is widely used, the coefficient only measures how linear a correlation is; it is therefore sometimes insufficient since two variables can be highly correlated without their correlation being linear. Many correlations in astronomy are power laws or exponential, and Pearson’s correlation coefficient is insensitive in these situations. As such, we also analyse the data by looking at the Maximal Information Criteria (MIC; Reshef et al. 2011). Similar to Pearson’s correlation coefficient, the MIC ranges between 0 and 1 , with 0 indicating no correlation, and higher values corresponding with stronger correlations. The MIC is defined as

$$\text{MIC}(D) = \max_{XY < B(n)} \frac{I^*(D, X, Y)}{\log(\min(X, Y))}, \quad (5)$$

where $D = D(x, y)$ is the set of n ordered pairs of elements of x and y in a space partitioned by an $X \times Y$ grid, grouping the x and y values in X and Y bins, respectively, $B(n) = n^\alpha$ is the search-grid size, with the α parameter defined by the sample size (here, $\alpha = 0.75$; Albanese et al. 2018), and $I^*(D, X, Y)$ is the maximum mutual information over all grids $X \times Y$ of the distribution induced by D on a grid having X and Y bins.

The MIC and other maximal information-based non-parametric exploration methods were introduced by Reshef et al. (2011), who showed that the MIC for a random relationship was approximately 0.18 , with anything near or below that number indicating no correlation. For this work, we do not use an MIC cut-off, but instead rely on p -values generated as described in Section 6.1.

3.4 Managing upper limits and uncertainties

NF is often reported in the literature with associated upper and lower limits. In addition, some pulsars have only been seen to null once, leading to an NF that is only a moving upper limit until the pulsar is observed to null again. These uncertainties and upper limits must be addressed to ensure that any discovered correlations are statistically significant, so we account for them in the following manner. For pulsars with an uncertainty and a nominal value, we draw NF values for each pulsar from a normal distribution centred on the nominal value, with a standard deviation of the published uncertainty. If the published NF is an upper limit, we instead draw the NF from a log uniform distribution of $U[0, NF_u]$, where NF_u is the upper limit of the NF. We chose the log uniform distribution to account for the skewness of the NF distribution. We test the uniform distribution as well to ensure that the choice of distribution is not significantly affecting the results, as discussed in Section 6.1.

We produce an uncertainty-adjusted value for each NF as described above. When exploring correlations between NF and other parameters, we produce 1000 of these sets of corrected NFs and calculate the median of the correlation measurements from each of these 1000 sets. To evaluate the significance of that median, we permute our data as explained in Section 6.1.

4 TWO DIFFERENT POPULATIONS ABOUT A NULLING FRACTION OF ~ 40 PER CENT

4.1 Forming two populations

Recently, Konar & Deka (2019) suggested a gap in nulling fraction (NF) at ~ 40 per cent. There are two separate claims that we would

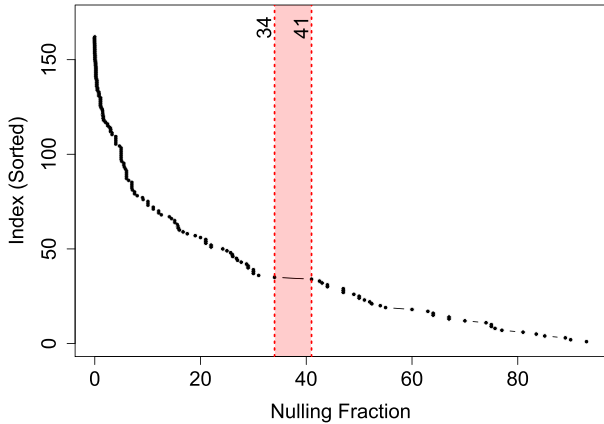


Figure 1. Results of Jenks natural breaks optimization (Sections 3.1.2) with two classes run on the uncorrected nulling fraction distribution of 162 pulsar entries. All methods identified the class breakpoint to be in the gap between 30 per cent and 45 per cent for the uncorrected NF distribution, as well as for 1000 sets of corrected NF, in accordance with Konar & Deka (2019). Bagged clustering and hierarchical clustering also identified a secondary breakpoint at 58 per cent. Geometrical interval classification, which does not assume a gap as the other methods do, concludes that both the corrected and uncorrected NF distributions are continuous and absent of true breaks. This suggests that the gap seen at 40 per cent is an artefact of small sample size.

like to test: (1) that a ‘gap’ or dearth of data exists at 40 per cent and (2) that this gap separates two statistically distinct populations. To confirm and characterize this gap, we use the CRAN package `classInt`⁴ (Bivand 2019) to run a series of univariate class interval algorithms with two classes on 1000 sets of corrected NF (see Section 3.1 for a description of the methods). Hierarchical clustering, Jenks–Fisher, bagged clustering, and Jenks natural breaks optimization methods all identified the natural breakpoint in the data at $47.0 \text{ per cent} \pm 11.2 \text{ per cent}$, $34.4 \text{ per cent} \pm 2.1 \text{ per cent}$, $46.4 \text{ per cent} \pm 12.4 \text{ per cent}$, and $32.7 \text{ per cent} \pm 1.5 \text{ per cent}$, respectively. We show the result of a gap identification using Jenks natural breaks optimization in Fig. 1. Hierarchical clustering and bagged clustering alternate between identifying two different minima in the data: one at 38 per cent and another at 58 per cent, leading to the larger estimates and standard deviations. Geometric interval classification, using the CRAN package `cartography`, did not identify a break or gap near 40 per cent. In fact, the algorithm returned that all 1000 sets of corrected NF were continuous distributions, suggesting that the gap that we see and the breaks that are identified by the other methods are artefacts of an undersampled population.

In order to directly compare to Konar & Deka (2019), we also test this method on the uncorrected NF, recovering gaps at 37 per cent for each method, except geometric interval classification which again does not identify any breaks in the data.

