ON ROBUST DETECTION OF BRAIN STIMULI WITH RAMANUJAN PERIODICITY TRANSFORMS

Pouria Saidi, George Atia and Azadeh Vosoughi

Department of Electrical and Computer Engineering, University of Central Florida,
Orlando, FL 32816, USA
Emails: pouria.saidi@knights.ucf.edu, George.Atia@ucf.edu, Azadeh@ucf.edu

ABSTRACT

Visual Evoked Potentials (VEPs) are the brain responses to repetitive visual stimuli. The ability to detect the underlying frequencies of VEPs is crucial to advancing Brain Computer Interfaces (BCIs). This paper considers the detection of such frequencies using a Ramanujan Periodicity Transform based model. We analyze the performance of a generalized likelihood ratio detector and derive the exact distributions of the sufficient statistics under hypotheses corresponding to different stimulus frequencies using confluent hypergeometric functions, along with flexible approximate distributions. Choosing stimulation periods with non-overlapping divisors is shown to enhance the detection performance. Our analysis provides guidelines for efficient design of stimulus waveforms for BCIs.

***Index Terms*—** Ramanujan Subspace, Ramanujan Periodicity Transform, Nested Periodic Matrices, Steady-state visual evoked potentials, Brain computer interface

1. INTRODUCTION

Repetitive visual stimulation of the brain can induce electric potentials on the occipital region of the cortex. Sending an outer stimulus at fixed frequency induces a periodic brain activity – with similar frequency to that of the stimulus [1] – known as the Steady State Visual Evoked Potential (SSVEP). SSVEPs are accompanied with background activity of the brain, which makes their detection a rather challenging task. In [1], a method for detecting such responses was proposed and a decision rule based on power statistics was derived. Similar to [1], we use white Gaussian noise to model the background electroencephalogram (EEG) activity. While this assumption does not accurately capture the actual statistics of the background neural noise by virtue of its non-stationarity, we use it here for simplicity.

Brain Computer Interfaces (BCIs) exploit SSVEPs to construct a communication pathway between the human brain and external devices. Of primary importance to advancing BCI technology is to enhance our ability to perform real-time processing of EEG signals and classify them based on their corresponding frequencies. Power Spectral Density Analysis (PSDA) is one of the conventional methods to detect the underlying frequency of a signal and has been widely used in EEG processing [2]. However, PSDA and related spectral-based methods require recording EEG signals of fairly long durations to enhance frequency resolution, which limits their usability in real-time applications. In [3], a powerful method to detect the underlying frequency of SSVEPs based on Canonical Correlation Analysis (CCA) was proposed. Since CCA is among the state-of-the-art techniques for SSVEP detection, we compare the classification results of our proposed approach to that of CCA in Section 3.

The authors in [4] developed a new family of matrices dubbed Nested Periodic Matrices (NPM) to estimate the periodicity in periodic sequences. They also introduced the so-called Ramanujan Periodicity Transforms (RPT) as an instance of NPMs. The RPT-based model was shown to exhibit robustness to noise and phase shifts. We have used it for detection of SSVEPs for the first time in [5], where we demonstrated its ability to capture the underlying periodicity in SSVEPs and its robustness to latencies naturally present in the brain response.

In this work, we study the performance of the RPT-based model in a composite binary hypothesis testing problem when the background EEG is modeled using additive white Gaussian noise. As mentioned earlier, the SSVEP is periodic with the same frequency as that of the stimulus. We leverage this prior information to design, and analyze the performance of, a generalized likelihood ratio (GLR) detector of the underlying periodicity based on ML estimates of the signal on the known support sets. Our analysis of the distribution of the sufficient statistics under the different hypotheses provides guidelines for efficient design of the external stimuli in BCIs.

The paper is organized as follows. In Section 2, we briefly review some of the properties of Ramanujan sums and RPT matrices. Then, we formulate the detection problem and present the performance analysis for the proposed GLR de-

This work was supported in part by NSF Grants CCF-1525990, CCF-1341966, CCF-1552497 and CCF- 1320547 .

tector. We provide our results in Section 3. Section 4 provides a discussion and a conclusion.

