Oscillation detection with period uncertainty and a limited sampling budget

Turner Silverthorne[1,2,3], Matthew Carlucci[2,3], Arturas Petronis[2,3], and Adam R Stinchcombe[1]

[1] Department of Mathematics, University of Toronto, Toronto, Canada
[2] The Krembil Family Epigenetics Laboratory, The Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Canada
[3] Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius, Lithuania

May 19, 2024

Abstract

Most rhythm detection experiments involve measurements collected at regular time-intervals. These equispaced designs will fail to detect oscillations at particular acrophases and frequencies even when oscillation amplitude is large. The spurious false negatives pose a challenge for studies aiming to determine the presence or absence of oscillations across a range of parameters. Here, we present a method to construct sampling schemes that are robust to parameter uncertainty, and demonstrate their improvements relative to equispaced designs. We prove that maximizing the worst-case statistical power is equivalent to a mixed-integer quadratically-constrained programming problem with convex relaxation. Using this equivalence, we construct optimal and near-optimal designs for a range of experimental conditions. Our method also allows us to include regularization and sampling constraints, ensuring that improving statistical power via measurement scheduling does not introduce unnecessary complexity in the design.

Contents

1 Introduction **2**

2 Background **4**
2.1 Power analysis for harmonic regression 4
2.2 Fisher information matrix 5

3 Results **6**
3.1 Optimized sampling improves worst-case statistical power 6
3.2 Periodogram analysis with balanced and threshold designs 8
3.3 Proof of the equivalence theorem and derivation of the method 9
3.4 Numerical evidence for optimality in the regularization-only problem 11

4 Discussion **18**

5 Methods **18**

6 Acknowledgements **23**

1 Introduction

Biological oscillations serve various essential roles in living systems [1]. When studying these systems experimentally, measurement timing is crucial for ensuring that oscillations are not overlooked as random noise. In general, protocols which measure at regular time intervals, known as equispaced or uniform designs, are treated as the gold standard for rhythm detection. If the period of the oscillations is known, this preference is justifiable since equispaced measurements have optimal statistical properties for fixed-period harmonic regression [2, 3]. On the other hand, systems with an unknown periodicity are outside the scope of this justification. In addition, ethical constraints may prohibit the collection of measurements at regular time intervals. For example, a study with human participants may be required to only collect data during waking hours e.g. if the participants are from a vulnerable population. Hence, systems with period uncertainty or sampling constraints pose new challenges in the optimal design of rhythm detection experiments.

When considering period uncertainty, a shortcoming of equispaced measurements can be illustrated in a simulated rhythm detection experiment (Fig. 1). All of the simulated oscillations have the same amplitude, and are randomly assigned an acrophase and a long, medium, or short period. When measurements are collected at regular time intervals, the rhythm detection results are artificially depleted for certain acrophases, even though oscillations are equally present across all acrophases. This depletion occurs because short period oscillators of certain acrophases are always near the same value when measurements are collected (Fig. 1A). On the other hand, if measurements are collected at irregular time-intervals, the short-period oscillations are detected without any evidence of phase depletion, while the medium and long period oscillations continue to be detected reliably (Fig. 1B). Examining other measurement budgets, phase-dependent power consistently emerges when the frequency of the oscillation is close to half the measurement budget (Fig. 1C). This critical frequency is known as the Nyquist rate of the design. While it is not unusual for spectral features to behave poorly near the Nyquist rate, we investigated whether this loss of performance could be routinely circumvented through irregular sampling.

Using an exact expression for statistical power [4], we investigate how parameter uncertainty and the measurement budget determine the applicability of irregular sampling. We find that the worst-case statistical power, meaning the lowest value of the power across all signals of a given amplitude, is a reasonable figure of merit for this problem. We prove that worst-case power maximization can be reformulated as a mixed integer quadratically constrained (MIQCP) programming problem with convex relaxation. This MIQCP is closely related to Elfving optimality (or E-optimality), a topic in optimal experimental design which has been well-studied for a variety of statistical models including harmonic regression [2, 5].

The remainder of this paper is structured as follows. Section 2 provides background necessary for understanding our approach to power maximization. The numerical consequences of our main theoretical result are investigated in Sec. 3.1. Using a standard MIQCP-programming method, we perform worst-case power optimization for a range of budgets and period uncertainties. These solutions perform well in worst-case scenarios but introduce unwanted variability in the power. We reduce power variability by adding a regularization term to the objective and enforcing equispaced constraints on a subset of the measurements. In Sec. 3.2 we demonstrate that the regularization and equispaced constraints lead to more robust designs by showing that they perform better when periodogram-based hypothesis testing is performed. Sec. 3.3 provides a proof of our equivalence theorem. We conclude in Sec. 3.4 by investigating convergence properties of the pure-regularization problem. For certain values of the grid coarseness, we find numerical evidence that the irregular candidate solutions are globally optimal. To make this work accessible to all readers, we provide an R package known as Power-CHORD (*P*ower optimization via *C*onvex programming for *HO*moscedastic *R*hythm *D*etection) for maximizing power using our equivalence theorem, available from the first author's GitHub [6].