We then estimate the distribution of NF using a 1D kernel density estimator (KDE), using the uncorrected NF distribution. Using 100 bandwidths drawn from a uniform distribution $U[1.0, 100.0]$,⁵ we confirm a minimum in the KDE of 37.23 ± 0.19 . We draw multiple bandwidths such that the location of the minimum is not dependent on our bandwidth.

⁴All statistics in this paper were performed using R 3.4.1 (R Core Team 2019).

⁵Bandwidths drawn from $U[1.0, 100.0]$ span the entire range of NF

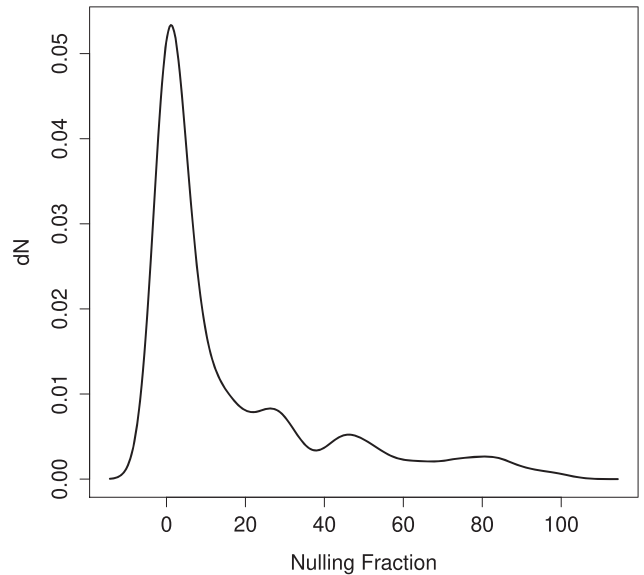


Figure 2. KDE of 16000 samples of corrected nulling fraction, with a bandwidth of 15.0. There does exist a local minimum in the distribution at around 40 per cent, as noted by Konar & Deka (2019), but this minimum is accompanied by two other local minima, at 20 per cent and at 60 per cent. This suggests that the gap seen in the data is due to small sample size.

We then perform a KDE on 16000 samples of the corrected NF, again using 100 bandwidths drawn from a uniform distribution $U[1.0, 100.0]$. Regardless of bandwidth, there does exist a minimum in the KDE around 37 per cent, but there also exist minima at ~ 20 per cent and at ~ 60 per cent (see Fig. 2). We thus find that the gap seen by Konar & Deka (2019) is not a boundary between two distinct populations, but is simply an artefact of the uncertainties on NF estimates and an undersampled population. Once uncertainties are accounted for, there does appear to be a local minimum at around 40 per cent, but other minima exist, suggesting that this minimum is solely due to the small sample size.

4.2 Confirming two populations

We run both K–S and A–D two-sample tests (see Section 3.2) on 1000 sets of corrected nulling fraction measurements. We define our two populations as pulsars with $NF < 37.5$ per cent and pulsars with $NF > 37.5$ per cent. We compare the two populations in 13 parameters ($P, \dot{P}, B_s, DM, \log \tau_c, \log \tau_K, L_R, \dot{E}, B_{lc}, Q, W50, W10,$ and α) and report the median p -values from the 1000 sets in Table 2.

Given the large p -values resulting from both of these tests, we fail to reject the null hypothesis that nulling pulsars above and below a nulling fraction of ~ 40 per cent are drawn from the same population. We do find that the two samples are distinct ($p < 0.05$) in DM and α and potentially distinct in $\dot{E}, B_{lc},$ and Q . Although these factors indicate separate populations, the other parameters, such as spin period and pulse width, have large p -values even though they are correlated with nulling fraction (see Section 6). This indicates that the apparent difference in populations is an artefact of selection bias or a noise limit. We additionally run these two-sample tests on the uncorrected NF measurements. We find that the two uncorrected samples are distinct in $DM, \log \tau_c, L_{R400}, \dot{E}, Q,$ and α , but again not distinct in the more characteristic parameters. We therefore conclude that the two sets of pulsars are drawn from the same population even though they are statistically distinct in some parameters.

Table 2. P values resulting from Kolmogorov–Smirnov (K–S) and Anderson–Darling (A–D) two-sample tests comparing the proposed populations of pulsars with $NF < 37.5$ per cent and pulsars with $NF > 37.5$ per cent for spin period in seconds (P), magnetic field in Gauss (B_s), spin-down rate in seconds per second (\dot{P}), dispersion measure (DM), characteristic age (τ_c), kinematic age (τ_k), radio luminosity at 400 MHz (L_{R400}), spin-down luminosity (\dot{E}), pulse widths at 50 per cent (W50) and 10 per cent (W10), and magnetic inclination angle α in 162 nulling pulsar entries with measured values of NF (per cent) taken from Konar & Deka (2019). We run these tests with 1000 sets of corrected NF measurements and report the median p -value. Given the above values, we fail to reject the null hypothesis that the two samples are drawn from the same population.

Parameter	K–S	A–D
P	0.18	0.066
B_s	0.83	0.65
\dot{P}	0.89	0.74
DM	0.018	0.012
$\log \tau_c$	0.049	0.11
$\log \tau_k$	0.52	0.57
L_{R400}	0.18	0.072
\dot{E}	0.10	0.042
B_{lc}	0.20	0.022
Q	0.10	0.023
W50	0.44	0.29
W10	0.31	0.062
α	0.036	0.026

5 AN ANALYSIS OF MORPHOLOGICAL CLASSIFICATIONS

We divide our sample into the morphological classifications of pulse profile (Rankin 1983). Using K–S and A–D tests, we investigate whether the populations are statistically distinct in nulling fraction and characteristic age, both of which have been claimed in the literature (e.g. Hesse & Wielebinski 1974; Rankin 1986).