2. PROBLEM DESCRIPTION

2.1. Ramanjuan Sums and RPT

Ramanujan sums [6] are the building blocks of the RPT matrices developed in [4]. Some of the key properties of Ramanujan sums and RPT matrices are investigated in [7] and [8]. A Ramanujan sum is defined as

$$c_q(n) = \sum_{\substack{k=1 \\ (k,q)=1}}^{q} e^{j2\pi kn/q}, \tag{1}$$

where (k, q) is the greatest common divisor (gcd) of k and q. The sequence $c_q(n)$ is an all integer, symmetric and periodic sequence with period q, i.e., $c_q(n+q) = c_q(n)$.

Orthogonality: In [7], it was shown that $c_{q_1}(n)$ and $c_{q_2}(n)$ for $q_1 \neq q_2$ are orthogonal over the sequence length $L = \mathrm{lcm}(q_1, q_2)$, the least common multiplier of q_1 and q_2. Therefore,

$$\sum_{n=0}^{L-1} c_{q_1}(n)c_{q_2}(n-k) = 0, \quad q_1 \neq q_2 \tag{2}$$

for any integer $0 \leq k \leq L-1$. The Authors in [9] defined periodicity dictionaries of order $P_{\max}$ as the set of signals $\mathcal{B}$ that can represent all periodic sequences with period $1 \leq p \leq P_{\max}$, i.e., such sequences can be represented through linear combinations of signals in the set $\mathcal{B}$. They proved that a periodic dictionary of order $P_{\max}$ must have at least $\sum_{p=1}^{P_{\max}} \phi(p)$ linearly independent signals, where $\phi(p)$ is the Euler totient function of p, i.e., the number of integers from 1 to p that are co-prime to p. It was shown that a dictionary that spans all subspaces of periodic signals of periods 1 to $P_{\max}$ must contain at least $\phi(p)$ linearly independent signals with period p for each p. Define the sequence

$$\mathbf{c}_q = \begin{bmatrix} c_q(0) & c_q(1) & \dots & c_q(P-1) \end{bmatrix}^T$$

of length P. For each $1 \leq q \leq P_{\max}$, a submatrix $\mathbf{C}_q$ is constructed from columns that are circularly downshifted versions of $\mathbf{c}_q$, i.e.,

$$\mathbf{C}_q = \begin{bmatrix} \mathbf{c}_q & \mathbf{c}_q^{(1)} & \dots & \mathbf{c}_q^{(\phi(q)-1)} \end{bmatrix}, \tag{3}$$

where $\mathbf{c}_q^{(1)}$ is the circularly downshifted version of $\mathbf{c}_q$, i.e.,

$$\mathbf{c}_q^{(1)} = \begin{bmatrix} c_q(P-1) & c_q(0) & c_q(1) & \dots & c_q(P-2) \end{bmatrix}^T. \tag{4}$$

The versions $\mathbf{c}_q^{(i)}$ are similarly defined with higher order downshifts.

The RPT dictionary K can be readily constructed by concatenating all the submatrices $\mathbf{C}_q$ for all q from 1 to $P_{\max}$, where $P_{\max}$ is the largest possible value for q. Thus,

$$K = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 & \mathbf{C}_3 & \dots & \mathbf{C}_{P_{\max}} \end{bmatrix}. \tag{5}$$

2.2. RPT Detector

In this work, we study a binary hypothesis testing problem in which the hypotheses H_0 and H_1 are associated with SSVEPs of frequencies f_0 and f_1 (equivalently, with periods T_0 and T_1), respectively. To capture the periodicity of the SSVEPs, the signals are represented in an RPT dictionary. The measurements under each of the hypotheses are

$$\begin{aligned} H_0 : y &= Kx_0 + w \\ H_1 : y &= Kx_1 + w \end{aligned} \tag{6}$$

where K is the periodicity RPT dictionary, $w \sim \mathcal{N}(0, \sigma^2 I)$ an additive white Gaussian noise representing the background neural noise, and x_0 and x_1 the representations of the SSVEPs in the RPT dictionary. Without loss of generality, in the performance analysis we set $\sigma^2 = 1$, and account for different SNRs by varying the magnitude that the signal part assumes on the support. In contrast to the periodicity estimation problem in [4] where there is no information about the signals' support (i.e., the locations of the nonzero entires of x_0 and x_1), we incorporate prior information about the supports under both hypotheses knowing that the period of the SSVEP should match that of the stimulus[1]. To this end, we restrict the support under each hypothesis to the atoms of the RPT dictionary that span the subspaces of periodic signals with periods T_0 and T_1, i.e., $\mathrm{supp}(x_0) = S_0$ and $\mathrm{supp}(x_0) = S_1$ for known sets S_0 and S_1.