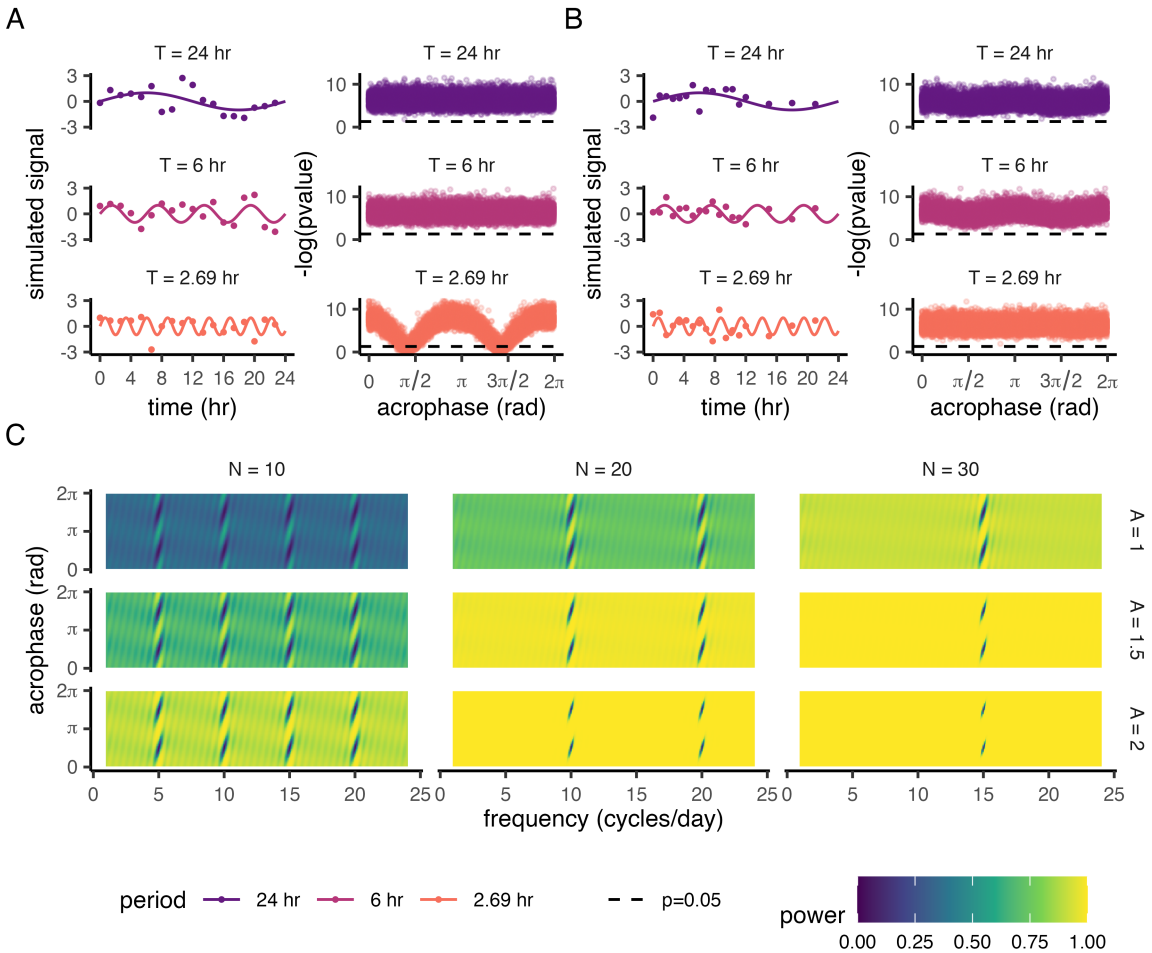

Figure 1: **Power fluctuates near the Nyquist rate of equispaced measurements.** **(A Left)** Examples of signals with long, medium, and short periods relative to the Nyquist rate of an equispaced design with $N = 18$ measurements. Points represent measurements of the oscillations with noise simulated from a normal distribution. The shortest period oscillations occur at 99% of the Nyquist rate of the design. **(A Right)** An empirical p-value distribution obtained from a fixed-period cosinor model (12,600 oscillators are simulated for each period) with the significance threshold $\alpha = 0.05$ indicated by the dashed line. Oscillations are detected robustly at all phases for the 24 hr and 6 hr period oscillators, but the 2.69 hr oscillations have a p-value distribution heavily dependent on acrophase even though the oscillations are equally prevalent at all acrophases. **(B Left)** To improve detection of fast oscillations, $N = 14$ measurements are collected in the first half of the experiment and $N = 4$ in the second half of the experiment. **(B Right)** The p-value distributions obtained from the new design no longer suffer from artificial acrophase depletion. Oscillations of all periods can now be detected robustly, even though the measurement budget is the same as in the equispaced case. **(C)** Even when the measurement budget is large, phase-dependent power emerges near the Nyquist rate when power is plotted as a function of acrophase and frequency.

2 Background

2.1 Power analysis for harmonic regression

Harmonic regression is a popular statistical framework for studying systems with a dominant frequency [7]. The single-frequency cosinor model with MESOR (*M*idline *E*stimating *S*tatistic of *R*hythm) Y_0, amplitude A, acrophase ϕ, and frequency f is given by

$$Y(t) = Y_0 + A\cos(2\pi f t - \phi) + \varepsilon(t), \tag{1}$$

in which $\varepsilon(t) \sim N(0, \sigma)$ is homoscedastic Gaussian white noise. For a known frequency f, Eq. (1) can be rewritten as

$$Y(t) = \beta_0 + \beta_1 \cos(2\pi f t) + \beta_2 \sin(2\pi f t) + \varepsilon(t). \tag{2}$$

so that all unknown parameters appear linearly. Given data $\mathbf{y} = \{y_i\}_{i=1}^N$ measured at distinct times $\mathbf{t} = \{t_i\}_{i=1}^N \subset [0,1]^N$, the optimal least squares solution for estimating the coefficients $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)$ appearing in Eq. (2) is

$$\hat{\boldsymbol{\beta}} = (X^T X)^{-1} X^T \mathbf{y}, \qquad X = \begin{bmatrix} \mathbf{1} & \cos(2\pi f \mathbf{t}) & \sin(2\pi f \mathbf{t}) \end{bmatrix}.$$

The amplitude and acrophase of the signal can be estimated from $\hat{\beta}$, to obtain $\hat{A} = \sqrt{\hat{\beta}_1^2 + \hat{\beta}_2^2}$ and $\hat{\phi} = \arctan(\hat{\beta}_2/\hat{\beta}_1)$.