We find that the relation between NF and morphological class is equivocal at best. Pulsars with double profiles (class ‘D’) are a statistically distinct population in NF when compared to all other classes, with p -values < 0.05 , except multiple profiles (class ‘M’), but no other classes are statistically distinct from each other.

The relation between characteristic age and morphological class is extant in our sample, but subtle. We find that S_r nulling pulsars are the youngest morphological class by median and mean, followed by T , then S_d , then D, then M, which is consistent with the age-morphology findings in Rankin (1986). However, in characteristic age, we cannot prove that neighbouring classifications are distinct (e.g. S_r and T), but any given class is statistically distinct from all non-neighbouring classes (e.g. S_r is distinct from S_d , D, and M). The gradient of age across profile classes can be seen in Fig. 3, but neighbouring CDFs are too similar to be distinguished by the K–S and A–D tests.

6 CORRELATIONS BETWEEN PULSAR PARAMETERS WITHIN THE NULLING POPULATION

6.1 Searching for correlations with null fraction

We calculate the Pearson- R correlation value and the maximal information coefficient (MIC) between nulling fraction and each of the fifteen parameters described in Section 2.

We determine the significance of the MIC for a parameter x by permuting the NF and x values within 1000 bootstrapped subsamples

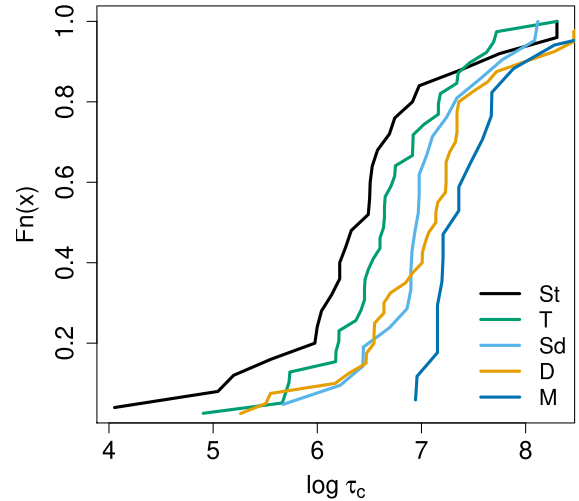


Figure 3. Cumulative distribution functions of characteristic age for the five morphological classifications S_r , T , S_d , D , and M . We find classes to be distinct from other classifications, but cannot prove that neighbouring classifications (e.g. S_r and T) are distinct from one another via K–S or A–D tests.

of the data set. We then calculate the p -value as the percentage of permuted outcomes that meet or exceed the median MIC value of the unscrambled and uncertainty-adjusted (corrected NF) data (see Section 3.4). We consider correlations with p -values less than 0.05 as statistically significant.

In the past, only linear correlations have been investigated for nulling pulsars, but in this paper we are also able to investigate non-linear correlations by using the MIC (Section 3.3). The difference between the two correlation measures (MIC – R^2) describes the non-linearity of a correlation.

We find no correlations shown by Pearson- R between NF and any other variable. However, MIC reveals weak non-linear correlations between NF and: P (MIC = 0.328, p -value = 0.009), \dot{E} (MIC = 0.334, p -value = 0.005), B_{lc} (MIC = 0.326, p -value = 0.007), Q (MIC = 0.306, p -value = 0.012), and W10 (MIC = 0.337, p -value = 0.003).

The correlations with \dot{E} , B_{lc} , and Q are not particularly interesting because all three parameters probably take their correlation from their heavy P -dependence. However, the weak non-linear trends with P and pulse width align with reports in the previous literature (Biggs 1992; Li & Wang 1995) and could point to fundamental physical nulling behaviour. We show the median NF values (for those with upper limits) in Figs 4 and 5. The measured trends might be driven by selection effects, as illustrated by the formation of very similar distributions by the nulling pulsar population and randomly scrambled samples from the non-nulling pulsar population in Figs 4 and 5. We therefore conclude that, while these correlations are statistically significant, they are not necessarily scientifically significant and caution against drawing conclusions from them until we have more data. In the meantime, future theories of pulsar nulling should account for these trends with period and pulse width.

To determine whether our method for accounting for the uncertainties was affected by the choice of distribution, we also perform the same test using a uniform distribution for the uncertainty adjustment instead of a log uniform distribution for NFs reported as an upper limit. We find that two additional correlations, with L_{R400} and W50, are only significant in the uniform and log uniform distributions, respectively. Because these correlations are dependent

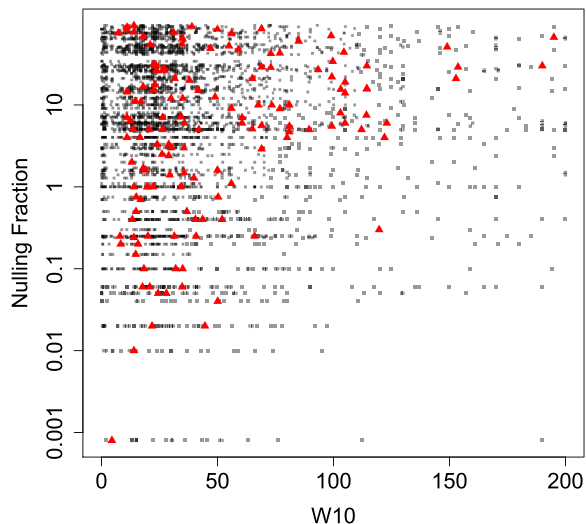


Figure 4. Nulling fraction versus the pulse width (in milliseconds) for 162 nulling pulsar measurements in red triangles. Nulling fractions that have only been measured as upper limits are displayed as this upper limit. For context, we plot the periods of the non-nulling populations against randomly assigned nulling fractions from the nulling population in semitransparent grey squares to show the expected density with no correlation. Given that the randomly sampled non-nulling pulsar population shows a similar distribution to the nulling pulsar distribution, the correlation is most likely due to selection effects. However, the paucity of points in the lower right quadrant is different between the two distributions and could be suggestive of a scientifically interesting trend.

on the uncertainties, and there is no clear scientific basis to choose one distribution over the other, we do not report these parameters as correlated as to NF.