Assuming a deterministic model for the representation vectors (i.e., x_0 and x_1 unknown but not random), we compute the restricted ML estimates of x_0 and x_1 for the composite hypothesis test and conduct a Generalized Likelihood Ratio Test (GLRT)

$$\mathcal{L}(y) = \frac{\max\limits_{x_1:\mathrm{supp}(x_1)=S_1} f(y|H_1)}{\max\limits_{x_0:\mathrm{supp}(x_0)=S_0} f(y|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta\,. \tag{7}$$

Given the presumed Gaussian model for the background noise, the restricted ML estimate under $H_k, k = 0, 1$, is the solution to

$$\min_{x_k:\mathrm{supp}(x_k)=S_k} \|y - Kx_k\|_2^2\,, k = 0, 1, \tag{8}$$

such that $x_0 = 0$ on the complement set S_0' and $x_1 = 0$ on S_1'. The solutions to (8) are given by

$$x_{S_0} = \quad (K_0^T K_0)^{-1} K_0^T y \qquad x_{S_1} = (K_1^T K_1)^{-1} K_1^T y\,, \tag{9}$$

where K_0 and K_1 consist of the columns of K supported on S_0 and S_1, respectively. The vectors x_{S_0} and x_{S_1} are the restrictions of x_0 and x_1 to the support sets S_0 and S_1 specifying the new representations of the measurement y in the RPT

[1] In addition to the main frequency of the stimulus, spectral-based methods (e.g., PSDA) suggest that energy may also be observed in the harmonics.

dictionary. These estimates are subsequently replaced in the GLRT (7) giving the decision rule

$$\ell(y) = y^T K_1 (K_1^T K_1)^{-1} K_1^T y - y^T K_0 (K_0^T K_0)^{-1} K_0^T y \underset{H_0}{\overset{H_1}{\gtrless}} 2\sigma^2 \log \eta = \gamma \,, \qquad (10)$$

where $\ell(y)$ is the sufficient statistic. If we maintain a sequence length $L = n \cdot \mathrm{lcm}(T_0, T_1), n \in \mathbb{N}$, for the atoms of K, the orthogonality of the submatrices corresponding to different divisors will be preserved [7]. The matrices K_0 and K_1 comprise all the submatrices corresponding to the divisors of T_0 and T_1. If T_0 and T_1 share any divisors, $S_0 \cap S_1 \neq \emptyset$, i.e., the support sets are non-disjoint. Therefore, only the submatrices supported on $S_0 \setminus S_1$ and $S_1 \setminus S_0$ contribute to $\ell(y)$ in (10), where $S_i \setminus S_j$ denotes the difference set of sets S_i and S_j. To simplify notation, we use $K_{i\setminus j} := K_{S_i \setminus S_j}$ to denote the submatrix of K indexed by the difference set $S_i \setminus S_j$. Therefore, the decision rule in (10) reduces to

$$\ell(y) = y^T B y - y^T A y \underset{H_0}{\overset{H_1}{\gtrless}} 2\sigma^2 \log \eta = \gamma \,, \qquad (11)$$

where A and B can be expressed as

$$\begin{aligned} A &= K_{0\setminus 1} (K_{0\setminus 1}^T K_{0\setminus 1})^{-1} K_{0\setminus 1}^T \\ B &= K_{1\setminus 0} (K_{1\setminus 0}^T K_{1\setminus 0})^{-1} K_{1\setminus 0}^T \,. \end{aligned} \qquad (12)$$

2.3. Performance Analysis

We identified more properties for the matrices A and B from the properties of Ramanujan sums, based on which we derived the probability density functions (pdfs) of the sufficient statistic $\ell(y)$ under both hypotheses. For brevity, we state the following results without proof and defer the proofs to a longer version of this work.