Using the cosinor model, rhythm detection can be formulated as a hypothesis test in which the null hypothesis is $\beta_1 = 0 = \beta_2$ and the alternative is $\beta_1 \neq 0$ or $\beta_2 \neq 0$. Equipped with a formal statistical test, the statistical power can be used to quantify how reliably the experimental design detects oscillations. Since the power of a study is influenced by factors that are unknown prior the experiment, it is useful to construct designs that achieve high power in a range of conditions. For rhythm detection, assuming a prior distribution on the frequency and a minimum effect size, one would then be able to determine how many samples are necessary to achieve at least a given statistical power. We adopt this framework here, and seek experimental designs that achieve high power for a range of frequencies and acrophases.

Monte Carlo simulation is a simple technique for estimating the power of a proposed study. This type of strategy is too computationally expensive for numerically optimizing the power. Fortunately, the statistical power of the one-frequency cosinor model can be evaluated analytically under the assumption of homoscedastic white noise.

Theorem 2.1 (Power of one-frequency F test, [4])**.** *Consider the one-frequency cosinor model applied to data $\{y_i\}_{i=1}^N$ collected at times $(t_i)_{i=1}^N$, with amplitude A, noise level σ, frequency f, acrophase ϕ, and type I error rate $\alpha \in (0,1)$. Suppose we test the hypotheses*

$$H_0 : \beta_1 = 0 = \beta_2, \qquad H_1 : \beta_1 \neq 0 \text{ or } \beta_2 \neq 0$$

with H_0 as the null and H_1 as the alternative, using the test statistic

$$F = \frac{\frac{TSS - RSS}{r-1}}{\frac{RSS}{N-r}},$$

in which $TSS = ||\mathbf{y} - \langle \mathbf{y} \rangle||_2^2$ is the total sum of squares, $RSS = ||\mathbf{y} - X\hat{\beta}||_2^2$ is the residual sum of squares, and $r = 3$ is the number of degrees of freedom in the model. Then, the power $\gamma(A, \sigma, f, \phi)$ is given by the expression

$$\gamma(A, \sigma, f, \phi) = 1 - F_{\lambda(\mathbf{t}; A, \sigma, f, \phi)} \left(F_0^{-1}(1 - \alpha; 2, N-3); 2, N-3 \right), \tag{3}$$

in which

$$\lambda(\mathbf{t}; A, \sigma, f, \phi) = \frac{A^2}{\sigma^2} \sum_{i=1}^{N} \cos^2(2\pi f t_i - \phi), \tag{4}$$

$F(x; n_1, n_2)$ is the F-distribution on (n_1, n_2) degrees of freedom, and $F_\lambda(x; n_1, n_2)$ is the noncentral F distribution with (n_1, n_2) degrees of freedom with noncentrality parameter $\lambda \geq 0$. For $\lambda = 0$, the noncentral distribution F_0 agrees with the standard F distribution.

Equation (3) is significantly cheaper to evaluate than Monte-Carlo-based power estimation. Beyond accelerating our computations, Eq. (3) also influences our work due to its mathematical structure. As we will show in Sec. 3.3, analysis of Eq. (3) reveals a convex optimization problem underlying worst-case power maximization.

2.2 Fisher information matrix

The Fisher information matrix (FIM) appears in many areas of statistics, including the theory of optimal experimental design [8]. By encoding figures of merit in linear-algebraic properties of the FIM, design problems can be converted to convex programming problems [9]. Our work makes use of the FIM in a similar way, and so we recall some basic properties and definitions relevant to our model.

Definition 2.1 (Fisher information matrix)**.** Given a linear regression model $u(t) = \boldsymbol{\beta}^T \mathbf{f}(t)$ with $\mathbf{f}(t), \boldsymbol{\beta} \in \mathbb{R}^p$ and input data $\mathbf{t} = (t_1, \ldots, t_N)$ construct the design matrix $X_{\mathbf{f}(\mathbf{t})} = \mathbf{f}(\mathbf{t}) \in \mathbb{R}^{N \times p}$. The associated Fisher information matrix $M \in \mathbb{R}^{p \times p}$ is given by

$$M_{\mathbf{f}(\mathbf{t})} = X_{\mathbf{f}(\mathbf{t})}^T X_{\mathbf{f}(\mathbf{t})}.$$

We consider two Fisher information matrices in this paper: the Fisher information matrix corresponding to the one-frequency cosinor model $\mathbf{f}(t) = [1 \;\; \cos(2\pi f t) \;\; \sin(2\pi f t)]$, which we denote by

$$M = M_{\mathbf{f}(\mathbf{t})},$$

and the Fisher information matrix corresponding to the de-meaned model $\tilde{\mathbf{f}}(t) = [\cos(2\pi f t) \;\; \sin(2\pi f t)]$, denoted by

$$\tilde{M} = M_{\tilde{\mathbf{f}}(\mathbf{t})} = \sum_{t \in \mathbf{t}} \begin{pmatrix} \cos(2\pi f t)^2 & \cos(2\pi f t)\sin(2\pi f t) \\ \cos(2\pi f t)\sin(2\pi f t) & \sin(2\pi f t)^2 \end{pmatrix}. \tag{5}$$

We refer to $\tilde{M}$ as the reduced Fisher information matrix. When studying the Fisher information matrix in an optimization setting, it is common to restrict the measurement times $\mathbf{t}$ to a fixed partition $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_K) \subset [0, 1]$ for some large $K \gg N$. The definition of the Fisher information matrix can be modified accordingly.