We find numerous correlations between NF and the other parameters within morphological classes. However, as the sample sizes are extremely small (13–36 pulsars), we note only that all classes uniformly show statistically significant correlations with the parameters found for the entire population: P , \dot{E} , Q , $W10$, and $W50$, with the exception of B_{lc} which is correlated with NF in all classes except D. We also recover two potentially interesting correlations not seen in the overall population: B_s and \dot{P} are non-linearly correlated with NF in all classes except S_s . However, because B_s is derived from P and \dot{P} , this correlation is probably just reflective of a trend with \dot{P} .

6.2 Searching for correlations between other parameters

We calculate the Pearson- R correlation value and the MIC for every combination of the fifteen parameters described in Section 2, with the exception of morphology.

Many correlations between pulsar parameters within the nulling pulsar data set are expected because the quantities are derived from one another. For example, we expect to see (and do see) a correlation between period P and characteristic age τ_c . This class of derived correlations accounts for the majority of correlations that we find, and while scientifically unoriginal, is still valuable to confirm our methods.

We find a weak but significant non-linear correlation between DM and surface magnetic field B_s (MIC = 0.372, p -value = 4×10^{-4}), however we do not see this correlation in the overall non-nulling population. There is a paucity of pulsars with low surface magnetic fields at large DMs, suggesting that the correlation we find is due

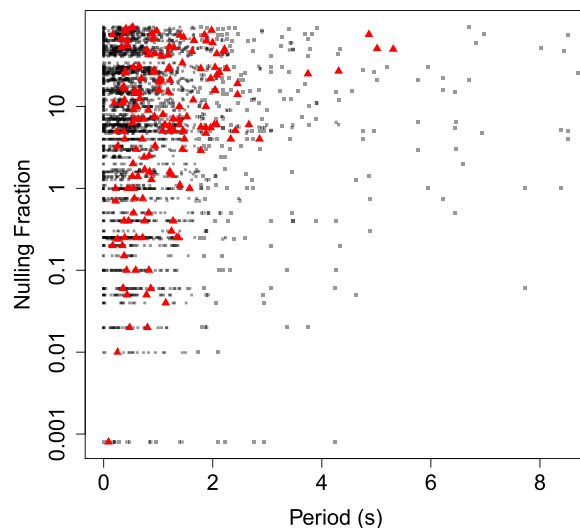


Figure 5. Nulling fraction versus the pulse period for 162 nulling pulsar measurements in red triangles. Nulling fractions that have only been measured as upper limits are displayed as this upper limit. For context, we plot the periods of the non-nulling populations against randomly assigned nulling fractions from the nulling population in semitransparent grey squares to show the expected density with no correlation. Since the randomly sampled and scrambled non-nulling pulsar population shows a similar distribution to the nulling pulsar distribution, the correlation is most likely due to selection effects.

to a detection bias against distant, low energy pulsars. The nulling population is most likely more sensitive to this bias because nulling pulsars must be observed for a longer span of time to reach the same signal-to-noise ratio as non-nulling pulsars during pulsar discovery surveys.

There is also a weak but significant non-linear correlation between spin-down rate \dot{P} and width of the pulse $W10$ (MIC = 0.318, p -value = 0.006), a correlation that we also see in the non-nulling population. This relationship has been previously reported within alternating spin-down states in individual pulsars (e.g. Lyne et al. 2010; Perera et al. 2016).

7 THE NULLING VERSUS NON-NULLING POPULATION

To verify that nulling and normal, radio pulsars are two distinct populations, we perform K–S and A–D two-sample tests, comparing 162 nulling pulsar entries (see Section 2) with 2307 radio pulsars in nine parameters that are available via ATNF (Manchester et al. 2005). We present the resulting p -values from both tests in Table 3. We find that p -values from all parameters except $W10$ and $W50$ are less than $\alpha = 0.05$ and are distinct from one another. We therefore reject the null hypothesis and confirm that nulling pulsars are not a subset of normal, radio pulsars but are instead a separate population. We compare the cumulative distribution functions for each parameter between the two populations in Fig. 6.

8 DISCUSSION

To put this work in context, we compile the correlations that have been claimed or suggested by all nulling pulsar statistical studies, including this one, in Table 4. Table 4 highlights the lack of a consensus in the literature about correlations between nulling

Table 3. Resulting p -values from Kolmogorov–Smirnov (K–S) and Anderson–Darling (A–D) two-sample tests. Here, P is spin period, \dot{P} is spin-down rate, B_s is the surface magnetic field strength, DM is the dispersion measure, τ_c is characteristic age, L_R is the radio luminosity at 400 MHz, \dot{E} is the spin-down luminosity, $W50$ is the pulse width at 50 per cent peak brightness, and $W10$ is the pulse width at 10 per cent brightness. Given our small p -values, we reject the null hypothesis that these two samples are from the same population, even though we cannot claim they are distinct in $W10$ or $W50$. We instead conclude that nulling pulsars are not a subset of radio pulsars and are a distinct and separate population. See Fig. 6 for the cumulative distribution functions of the two populations in all nine parameters.

Parameter	No. Nulling	No. Radio	K–S	A–D
P	162	2307	1.313e-05	4.196e-09
\dot{P}	162	1958	0.03879	0.02330
B_s	162	1958	0.007411	0.005287
DM	162	2306	4.408e-14	7.058e-23
τ_c	162	1958	0.002137	0.001875
L_R	131	700	0.0293	0.04452
\dot{E}	161	1958	0.0002043	1.714e-06
$W50$	158	1971	0.0363	0.09156
$W10$	145	1154	0.1802	0.08264

fraction (NF) and intrinsic pulsar parameters. We are the only study to have searched for correlations in every previously suggested parameter (see Section 2 for a more thorough description of the included parameters). Only recently have enough nulling pulsars been discovered to draw statistically significant conclusions about the population – this can be seen in the chronological ordering of Table 4.

