The Metrewavelength Sky
ASI Conference Series, 2014, Vol. 13, pp 79 – 81
Edited by J. N. Chengalur & Y. Gupta

Inferring pulsar null fraction using Gaussian mixtures

Mihir Arjunwadkar[1*], Kaustubh Rajwade[2] and Yashwant Gupta[3]
[1]*Centre for Modeling and Simulation, University of Pune, Pune 411007 India*
[2]*West Virginia University, Morgantown, WV 26506 USA*
[3]*National Centre for Radio Astrophysics, University of Pune Campus, Pune 411007 India*

Abstract. We present a Gaussian-mixture-based rigorous statistical methodology for estimating the null fraction of a pulsar, and validate it against a number of data sets from well-characterized regular pulsars.

Keywords : pulsar nulling – null fraction – Gaussian mixture models

1. Introduction

Many pulsars exhibit *nulling*, where their pulsed radio emission ceases abruptly for one or several consecutive pulses. The fraction NF of absent (or null) pulses is estimated using pulse energy data. The null-pulse energy distribution is assumed to have the same shape as the off-pulse receiver noise, assumed Gaussian. The distribution of non-null pulse energies can, however, assume a wide variety of shapes depending on pulsar characteristics. Conventional methods for estimating NF (e.g., Ritchings (1976)) suffer when the signal-to-noise ratio (SNR) is low, or may face difficulties because of arbitrary shapes of the on-pulse distribution (e.g., Burke-Spolaor et al. (2012)). Further, obtaining well-justified confidence intervals (CI) on NF is difficult in these approaches. Here, we demonstrate a rigorous statistical approach to estimation and inference about NF.

2. Methodology

We model the on-pulse energy distribution as a mixture of two or more univariate Gaussians, including a possible null-pulse component. This is an instance of the classic mixture model problem (McLachlan & Peel 2000) where each Gaussian is

*email: [1]mihir@cms.unipune.ac.in

Table 1. NF estimates for a few well-known pulsars. Literature references: (1) J1543-0620 is a known non-nuller; (2) Bhat et al. (2007), Herfindal & Rankin (2007); (3) Ritchings (1976); (4) Biggs (1992).

Dataset	SNR	Number of Pulses	Model-conditioned %NF (CI; K_*)	Multimodel %NF (CI)	Literature %NF
J1543-0620 (B1540-06)	11.3	1700	0.0 (0.0–0.0; 4)	0.0 (0.0–0.0)	0[1]
J1136+1551 (B1133+16)	2.7	1000	21.6 (16.8–26.6; 3)	13.2 (0.0–25.9)	14–25[2,3]
	30	1000	22.5 (19.8–25.1; 3)	20.7 (19.8–24.2)	
J2321+6024 (B2319+60)	2.9	3500	28.8 (26.8–30.7; 2)	29.0 (26.8–31.2)	25[3]
	25	3500	28.5 (24.2–32.6; 3)	21.0 (24.2–32.4)	
J1745-3040 (B1742-30)	2.7	3200	28.8 (13.8–39.6; 4)	31.8 (17.3–39.9)	$\leq$ 17.5[4]
	5.3	3200	19.3 (16.1–22.5; 3)	9.9 (2.5–19.5)	

characterized by its location, scale, and weight in the mixture. The location and scale for the null-pulse Gaussian are constrained by the measured off-pulse energy distribution; the mixture weight of this Gaussian is the NF. Remaining components of the mixture together represent energy distribution of non-null pulses. Model parameters are estimated from the on-pulse data by maximizing the log-likelihood via the EM algorithm (McLachlan & Peel 2000). The optimal number K_* of mixture components is obtained by minimizing the Bayesian information criterion (BIC) (Burnham & Anderson 2002). We use the following two bootstrap strategies to obtain quantile-based 95% confidence intervals (Wasserman 2004): (a) Model-conditioned: For each bootstrap data replicate, we fit a mixture with K_* components and record NF; (b) Multimodel (Burnham & Anderson 2002): Model selection is performed afresh for each bootstrap replicate and NF for the BIC-optimal model is recorded.

3. Results

We used single-pulse off- and on-pulse energies derived from select data sets in the GMRT archives. We first validated our methodology on a sample of pulsars with known nulling properties (Table 1). We see that, by and large, we are able to reproduce known results for both nulling and non-nulling pulsars. The two J1745-3040 datasets yielded mutually inconsistent estimates, one of which is similar to the literature estimate: we believe that this is partly due to the peculiar nature of this pulsar (viz., large variations in single pulse properties, intermittent emission, occasional strong single pulses, etc.), and partly due to the extreme overlap of the null- and non-null-pulse distributions (not shown). The methodology is being applied to a set of millisecond pulsars (not shown), and our preliminary results (Rajwade et al., this volume) suggest that these millisecond pulsars null rarely, if at all.

4. Conclusion

While Gaussian mixture models have been around for decades, their use for estimating pulsar null fractions is novel here. Constraining the null-pulse location and scale, together with BIC-based model selection, makes our Gaussian mixture fits by and large well-determined. A combination of BIC-based model selection and the bootstrap inferential machinery makes our approach reasonably immune to low SNR. Occasional cases of lack of model identifiability do occur in practice. Pulse energy data may also be affected by radio frequency interference. Work on improving performance of our methodology in such situations is in progress.

References

Bhat et al., A&A, 462, 257, 2007
Biggs J.D., 1992, ApJ, 394, 574
Burke-Spolaor S., et al., 2012, MNRAS, 423, 1351
Burnham K.P., Anderson D.R., 2002, Model Selection and Multimodel Inference, Springer
Herfindal J. L., Rankin J. M., MNRAS, 380, 430, 2007
McLachlan G., Peel D., 2000, Finite Mixture Models, Wiley (New York)
Ritchings R. T., 1976, MNRAS, 176, 249
Wasserman L., 2004, All of Statistics, Springer (New York)