Definition 2.2 (Fisher information matrix associated to a partition)**.** Given a linear model $y(t) = \boldsymbol{\beta}^T \mathbf{f}(t)$, partition $\boldsymbol{\tau} \subset [0, 1]$ with $|\boldsymbol{\tau}| = K$, and weight function $\mu \in \{0, 1\}^K$, let $M(\boldsymbol{\mu}; \boldsymbol{\tau}, \mathbf{f})$ denote the Fisher information matrix

$$M(\boldsymbol{\mu}; \boldsymbol{\tau}, \mathbf{f}) = \sum_{i=1}^{K} \mu_i \mathbf{f}(\tau_i) \mathbf{f}(\tau_i)^T.$$

We can use Defn. 2.2 to rewrite Eq. (5) as

$$\tilde{M} = \sum_{i=1}^{K} \mu_i \begin{pmatrix} \cos(2\pi f \tau_i)^2 & \cos(2\pi f \tau_i)\sin(2\pi f \tau_i) \\ \cos(2\pi f \tau_i)\sin(2\pi f \tau_i) & \sin(2\pi f \tau_i)^2 \end{pmatrix}$$

in which the weights $\mu_i \in \{0, 1\}$ obey $\sum_{i=1}^{K} \mu_i = N$. By working with a fixed partition and searching for an optimal choice of weights, the decisions variables are now discrete instead of continuous real numbers. Still, this discrete formulation is preferable because the decision variables μ_i appear linearly.

3 Results

For the remainder of the paper, we scale time so that the duration of the entire study is $T = 1$ and set the noise strength $\sigma = 1$ to reduce the number of parameters in the model. In these units, the amplitude parameter A can be interpreted as a measure of the signal-to-noise ratio. We also follow the convention that the acrophase $\phi \in [0, 2\pi)$ is expressed in radians so that it is easier to compare phase-dependent features between signals of different frequencies.

3.1 Optimized sampling improves worst-case statistical power

We construct two types of irregular measurement schedules: threshold designs and balanced designs. Threshold designs are constructed by maximizing the worst-case power for a fixed amplitude and a range of frequencies and acrophases. We use the terminology "threshold" because such designs are generated with the aim of scheduling measurements such that the power is above the highest possible threshold for all parameters of interest. In contrast, balanced designs are constructed to achieve a favorable trade-off between worst-case power and variability in power.

Our focus on worst-case power in the construction of threshold designs is motivated by our model of parameter uncertainty. We assume that N measurements are to be collected over the course of an experiment that seeks to investigate all frequencies f in a given window $f_{\min} \leq f \leq f_{\max}$ and no information is available about the acrophase of the signal before the experiment takes place. Without any knowledge of the acrophase, it is impossible to alter the average power by rescheduling measurements.

Proposition 3.1 (Conservation of average power)**.** *For a prior distribution on frequency and acrophase of the form $p(f, \phi) = \frac{1}{2\pi}\rho(f)$, the average value of the noncentrality parameter is given by*

$$\mathbb{E}_{p(f,\phi)}\left[\lambda(\mathbf{t}; A, \sigma, f, \phi)\right] = N\frac{A^2}{2},$$

in which N is the measurement budget.

Proof. By Theorem 2.1, the timing of measurements only influences the power through the noncentrality parameter given in Eq. (4). Integration against a uniform prior shows that the noncentrality parameter is only sensitive to the number of measurements. □

Corollary 3.0.1. *Provided that the acrophase prior is uniform, the average power with respect to frequency and acrophase depends only on the number of measurements and not when they are scheduled.*

For the remainder of this work, we model parameter uncertainty using a uniform prior in phase and frequency

$$p(f, \phi) = \frac{1}{2\pi(f_{\max} - f_{\min})}.$$

Since it is not meaningful to optimize average power in this setting, we focus instead on worst-case power. Our main theoretical result converts worst-case power maximization into an equivalent but better studied optimization problem. We defer the proof of Theorem 3.1 to Sec. 3.3 and focus first on the numerical consequences.

Theorem 3.1. *For a rhythm detection experiment with $N > 3$ distinct measurements and homoscedastic noise, finding a design $\mathbf{t}^*$ that maximizes power uniformly in acrophase and frequency*

$$\mathbf{t}^* = \underset{\{\mathbf{t}\in[0,1]^N \mid t_1<t_2<\cdots<t_N\}}{\arg\max} \quad \underset{(f,\phi)\in[f_{\min},f_{\max}]\times[0,2\pi)}{\min} \gamma(\mathbf{t}; A, f, \phi), \tag{6}$$

is equivalent to solving a convex mixed-integer quadratically constrained programming problem

$$(\boldsymbol{\mu}^*, s^*) = \underset{\boldsymbol{\mu}\in\{0,1\}^M, \sum_{i=1}^M \mu_i = N, s\in\mathbb{R}}{\arg\min} \left\{ s \mid \boldsymbol{\mu}^T Q_1 \boldsymbol{\mu} \leq s, \ldots, \boldsymbol{\mu}^T Q_K \boldsymbol{\mu} \leq s \right\}, \tag{7}$$

in which $Q_1, \ldots, Q_K$ are symmetric positive semi-definite matrices.

The optimization problems in Eq. (6) and Eq. (7) are equivalent in theory, but their structural differences have computational consequences which influence their applicability. Since the cost function in Eq. (6) is cheap to evaluate, various stochastic optimization methods [10] would be reasonable choices for constructing near-optimal designs. On the other hand, it may be useful to estimate how close a proposed design is to the theoretical optimal performance. From this perspective, the formulation in Eq. (7) is advantageous because it allows the problem to be studied by standard MIQCP programming methods. These methods provide estimates of the difference in performance between a candidate design and the theoretical maximum. Such estimates can be time-consuming to produce, with run-times heavily dependent on the parameters under consideration. We find that after adding additional structure to the optimization problem, we are able to obtain numerical evidence of the optimality of certain candidate designs. Since our work aims to understand the theoretical limitations of power optimization while also constructing high-quality numerical solutions, Eq. (7) is best-suited to our purposes.