In Section 4.2, we conclude that pulsars with NF below 37.5 per cent and pulsars with NF above 37.5 per cent are drawn from the same underlying population. In some variables, however, we find the two populations to be distinct such as in DM. DM is dependent on distance which reveals a potential observational bias: pulsars that null for longer are less likely to appear in e.g. a drift scan survey, so we would expect discovered pulsars with higher nulling fractions to be those that are brighter and/or closer to Earth. To confirm that our results are not being affected by nulling pulsars with multiple NF measurements,⁶ we re-run our analysis a few times by randomly choosing a singular measurement of NF for each pulsar and by omitting these pulsars entirely. We find this does not change our results.

In Fig. 7, we show a $P - \dot{P}$ plot for all of the pulsars in the ATNF database (Manchester et al. 2005), with nulling pulsars highlighted in red triangles. The ‘tail’ in the plot towards the bottom left-hand corner is comprised of the millisecond pulsar population, which has extremely short spin periods and very small \dot{P} values. Rajwade et al. (in preparation) suggested that millisecond pulsars might not null at all. The current nulling pulsar population, as shown in Fig. 7 supports this conclusion.

Given the non-linear correlation between NF and P that we found in this work, the absence of nulling millisecond pulsars points to a fundamental relationship between spin period and nulling behaviour. This motivated the decision to remove both millisecond pulsars and all other non-normal pulsars from the comparison between nulling and non-nulling pulsars in Section 7.

⁶Our sample contains 141 individual pulsars, but numerous pulsars have different estimates of nulling fraction. We therefore typically treat each individual nulling fraction measurement as a different pulsar.

To ensure that the uncertainties resulting from pulse widths measured at different frequencies are not obscuring our results, we perform the following. First, we find that our pulse widths were measured at the following frequencies: roughly one third of the measurements (both $W10$ and $W50$) were made at 1400 MHz, one third at 408 MHz, and the rest between those frequencies, with only a single outlier at 5000 MHz (the $W10$ value for B1641-45). This factor of 3.5 in observing frequency translates to differences in measured pulse widths. Previous literature shows that, because observing frequency is related to pulse width by a power law, differences in observing frequency of factors <10 do not greatly affect pulse widths in most cases (e.g. D’Amico et al. 1998; Chen & Wang 2014). Rankin (1983) found that a change of pulse width by a factor of 2 requires multiple orders-of-magnitude changes in observing frequency. To prove that this potential source of error does not affect our results, we add in errors of 60 per cent⁷ and perform a bootstrapped analysis. We find that, while the strength of the correlation decreases in some of our trials, our conclusions do not change. We still find a non-linear, weak, but statistically significant correlation between pulse width and NF.

While we searched for correlations between NF and other parameters within pulse profile morphological classes in Section 6.2, we were limited by small-number statistics. In order to determine whether nulling behaviour depends on morphological class, a much larger nulling population will be needed. One indirect way to increase the sample of nulling pulsars is to acknowledge the link between mode-changing behaviour and nulling behaviour (e.g. Timokhin 2010). The fraction of time that a mode-changing pulsar spends in a less-energetic mode can be used as a de facto nulling fraction.

Selection effects influence the population of nulling pulsars more than the overall pulsar population, as discussed in Biggs (1992). For example, an absence of emission from a pulsar is easier to detect if the pulsar has a higher flux density (i.e. is nearer to Earth or more luminous). The issue is complicated by an additional selection bias due to lack of sensitivity towards single pulses. For the majority of pulsars, multiple pulses must be integrated in order to recover a pulse profile. Single pulse nulls can thus easily be missed. Nearer pulsars will have lower DMs, explaining the centre left plot in Fig. 6. Thus it is hard to characterize intrinsic differences between nulling and non-nulling pulsar parameters until the selection effects have been characterized. This could perhaps be accomplished via simulated injection-recovery testing.

Nulling fraction is only a single, and admittedly the crudest, measure of pulsar nulling behaviour. Gajjar (2017) suggested that the lack of a consistent trend of correlation between pulsar parameters and NF was indicative of NF being a poor metric rather than any inherent behavioural cause. Nulling behaviour can also be characterized with nulling length, which has been reported to be correlated with age (Lazaridis & Seiradakis 2006). Yang, Han & Wang (2014) defines more complex parameters to quantify nulling that relate the lengths of the on and off states, and finds that spin-down rate and spin-down luminosity are correlated with these parameters. Nulling randomness is yet another way to parametrize nulling behaviour (Gajjar 2017). Polarization parameters and timing noise, as of yet uncompiled for these pulsars, could also be important to characterizing nulling behaviour. Future statistical studies should incorporate more nulling-related parameters once the data are available to do so.

⁷60 per cent corresponds to the largest errors from Chen & Wang (2014), assuming pulse widths varied by a factor of 10, and ~ 2.5 greater than we observe in our sample