Lemma 1 *The matrices A and B in (12) are idempotent.*

Lemma 2 *(Adopted from [10]) Let $x \sim \mathcal{N}(\mu, \Sigma)$, with Σ positive definite and let A be a deterministic and symmetric matrix. Then, $x^T A x$ has a non-central χ_f^2 distribution with f degrees of freedom and non-centrality parameter λ^2 if and only if $A\Sigma A = A$, where $f = \mathrm{trace}(A)$ and $\lambda^2 = \mu^T A \mu$.*

From Lemma 1 and Lemma 2, $y^T A y$ and $y^T B y$ have non-central χ^2 distributions. Since A and B are idempotent by Lemma 1, the degrees of freedom and non-centrality parameters for these distributions are obtained as in (13). Furthermore, using [10, Theorem 1], we can show that the random variables $y^T A y$ and $y^T B y$ are independent. Therefore, the sufficient statistic $\ell(y)$ for the GLRT in (11) is the difference of two quadratic and mutually independent chi-squared random variables with different non-centrality parameters. From the statistical theory of linear combinations of non-central chi-squared variates (e.g., see [11, Theorem 4.1]), we can obtain the closed-form expressions for the distributions of the sufficient statistic under both hypotheses in terms of confluent hyper-geometric functions. The proof of this result is omitted for brevity. However, in this work we approximate these distributions with Gaussian distributions by fitting the first and second order moments. Recalling that a non-central chi-squared distribution with f degrees of freedom and non-centrality parameter λ^2 has a mean $\mu = f + \lambda^2$ and variance $\sigma^2 = 2(f + 2\lambda^2)$, the two quadratic terms $y^T A y$ and $y^T B y$ have a $\chi^2(m, \lambda_{kA}^2)$ and $\chi^2(n, \lambda_{kB}^2)$ distribution under hypotheses $H_k, k = 0, 1$, where $(m, n, \lambda_{0A}^2, \lambda_{1B}^2)$ are given in (13) below. Following from the orthogonality of the submatrices associated with different divisors, $\lambda_{1A}^2 = \lambda_{0B}^2 = 0$.

$$\begin{aligned} \lambda_{0A}^2 &= \sum_{\substack{i=1 \\ d_i \notin \mathcal{T}_1}}^{v} x_{0,i}^T K_{0,i}^T K_{0,i} x_{0,i} \quad \lambda_{1B}^2 = \sum_{\substack{j=1, \\ d_j \notin \mathcal{T}_0}}^{u} x_{1,j}^T K_{1,j}^T K_{1,j} x_{1,j} \\ n &= \sum_{\substack{i=1 \\ d_i \notin \mathcal{T}_1}}^{v} \phi(d_i | T_0) \qquad m = \sum_{\substack{j=1 \\ d_j \notin \mathcal{T}_0}}^{u} \phi(d_j | T_1) \end{aligned} \qquad (13)$$

where $d|T$ denotes that d is a divisor of T, $\mathcal{T}_0$ and $\mathcal{T}_1$ are the sets of divisors of T_0 and T_1, v and u are the number of divisors of T_0 and T_1, respectively and $K_{0,i}, i = 1, \ldots, v$, and $K_{1,j}, j = 1, \ldots, u$, the corresponding submatrices. Following the same notation, $x_{k,i}$ and $x_{k,j}$ denote the ML estimates of x_k derived from (8) restricted to the index set of the ith and jth divisors of T_0 and T_1.