Threshold designs are constructed by maximizing Eq. (7) numerically. When compared in a worst-case scenario, threshold designs can perform significantly better than equispaced designs (Fig. 2). For various measurement budgets, irregular sampling is only worthwhile when the Nyquist rate of the equispaced design is included in the frequency window under investigation. This introduces a threshold in the $(f_{\min}, f_{\max})$ plane, above which one can expect substantial benefits from irregular sampling (Fig. 2A). Examining the distribution of measurement times for a particular experimental setting ($N = 48$, $f_{\min} = 1$, $f_{\max} = 24$) reveals that measurements are pulled closer together by the optimization method, reducing the temporal coverage of the experiment but improving its ability to detect high frequency oscillations (Fig. 2B). Unlike the equispaced designs which undergo power oscillations at integer multiples of the Nyquist rate, the irregular designs do not have a strong dependence on the acrophase or frequency of the signal (Fig. 2C). At the same time, removal of the strong phase dependence is accompanied by weak power oscillations throughout the acrophase-frequency plane, especially at low measurement budgets. While irregular sampling improves worst-case power, there is a clear need for additional structure in the optimization problem.

We consider two strategies for adding structure to the optimization problem and thereby striking a better balance between worst-case power and power variability. The first strategy involves adding a penalty term to the objective function to introduce a bias away from solutions with a high degree of power variability. This is an example of regularization, a common technique for finding robust solutions to optimization problems. In our problem, we use the average gradient of the noncentrality parameter to construct a regularization term. This term takes the form given in Prop. 5.5 and is simplified using Cor. 5.0.1. Added to the optimization problem, we are left with an expression of the form

$$(\boldsymbol{\mu}^*, s^*) = \underset{\boldsymbol{\mu}\in\{0,1\}^M, \sum_{i=1}^M \mu_i = N, s\in\mathbb{R}}{\arg\min} \left\{ s \mid \boldsymbol{\mu}^T \left(Q_1 + w_{\text{reg}} R\right) \boldsymbol{\mu} \leq s, \ldots, \boldsymbol{\mu}^T \left(Q_K + w_{\text{reg}} R\right) \boldsymbol{\mu} \leq s \right\}, \tag{8}$$

in which the weight $w_{\text{reg}} > 0$ determines the strength of the regularization and R is the symmetric matrix given in Eq. (26).

The second strategy involves assigning a subset of the measurements to an equispaced measurement grid (Fig. 3). Since equispaced measurements have low worst-case power but appear to exhibit little power variability, we expect that a balance between power in a worst-case scenario and average power variability could be found by searching for designs that contain an equispaced sub-design. Equispaced constraints also simplify the final design. We use the terminology "balanced designs" to refer to optimal and near optimal solutions to Eq. (8) with optional equispaced sampling constraints.

To gauge the influence of regularization and equispaced constraints on balanced designs, we imposed equispaced constraints that require measurements to be collected either every 1 hr, 2 hr, or 3 hr. For a fixed amount of frequency uncertainty ($f_{\text{min}} = 1$,$f_{\text{max}} = 24$), the presence of equispaced constraints, explicit regularization, and the measurement budget together determine the relationship between worst-case power and variability (Fig. 4). As expected, the stricter equispaced constraints force the design to be more similar to an equispaced design and therefore limit performance in a worst-case setting. Explicit regularization slightly diminishes power variability while unconstrained designs generally achieve the best balance between worst case power and variability. Overall, the various balanced designs considered here perform comparably well to threshold designs in terms of worst-case power and power variability.

3.2 Periodogram analysis with balanced and threshold designs

Our study of threshold and balanced designs has been limited to the setting of a fixed-period cosinor model. If such designs were to be employed experimentally, the data they produce would likely be analyzed in broader statistical frameworks. For instance, periodogram analysis may be performed as an initial test of periodicity before fitting a fixed-period cosinor model to the data. In this section, we simulate periodogram analysis with threshold, balanced, and equispaced designs to determine if they remain applicable in this more general statistical context. Several approaches to periodogram analysis are commonly employed in the study of periodic data. We make use of the Lomb-Scargle periodogram, a method developed for the study of irregularly sampled data [11, 12]. Given a list of frequencies, the Lomb-Scargle periodogram returns a power for each frequency which can be interpreted as the intensity of that frequency component in the data. High intensity frequencies, which manifest as peaks in the periodogram, can then be investigated using hypothesis tests. We detect oscillations using a hypothesis test derived from the Lomb-Scargle periodogram [13] which is included in a standard periodogram library [14].

We investigated the reliability of equispaced, balanced, and threshold designs for Lomb-Scargle based rhythm detection study in a simulated dataset. The dataset was made up oscillatory and white noise signals of various amplitudes, acrophases, and frequencies (Fig. 5). Performance differences between designs were most noticeable at the highest and lowest frequencies under consideration. Importantly, the balanced design performed well at high and low frequencies, whereas the threshold design failed to detect the lowest frequency oscillations and the equispaced design could not resolve the highest frequency oscillations. These differences in performance persisted when the amplitude and abundance of oscillatory signals were increased. A randomly generated design was included for reference and was not reliable at detecting high frequency oscillations, similar to the equispaced design. This supports the argument that the performance improvements in the threshold and balanced designs are stronger than would be expected of a design that deviates randomly from equispaced measurements.

3.3 Proof of the equivalence theorem and derivation of the method

We first consider a single frequency and then generalize to multiple frequencies. We aim to choose a set of N measurement times $\mathbf{t} = (t_1, \ldots, t_N)$ such that

$$\mathbf{t}^* = \underset{\mathbf{t}\in[0,1]^N}{\operatorname{argmax}} \min_{\phi\in[0,2\pi)} \gamma(\mathbf{t}; A, f, \phi), \tag{9}$$

in which $\gamma(\mathbf{t}; A, f, \phi)$ is the power function given in Eq. (3). We first show that the optimization problem in Eq. (9) is equivalent to maximizing the noncentrality parameter from Eq. (4).