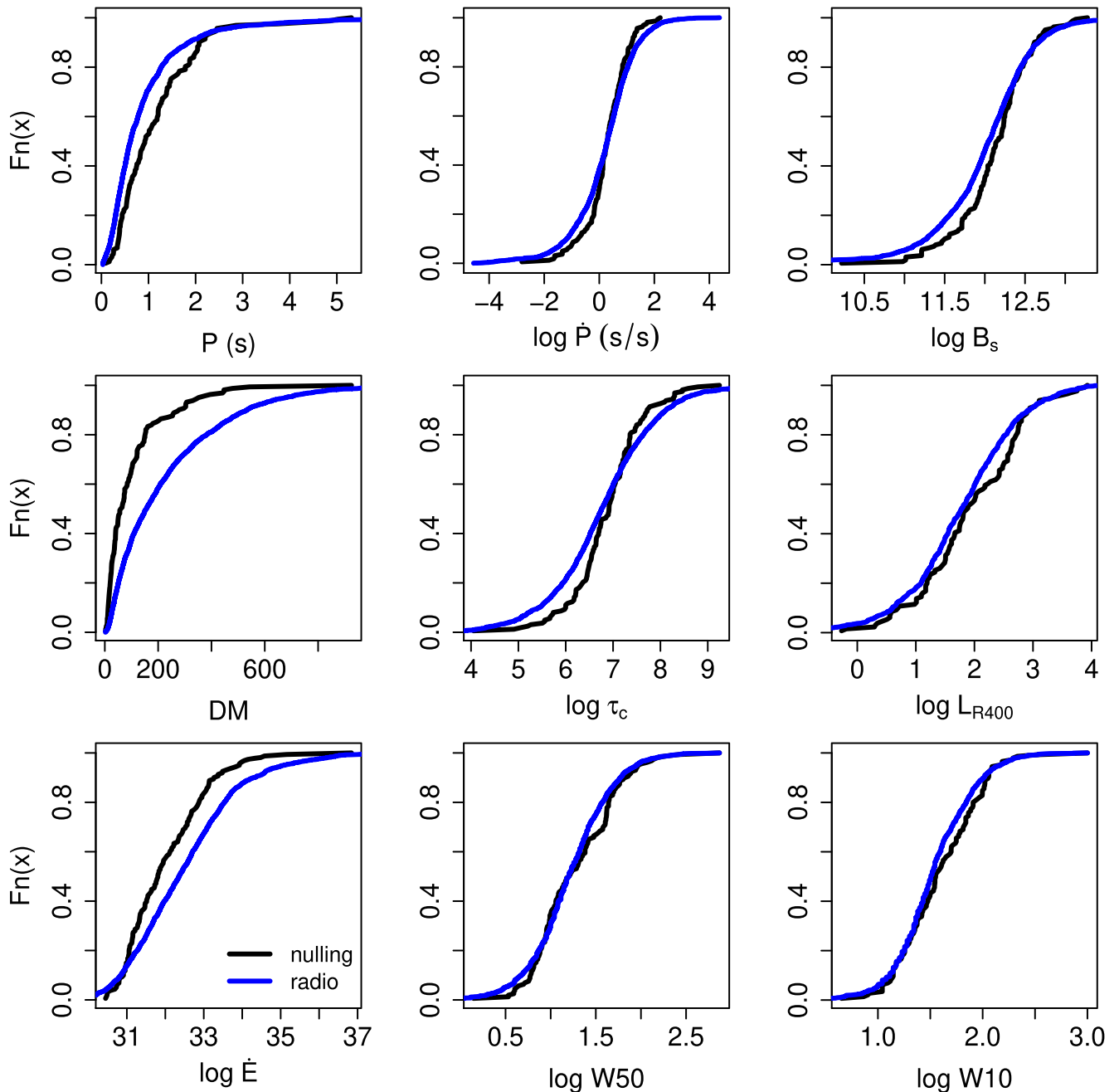


Figure 6. Cumulative distribution functions comparing nulling pulsars in black to normal, radio pulsars in blue. Here, P is spin period, \dot{P} is spin-down rate, B_s is the surface magnetic field strength, DM is the dispersion measure, τ_c is characteristic age, L_{R400} is the radio luminosity at 400 MHz, \dot{E} is the spin-down luminosity, $W50$ is the pulse width at 50 per cent brightness, and $W10$ is the pulse width at 10 per cent brightness. We conclude that these two populations are indeed two populations and not samples from a singular population.

9 SUMMARY AND CONCLUSION

About a tenth of all known pulsars temporarily cease their normal emission in a phenomenon known as ‘nulling.’ Here, we investigate this population of 141 pulsars, using a variety of techniques described in Section 3. We first explore whether there is a gap in the nulling fraction (NF) distribution around 40 per cent, as suggested by Konar & Deka (2019), with cluster analyses. Then we determine if this gap does indeed separate two distinct pulsar populations. Using both K–S and A–D two-sample tests, we then investigate the statistical

distinctness of different morphological classes within the overall nulling population. Following the suggestion in past studies that NF is correlated with a variety of pulsar parameters, we test for such correlations between fundamental pulsar parameters (see Section 2) and NF using both the MIC and Pearson’s R . We also investigate correlations between all of the parameters in the nulling pulsar population, independent of their behaviour with NF. Lastly, we explore whether nulling pulsars and non-nulling pulsars are distinct populations. Following these analyses, we conclude the following:

Table 4. Pulsar parameters studied for correlation with nulling fraction in the literature, along with the results of this work. P is spin period, \dot{P} is spin-down rate, B_s is the surface magnetic field strength, τ_C is characteristic age, α is the inclination angle between the magnetic and spin axes, L_R is the radio luminosity at 400 MHz, \dot{E} is the spin-down luminosity, B_{lc} is the magnetic field strength at the light cylinder, Q is a dimensionless plasma flow parameter from Beskin, Gurevich & Istomin (1988), M is a categorical variable representing the morphological class of the pulse profile (as per Rankin 1983), τ_k is the kinetic age of the pulsar, and W_p is the total pulse width. Cells with an ‘X’ indicate that the authors investigated that particular variable in their study and found no correlation with nulling fraction. The shaded cells with a checkmark indicate that a study suggested a correlation between the given variable and nulling fraction. The unshaded cells indicate that a study did not use that particular parameter. We investigated correlations with MIC and R^2 and found that the five correlations that we found were non-linear. Morphology is a categorical variable so we used K–S and A–D tests to determine the association between morphology and nulling fraction. Li & Wang (1995) used equivalent width for the pulse width parameter, while we saw a correlation with W10 but not W50.