Following from the aforementioned independence property of the two quadratic terms, we obtain the Gaussian approximations for $\ell(y)$ under both hypotheses as

$$\begin{aligned} H_0 : \ell(y) &\sim \mathcal{N}(m - n - \lambda_{0A}^2, 2(m+n) + 4\lambda_{0A}^2) \\ H_1 : \ell(y) &\sim \mathcal{N}(m - n + \lambda_{1B}^2, 2(m+n) + 4\lambda_{1B}^2) \end{aligned} \qquad (14)$$

where the symbol $\sim$ means 'statistically distributed as'. We can readily write the approximate probabilities of detection P_D and false alarm P_F as

$$\begin{aligned} P_D &= \int_{\ell(y)>\gamma} P(\ell(y)|H_1) dy = Q\left(\frac{\gamma - m + n - \lambda_{1B}^2}{\sqrt{2(m+n) + 4\lambda_{1B}^2}} \right) \\ P_F &= \int_{\ell(y)>\gamma} P(\ell(y)|H_0) dy = Q\left(\frac{\gamma - m + n + \lambda_{0A}^2}{\sqrt{2(m+n) + 4\lambda_{0A}^2}} \right) \end{aligned} \qquad (15)$$

where $P(\ell(y)|H_i), i = 0, 1$, denotes the distribution of $\ell(y)$ under the ith hypothesis, and $Q(.)$ the Q-function, the tail of the standard Normal distribution. Under the Neyman Pearson criteria, for a false alarm level $P_F \leq \alpha$, the maximum probability of detection is

$$P_D = Q\left(\frac{Q^{-1}(\alpha)\sqrt{2(m+n) + 4\lambda_{0A}^2} - \lambda_{0A}^2 - \lambda_{1B}^2}{\sqrt{2(m+n) + 4\lambda_{1B}^2}} \right) . \qquad (16)$$

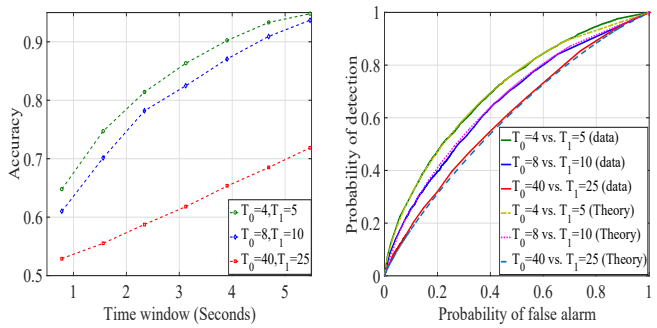

Fig. 1. The accuracy is obtained as function of the length of the time window for SNR $= -20$ dB. T_0 and T_1 are chosen such that the orthogonality is preserved (left). The ROC curves are shown for both the synthesized data and the theory prediction in (15) (right).

Since the Q-function is monotonically decreasing, the trade-off curve between P_D and P_F can improve for larger values of λ_{0A}^2 and λ_{1B}^2. In other words, if the two sets $\mathcal{T}_0$ and $\mathcal{T}_1$ share a smaller number of divisors, a better performance could be attained. As an example, we choose values for T_0 and T_1 from the set of divisors of 200. Preserving the orthogonality, we compare the performance of the GLR detector for three cases in Fig. 1. For the first case, $T_0 = 40$ and $T_1 = 25$, for the second $T_0 = 8$ and $T_1 = 10$ and for the third $T_0 = 4$ and $T_1 = 5$. In the first case, $\mathcal{T}_0$ and $\mathcal{T}_1$ share more divisors, which yields smaller values for λ_{0A}^2 and λ_{1B}^2, hence worse performance. Fig. 1 (right) illustrates the ROCs from both the synthesized data and the Gaussian approximation in (15), which are shown to exhibit very good agreement. The accuracy of the classifier for time windows of different lengths is depicted in Figure 1 (left).

3. RESULTS

Real data: In this section, we employ a publicly available SSVEP dataset [12]. The dataset contains 12 goal frequencies with sampling frequency of 256 Hz. 10 subjects have participated in the experiment and 15 trials are recorded for each of the subjects per frequency. Hence, the dataset has a total number of 150 trials for each of the classes. The onset of visual stimulation is at the 39th sample point and we only employ post-stimulus samples.