Proposition 3.2. *The noncentral F distribution $F_\lambda(x; n_1, n_2)$ with degrees of freedom $n_1 = 2$ and $n_2 = N - 3 > 0$ is a monotone decreasing function of the noncentrality parameter λ.*

Proof. Use Prop. 5.1 to express $F(x; n_1, n_2, \lambda)$ as

$$F(x; n_1, n_2, \lambda) = \sum_{k=0}^{\infty} T_k(z, n_1, n_2, \lambda)$$

in which $z = \frac{n_1 x}{n_1 x + n_2}$. Since $n_2 = N - 3 > 0$ by assumption, it follows that $z \in [0, 1]$ and we may apply Prop. 5.2 and Prop. 5.3 to justify term-wise differentiation of the series. Differentiate term-wise and apply the same reasoning as in the proof of Prop. 5.3 to find

$$\begin{aligned}\frac{\partial}{\partial\lambda} F(x; n_1, n_2, \lambda) &\leq -T_0(z; n_1, n_2, \lambda) + \sum_{k=1}^{\infty} \frac{1}{2}\left(-T_k(z; n_1, n_2, \lambda) + T_{k-1}(z; n_1, n_2, \lambda)\right) \\ &= \lim_{N\to\infty} -\frac{1}{2} T_N(z; n_1, n_2, \lambda) \leq 0,\end{aligned}$$

which verifies that $F(x; n_1, n_2, \lambda)$ is monotone decreasing for all $\lambda \geq 0$. □

It follows from Prop. 3.2 that the optimal $\mathbf{t}^*$ appearing in Eq. (9) can be found by simply maximizing the noncentrality parameter. A connection between the non-centrality parameter and the reduced design matrix further simplifies the problem.

Proposition 3.3. *Maximizing the noncentrality parameter uniformly over acrophase is equivalent to maximizing the smallest singular value of the reduced design matrix $\tilde{X} = [\cos(2\pi f \mathbf{t}) \;\; \sin(2\pi f \mathbf{t})]$.*

Proof. The parameter λ can be expressed as

$$\lambda(\mathbf{t}; A, f, \phi) = A^2 \sum_{i=1}^{N} \cos^2(2\pi f t_i - \phi) = ||\tilde{X}[\beta_1 \;\; \beta_2]^T||_2^2. \tag{10}$$

We reparameterize Eq. (10) to obtain

$$\max_{\mathbf{t}\in[0,1]^N} \min_{\phi\in[0,2\pi)} \lambda(\mathbf{t}; A, f, \phi) = \max_{\mathbf{t}\in[0,1]^N} \min_{\beta_1^2+\beta_2^2=A^2} ||\tilde{X}[\beta_1 \;\; \beta_2]^T||_2^2.$$

By homogeneity, we may assume $A^2 = 1$, reducing our problem to finding

$$\mathbf{t}^* = \underset{\mathbf{t}\in[0,1]^N}{\operatorname{argmax}} \min_{\beta_1^2+\beta_2^2=1} ||\tilde{X}[\beta_1 \;\; \beta_2]^T||_2^2. \tag{11}$$

The min-max principle for the singular values of a matrix $A \in \mathbb{C}^{n\times 2}$ is

$$\sigma_i(A) = \max_{\dim(U)=i} \min_{x\in U, ||x||=1} ||Ax||. \tag{12}$$

Here $\sigma_2 \leq \sigma_1$, so $i = 2$ gives the smallest singular value. When $i = 2$, the right hand side of Eq. (11) agrees with the square of the right hand side of Eq. (12). □

To further simplify the optimization problem, notice that $\tilde{M} = \tilde{X}^T \tilde{X}$ is symmetric and therefore has real eigenvalues.

Corollary 3.1.1. *Since $\sigma_2^2(\tilde{X}) = \lambda_{\min}(\tilde{M})$, maximizing power uniformly in acrophase is equivalent to maximizing the smallest eigenvalue of the reduced Fisher information matrix $\tilde{M}$.*

The remaining steps in reducing our optimization problem are motivated from a computational perspective. Corollary 3.1.1 reveals the convex structure of the problem, since maximizing the smallest eigenvalue of a symmetric positive semi-definite matrix is a concave programming problem [15]. We simplify further so that the problem can be solved using standard convex programming packages. We also assume for the remainder of the proof that the decision variables have been restricted to a discrete grid, as discussed in Sec. 2.2.

Since $\tilde{M}$ is low-dimensional, we can use an exact expression for its smallest eigenvalue and convert eigenvalue maximization to a quadratic programming problem.

Proposition 3.4. *A weight function $\boldsymbol{\mu}^* \in \{0,1\}^K$ on a partition $\boldsymbol{\tau} \in [0,1]^K$ satisfies*

$$\boldsymbol{\mu}^* = \underset{\boldsymbol{\mu} \in \{0,1\}^K}{\arg\max}\, \lambda_{\min}(M(\boldsymbol{\mu}; \boldsymbol{\tau}, \mathbf{f})),$$

if and only if

$$\boldsymbol{\mu}^* = \underset{\boldsymbol{\mu} \in \{0,1\}^K}{\arg\max}\, \mathbf{b}^T \boldsymbol{\mu} - \sqrt{\boldsymbol{\mu}^T A \boldsymbol{\mu}}, \tag{13}$$

in which $M(\boldsymbol{\mu}; \boldsymbol{\tau}, \mathbf{f})$ is the Fisher information matrix associated to the bi-variate linear model