Study	Sample size	P	\dot{P}	B_s	τ_C	α	L_R	\dot{E}	B_{lc}	Q	M	τ_k	W_p	DM
Hesse & Wielebinski (1974)	15	–	–	✓	–	–	–	–	–	–	✓	–	–	–
Ritchings (1976)	32	✓	X	–	✓	–	–	–	✓	–	–	–	–	–
Zhen-ru & Yi (1981)	32	✓	–	✓	–	–	–	–	✓	–	–	–	–	–
Rankin (1986)	59	X	X	X	X	–	–	–	X	–	✓	–	–	–
Biggs (1992)	72	✓	–	✓	✓	✓	✓	✓	✓	✓	✓	–	–	–
Li & Wang (1995)	72	–	–	–	–	–	–	–	–	–	–	–	✓	–
Lazaridis & Seiradakis (2006)	19	–	–	–	✓	–	–	–	–	–	–	–	–	–
Wang et al. (2007)	64†	X	–	–	✓	–	–	–	–	–	X	–	–	–
Cordes & Shannon (2008)	69†	–	–	–	–	✓	–	–	–	–	–	–	–	–
Konar & Deka (2019)	204*	–	X	X	–	–	–	–	–	–	–	–	–	–
This Work	162	✓	X	X	X	X	X	✓	✓	✓	X	X	✓	X

Notes. †The authors did not report the number of nulling pulsars, so we estimated the number from provided figures.

* This study also included pulsars that were known to null but did not have an estimated nulling fraction. Only 162 values were used in the correlation analyses.

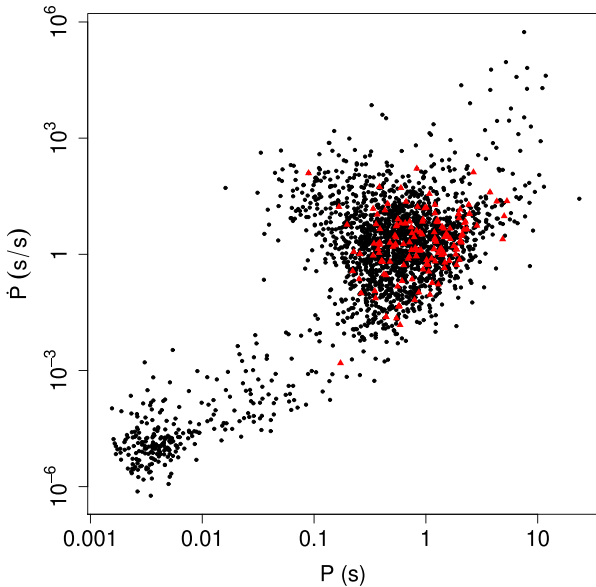


Figure 7. Period versus period derivative plot for non-nulling pulsars (black dots) and nulling pulsars (red triangles). The absence of red triangles in the bottom left of the plot indicates that there are no known millisecond pulsars that null.

(i) There is a local minimum in the nulling fraction of the nulling pulsar population at about 40 per cent, but we cannot identify a natural breakpoint, and this minimum does not divide two statistically distinct populations (Section 4).

(ii) In terms of nulling fraction, pulsars in morphological class D are statistically distinct from all other classes except M. No other classes are statistically distinct from one another (Section 5).

(iii) Morphological class is related to age in the nulling pulsar population, with the classes ordered, youngest to oldest, as S_r , T, S_d , D, M, which is consistent with Rankin (1986). This age gradient

is statistically distinguishable, but also subtle with large scatter (Section 5).

(iv) We find nulling fraction in the nulling pulsar population to be weakly correlated with pulse period P and full pulse width W_{10} , as well as with parameters that are derived from P (\dot{E} , B_{lc} , Q), however these trends should be taken with caution because of strong selection effects (Section 6.1).

(v) We find dispersion measure DM and surface magnetic field B_s to be weakly, non-linearly correlated in the nulling pulsar population but not in the non-nulling population, likely due to selection bias (Section 6.2).

(vi) We find spin-down rate \dot{P} and pulse width W_{10} to be weakly, non-linearly correlated in the nulling pulsar population as well as in the non-nulling population, consistent with previous studies (Section 6.2).

(vii) Nulling and normal, radio pulsar populations are statistically distinct in nine fundamental parameters: period, spin-down rate, surface magnetic field, dispersion measure, characteristic age, radio luminosity, spin-down luminosity, and 10 per cent pulse width and 50 per cent pulse width (Section 7).

This list of statistical outcomes might feel disjointed and detached from the physical sample, but each result gives us a clue about either (1) the fundamental physics behind the pulsar nulling phenomenon or (2) the biases that must be accounted for when such analyses are performed.

Future work must thoroughly investigate the selection effects present in the nulling pulsar sample. This can be achieved via creating synthetic populations and simulating detection rates for pulsars with different nulling fractions, dispersion measures, radio luminosities, etc. for a range of instruments and survey parameters. A study of this magnitude would allow us to disentangle the connection between nulling fraction and inherent pulsar properties from the effect of detection biases. As always, over time, the nulling pulsar sample will become better characterized by the acquisition of polarizations, nulling lengths, null distributions, timing noise, and other data, which

will lead to a better understanding of the emission mechanism and its relationship to nulling behaviour.