We filter the observations between 4-30 Hz using a band-pass filter to improve the SNR. From the available frequencies, we choose the set $\mathcal{F} = \{10.25, 12.25, 14.25, 12.75\}$ in units of Hz. Given the sampling frequency, the set of periods $\mathcal{T} = \{24.97, 20.89, 17.96, 20.07\}$. It is important to note that the RPT model is constructed based on integer periodicities, therefore we round these numbers to the nearest integer values. We obtain the accuracy of multiple binary detectors by choosing pairs of frequencies from the set $\mathcal{F}$.

Remark 1 *In our experiments, we picked those frequencies whose periods are close to integer numbers since the RPT model – which is based on the Ramanujan sums – is an all-integer model, hence, can best capture periodicity when the period of the sequence is an integer or very close to an integer.*

Remark 2 *Our analysis uses the orthogonality of the defined submatrices $K_{0\setminus 1}$ and $K_{1\setminus 0}$. To preserve orthogonality, we should maintain a length for the sequences that is equal to (multiple of) the lcm of the two periods, i.e., $n \cdot \mathrm{lcm}(T_0, T_1)$, where $n \in \mathbb{N}$. While we can also use shorter sequences (and lose orthogonality) to reduce the latency for real-time processing, the performance analysis for the non-orthogonal case is different and will be discussed in our future work.*

Remark 3 *A single electrode (electode Oz) is used to evaluate and compare all the methods. This electrode is located on the ocular region on the cortex. Extensions to multi-channel processing is also deferred to a longer version of this work.*

Figure 2 illustrates the accuracy of detecting the underlying periodicity using time windows of different lengths. The figure shows the accuracy when the length of the sequences is a multiple of the lcm of the two periods i.e. $L = n \cdot \mathrm{lcm}(T_0, T_1), n \in \mathbb{N}$, as well as for time windows of duration ranging from 1 to 4 seconds (where generally orthogonality is not preserved). The accuracy of the RPT detector is compared to PSDA as a conventional method and to CCA as the state-of-the-art technique. As shown, the RPT detector performance is comparable to CCA and substantially outperforms PSDA. Figure 2 (bottom-right) illustrates an example where one of the periods is not an integer ($T_1 = 20.89$), however the RPT-based detector succeeds at detecting the true underlying periodicity.

4. DISCUSSION AND CONCLUSION

In this work, we proposed a method based on Ramanujan Periodicity Transform to detect the underlying frequency in SSVEPs. The PSDA method failed to detect the frequencies for each class since higher frequency resolution was required for this dataset. However, both the RPT detector and CCA could detect the underlying periodicities in real data. Figure 2 illustrates the accuracy for different time windows. In all four examples, the accuracy of the RPT detector exceeds 80 percent by using 1 to 1.5 seconds of the data. The accuracy is about 90 percent using 2 seconds in three of the examples. This underscores the promise this approach holds for real-time BCI. Our results indicate that selecting periods with smaller number of shared divisors, can improve performance. For instance, when $f_0 = 10.25$ Hz and $f_1 = 12.75$ Hz (equivalently, $T_0 = 25$ and $T_1 = 20$), the sets of divisors are $\mathcal{T}_0 = \{1, 5, 25\}$ and $\mathcal{T}_1 = \{1, 2, 4, 5, 10, 20\}$, and for the 10.25 Hz vs. 12.25 Hz example (i.e. $T_0 = 25$, $T_1 = 21$), $\mathcal{T}_0 = \{1, 5, 25\}$ and $\mathcal{T}_1 = \{1, 3, 7, 21\}$. In the former case,