$$y(t) = \boldsymbol{\beta}^T \mathbf{f}(t) = \beta_1 f_1(t) + \beta_2 f_2(t),$$

A is the symmetric positive semi-definite matrix

$$\begin{aligned} \mathbf{b} &= \mathbf{a}_{11} + \mathbf{a}_{22}, \\ A &= \mathbf{a}_{11}\mathbf{a}_{11}^T + \mathbf{a}_{22}\mathbf{a}_{22}^T - \left(\mathbf{a}_{11}\mathbf{a}_{22}^T + \mathbf{a}_{22}\mathbf{a}_{11}^T\right) + 4\mathbf{a}_{12}\mathbf{a}_{12}^T, \\ \mathbf{a}_{11} &= f_1(\boldsymbol{\tau}) \circ f_1(\boldsymbol{\tau}), \\ \mathbf{a}_{12} &= f_1(\boldsymbol{\tau}) \circ f_2(\boldsymbol{\tau}), \\ \mathbf{a}_{22} &= f_2(\boldsymbol{\tau}) \circ f_2(\boldsymbol{\tau}), \end{aligned} \tag{14}$$

and $\circ$ denotes the Hadamard product.

Proof. The smallest eigenvalue is given by

$$\boldsymbol{\Lambda}_{\min}(M(\boldsymbol{\mu}; \boldsymbol{\tau}, \mathbf{f})) = \frac{1}{2}\left(\operatorname{tr}\left(M(\boldsymbol{\mu}; \boldsymbol{\tau}, \mathbf{f})\right) - \sqrt{\operatorname{tr}\left(M(\boldsymbol{\mu}; \boldsymbol{\tau}, \mathbf{f})\right)^2 - 4\det(M(\boldsymbol{\mu}; \boldsymbol{\tau}, \mathbf{f}))}\right). \tag{15}$$

$M(\boldsymbol{\mu}; \boldsymbol{\tau}, \mathbf{f})$ can be expressed as

$$M(\boldsymbol{\mu}; \boldsymbol{\tau}, \mathbf{f}) = \begin{pmatrix} \langle \boldsymbol{\mu}, \mathbf{a}_{11} \rangle & \langle \boldsymbol{\mu}, \mathbf{a}_{12} \rangle \\ \langle \boldsymbol{\mu}, \mathbf{a}_{12} \rangle & \langle \boldsymbol{\mu}, \mathbf{a}_{22} \rangle \end{pmatrix}$$

so the trace is given by

$$\operatorname{tr}\left(M(\boldsymbol{\mu}; \boldsymbol{\tau}, \mathbf{f})\right) = \langle \boldsymbol{\mu}, \mathbf{a}_{11} + \mathbf{a}_{22} \rangle \tag{16}$$

Replacing the trace terms in Eq. (15) with Eq. (16) gives

$$\begin{aligned}\boldsymbol{\Lambda}_{\min}(M(\boldsymbol{\mu};\boldsymbol{\tau},\mathbf{f})) &= \frac{1}{2}\left(\langle\boldsymbol{\mu},\mathbf{a}_{11}+\mathbf{a}_{22}\rangle - \sqrt{\langle\boldsymbol{\mu},\mathbf{a}_{11}+\mathbf{a}_{22}\rangle^2 - 4\left(\langle\boldsymbol{\mu},\mathbf{a}_{11}\rangle\langle\boldsymbol{\mu},\mathbf{a}_{22}\rangle - \langle\boldsymbol{\mu},\mathbf{a}_{12}\rangle^2\right)}\right)\\ &= \frac{1}{2}\left(\langle\boldsymbol{\mu},\mathbf{a}_{11}+\mathbf{a}_{22}\rangle - \sqrt{\langle\boldsymbol{\mu},\mathbf{a}_{11}\rangle^2 + \langle\boldsymbol{\mu},\mathbf{a}_{22}\rangle^2 - 2\langle\boldsymbol{\mu},\mathbf{a}_{11}\rangle\langle\boldsymbol{\mu},\mathbf{a}_{22}\rangle + 4\langle\boldsymbol{\mu},\mathbf{a}_{12}\rangle^2}\right)\\ &= \frac{1}{2}\left(\langle\boldsymbol{\mu},\mathbf{a}_{11}+\mathbf{a}_{22}\rangle - \sqrt{\boldsymbol{\mu}^T A\boldsymbol{\mu}}\right)\end{aligned}$$

in which A is as given in Eq. (14). Finally, A is a symmetric positive semidefinite matrix by Prop. 5.4. □

Corollary 3.1.2. *For the cosinor model, the trace of $M(\boldsymbol{\mu};\boldsymbol{\tau},\mathbf{f})$ obeys*

$$\sum_{i=1}^{K}\mu_i\left(\sin(2\pi f\tau_i)^2 + \cos(2\pi f\tau_i)^2\right) = N$$

and the optimization problem in Eq. (13) is reduced to

$$\boldsymbol{\mu}^* = \underset{\boldsymbol{\mu}\in\{0,1\}^K}{\arg\min}\ \boldsymbol{\mu}^T A\boldsymbol{\mu}. \tag{17}$$

Equation (17) is an example of an integer-constrained quadratic programming problem. When considering a range of candidate frequencies (f_1, f_2) simultaneously, Eq. (17) generalizes to

$$\boldsymbol{\mu}^* = \underset{\boldsymbol{\mu}\in\{0,1\}^K}{\arg\max}\ \min_{f\in(f_1,f_2)}\boldsymbol{\mu}^T A_f\boldsymbol{\mu}, \tag{18}$$

in which A_f is the quadratic form of the previous section corresponding to frequency f. To convert Eq. (18) to a form suitable for numerical study, we make use of Prop. 5.1 to obtain a linear objective with quadratic constraints. In this final form, the problem can be studied numerically using standard solvers such as [16] or open-source alternatives [17]. The same conversion is applicable to the regularized problem.