ACKNOWLEDGEMENTS

We would like to thank the referee for the helpful comments and suggestions that have improved this manuscript. We would like to thank Vishal Gajjar, Alex Wolszczan, and Jason Wright for the helpful discussions. MGM acknowledges that this material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE1255832. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation. We acknowledge use of the ATNF catalogue at <http://www.atnf.csiro.au/research/pulsar/psrcat/>. The Center for Exoplanets and Habitable Worlds is supported by the Pennsylvania State University, the Eberly College of Science, and the Pennsylvania Space Grant Consortium.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

- Albanese D., Riccadonna S., Donati C., Franceschi P., 2018, *GigaScience*, 7, g1y032
- Backer D., 1970, *Nature*, 228, 42
- Beskin V. S., 2009, *MHD Flows in Compact Astrophysical Objects: Accretion, Winds and Jets*. Springer Science & Business Media, Berlin
- Beskin V., Gurevich A., Istomin Y. N., 1984, *Astrophys. Space Sci.*, 102, 301
- Beskin V. S., Gurevich A. V., Istomin Y. N., 1988, *Astrophys. Space Sci.*, 146, 205
- Biggs J. D., 1992, *ApJ*, 394, 574
- Bivand R., 2019, classInt: Choose Univariate Class Intervals. Available at: <https://CRAN.R-project.org/package=classInt>
- Chen J. L., Wang H. G., 2014, *ApJS*, 215, 11
- Cordes J. M., Shannon R., 2008, *ApJ*, 682, 1152
- Coulson M. R., 1987, *Cartographica*, 24, 16
- Cox D. R., 1972, *J. R. Stat. Soc. B*, 34, 187
- D'Amico N., Stappers B. W., Bailes M., Martin C. E., Bell J. F., Lyne A. G., Manchester R. N., 1998, *MNRAS*, 297, 28
- Fisher W. D., 1958, *J. Am. Stat. Assoc.*, 53, 789
- Gajjar V., 2017, preprint ([arXiv:1706.05407](https://arxiv.org/abs/1706.05407))
- Gajjar V., Joshi B. C., Kramer M., 2012, *MNRAS*, 424, 1197
- Gajjar V., Joshi B. C., Wright G., 2014a, *MNRAS*, 439, 221
- Gajjar V., Joshi B. C., Kramer M., Karuppusamy R., Smits R., 2014b, *ApJ*, 797, 18
- Gold T., 1969, *Nature*, 221, 25
- Goldreich P., 1969, *PASA*, 1, 227
- Hankins T., 1984, in *Bulletin of the American Astronomical Society*. p. 468
- Hesse K., Wielebinski R., 1974, *A&A*, 31, 409
- Huguenin G., Manchester R., Taylor J., 1971, *ApJ*, 169, 97
- Jenks G. F., 1977, Department of Geography, University of Kansas Occasional Paper

- Jianhua X., 2002, *Mathematical Methods in Contemporary Geography*, 2nd edition, Higher Education Press, Beijing, p. 37
- Jokipii J., Lerche I., 1969, *ApJ*, 157, 1137
- Kendall M. G., 1938, *Biometrika*, 30, 81
- Kerr M., Hobbs G., Shannon R. M., Kiczynski M., Hollow R., Johnston S., 2014, *MNRAS*, 445, 320
- Konar S., Deka U., 2019, *J. Astrophys. Astron.*, 40, 42
- Kramer M., Lyne A. G., O'Brien J. T., Jordan C. A., Lorimer D. R., 2006, *Science*, 312, 549
- Lazaridis K., Seiradakis J. H., 2006, in Solomos N. H., ed., *AIP Conference Proceedings*, p. 309
- Leisch F., 1999, Technical Report 51, Bagged clustering. Vienna University of Economics and Business Administration
- Li X., Wang Z., 1995, *Chin. Astron. Astrophys.*, 19, 302
- Lyne A. G., Ashworth M., 1982, *MNRAS*, 204, 519
- Lyne A., Manchester R., 1988, *MNRAS*, 234, 477
- Lyne A., Ritchings R., Smith F., 1975, *MNRAS*, 171, 579
- Lyne A., Hobbs G., Kramer M., Stairs I., Stappers B., 2010, *Science*, 329, 408
- MacQueen J. et al., 1967, *Proc. fifth Berkeley Symp. Math. Stat. Prob.*, 1, 281
- Malov I. F., 1990, *Astron. Z.*, 67, 377
- Malov I., Nikitina E., 2011, *Astron. Rep.*, 55, 19
- Manchester R. N., Hobbs G. B., Teoh A., Hobbs M., 2005, *AJ*, 129, 1993
- Müllner D. et al., 2013, *J. Stat. Softw.*, 53, 1
- Naidu A., Joshi B. C., Manoharan P., Krishnakumar M., 2018, *MNRAS*, 475, 2375
- Nikitina E. B., Malov I. F., 2017, *Astron. Rep.*, 61, 591
- Ostriker J., Gunn J., 1969, *ApJ*, 157, 1395
- Peacock J., 1983, *MNRAS*, 202, 615
- Pearson K., 1895, *Proc. R. Soc.*, 58, 240
- Perera B. B., Stappers B. W., Weltevrede P., Lyne A. G., Rankin J. M., 2016, *MNRAS*, 455, 1071
- R Core Team, 2019, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: <https://www.R-project.org/>
- Rahadianto H., Fariza A., Hasim J. A. N., 2015, in *2015 International Conference on Data and Software Engineering (ICoDSE)*, IEEE, Yogyakarta, Indonesia, p. 195
- Rankin J. M., 1983, *Toward an Empirical Theory of Pulsar Emission. I. Morphological Taxonomy*. Available at: <https://patentimages.storage.googleapis.com/41/b5/ca/69ffeea861af61/US8949899.pdf>
- Rankin J. M., 1983, *ApJ*, 274, 359
- Rankin J. M., 1986, *ApJ*, 301, 901
- Rankin J. M., 1990, *ApJ*, 352, 247
- Reshef D. N. et al., 2011, *Science*, 334, 1518
- Ritchings R., 1976, *MNRAS*, 176, 249
- Robitaille T. P. et al., 2013, *A&A*, 558, A33
- Timokhin A. N., 2010, *A Model for Nulling and Mode Changing in Pulsars*. Available at: <https://academic.oup.com/mnrasl/article-lookup/doi/10.1111/j.1745-3933.2010.00924.x>
- Wang N., Manchester R., Johnston S., 2007, *MNRAS*, 377, 1383
- Yang A., Han J., Wang N., 2014, *Sci. China: Phys. Mech. Astron.*, 57, 1600
- Zhen-ru W., Yi C., 1981, in *Symp. – Int. Astron. Un., Vol. 95*. Cambridge University Press, p. 215

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.