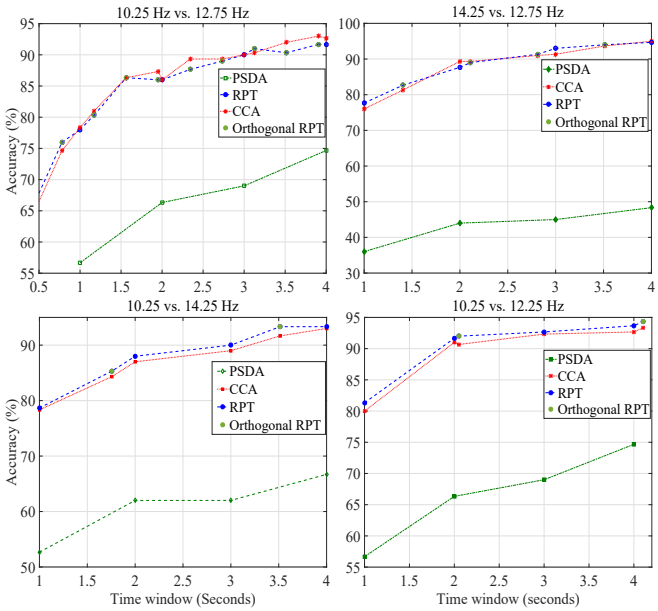

Fig. 2. Classification performance for frequencies 10.25 Hz vs. 12.75 Hz (Top-left), 10.25 Hz vs. 14.25 Hz (Bottom-left), 14.25 Hz vs. 12.75 Hz (Top-right), and 10.25 Hz vs. 12.25 Hz (Bottom-right)

the two sets share a larger number of divisors, hence the supports defined in (7) have a larger overlap, which reduces the separability of the hypotheses. Furthermore, when T_0 and T_1 are even, the RPT model spans the subspace of the first harmonic in addition to the main frequency, hence the accuracy improves. In most examples, the RPT detector could achieve high accuracy and slightly outperform the state-of-the-art CCA algorithm.

In designing waveforms for BCI, it is useful to select frequencies for the external stimuli such that the overlap of the corresponding support sets is minimized. Furthermore, orthogonal RPT designs can prove beneficial and are easy to analyze. However, our investigations have shown that non-orthogonal designs could sometimes yield performance gains over orthogonal ones – especially when reducing latency is contemplated in real-time or delay intolerant applications – albeit the performance analysis becomes more tedious. The study of non-orthogonal designs as well as their analysis are subject of future work.

5. REFERENCES

[1] A. P. Liavas, G. V. Moustakides, G. Henning, E. Z. Psarakis, and P. Husar, "A periodogram-based method for the detection of steady-state visually evoked potentials," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 2, pp. 242–248, 1998.

[2] A. Paris, A. Vosoughi, and G. Atia, "Whitening 1/f-type noise in electroencephalogram signals for steady-state visual evoked potential brain-computer interfaces," in *48th Asilomar Conference on Signals, Systems and Computers*, 2014, pp. 204–207.

[3] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for ssvep-based bcis," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2610–2614, 2006.

[4] S. V. Tenneti and P. Vaidyanathan, "Nested periodic matrices and dictionaries: new signal representations for period estimation," *IEEE Transactions on Signal Processing*, vol. 63, no. 14, pp. 3736–3750, 2015.

[5] P. Saidi, G. Atia, and A. Vosoughi, "Detection of visual evoked potentials using ramanujan periodicity transform for real time brain computer interfaces," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 959–963.

[6] S. Ramanujan, "On certain trigonometrical sums and their applications in the theory of numbers," *Trans. Cambridge Philosoph. Soc*, vol. XXll, no. 13, pp. 259–276, 1918.

[7] P. Vaidyanathan, "Ramanujan sums in the context of signal processing part i: Fundamentals," *IEEE Transactions on Signal Processing*, vol. 62, no. 16, pp. 4145–4157, 2014.

[8] ——, "Ramanujan sums in the context of signal processing part ii: For representations and applications," *IEEE Transactions on Signal Processing*, vol. 62, no. 16, pp. 4158–4172, 2014.

[9] S. V. Tenneti and P. Vaidyanathan, "A unified theory of union of subspaces representations for period estimation," *IEEE Transactions on Signal Processing*, vol. 64, no. 20, pp. 5217–5231, 2016.

[10] A. T. Craig, "Note on the independence of certain quadratic forms," *The Annals of Mathematical Statistics*, vol. 14, no. 2, pp. 195–197, 1943.

[11] S. J. Press, "Linear combinations of non-central chi-square variates," *The Annals of Mathematical Statistics*, pp. 480–487, 1966.

[12] M. Nakanishi, Y. Wang, Y.-T. Wang, and T.-P. Jung, "A comparison study of canonical correlation analysis based methods for detecting steady-state visual evoked potentials," *PLOS ONE*, vol. 10, no. 10, 2015.