3.4 Numerical evidence for optimality in the regularization-only problem

In this section, we investigate numerical solutions to the regularization-only problem

$$(\boldsymbol{\mu}^*, s^*) = \underset{\boldsymbol{\mu}\in\{0,1\}^M,\sum_{i=1}^M\mu_i=N,s\in\mathbb{R}}{\arg\min}\left\{s \mid \boldsymbol{\mu}^T R\boldsymbol{\mu} \le s\right\}. \tag{19}$$

The optimization problem in Eq. (19) has fewer quadratic constraints than the problem in Eq. (8) because the frequency averaging has been carried out analytically. This difference in structure makes easier to find numerical evidence for the optimality of candidate solutions and affects the degenerate critical points symmetry of the problem (Fig. 6)

When solving an MIQCP, it is common to impose a time limit for the computation. In some cases, the solver will terminate early because the value of the objective function at the candidate solution is sufficiently close to the optimal solution to the continuous relaxation. When this occurs, it provides numerical evidence for the optimality of the candidate solution. This type of numerical evidence was occasionally obtained for the balanced designs in the MIQCP studied here, but in most cases, the solver would not terminate until it reached the time limit.

We identified large regions of parameter space where the regularization-only problem halts because the solver has produced evidence of optimality (Fig. 7). The optimal solutions to the problem have a structure that depends on the boundaries of the frequency prior. Dense measurements are

collected early in the experiment and depending on the parameters of the prior, sparser measurements may be collected near the end (Fig. 7A). Varying the measurement budget together with the boundaries of the prior reveals regions of parameter space where irregular measurements are optimal (Fig. 7B). For a fixed measurement budget, there is a threshold above which equispaced measurements again appear to be optimal. It should also be noted that the coarseness of the underlying partition is of central importance in producing evidence for optimality. Determining if the evidence of optimality identified here remains valid at finer partitions remains to be understood.

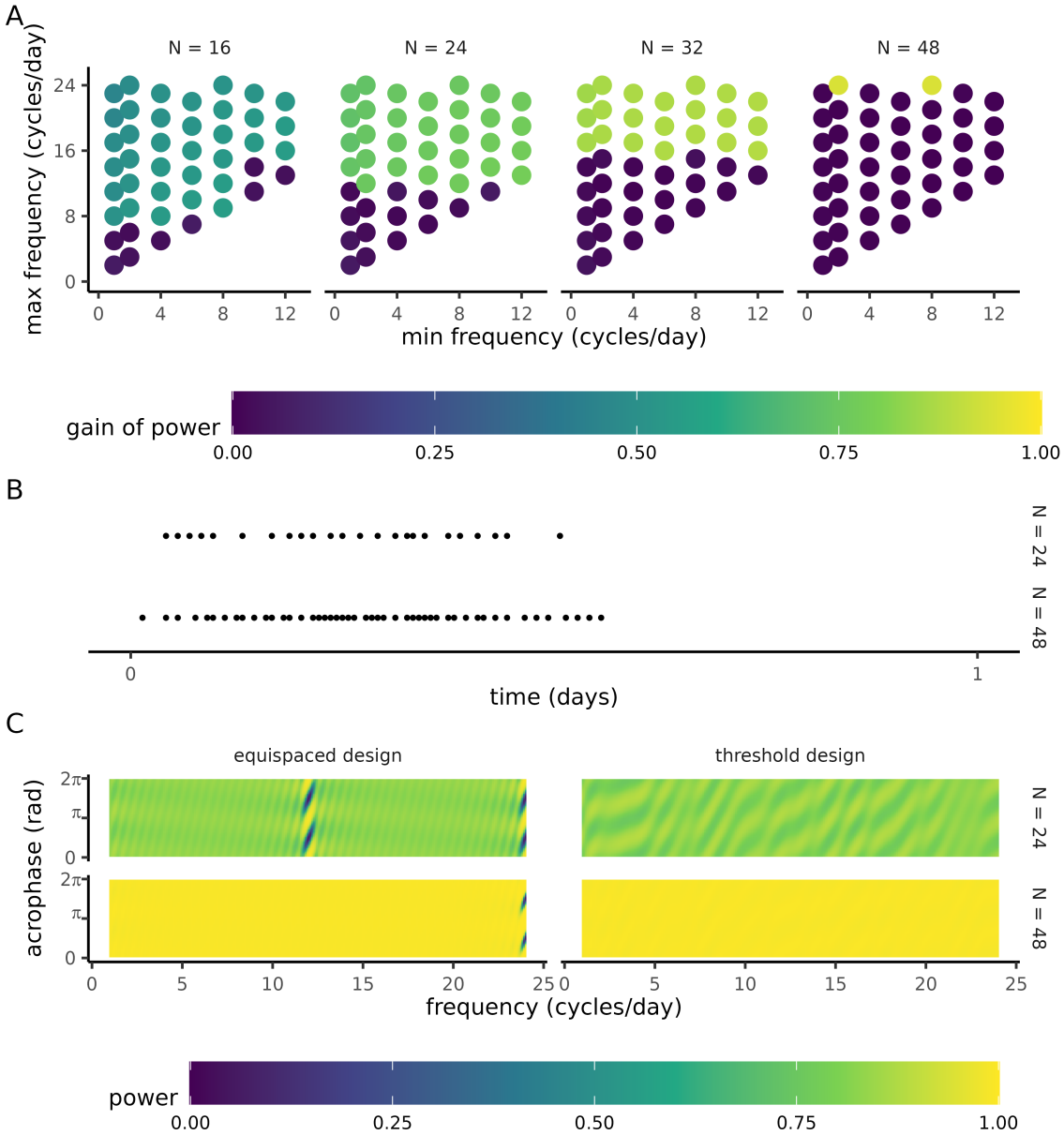

Figure 2: **Benefits of irregular sampling are dictated by period uncertainty and measurement budget.** **(A)** Threshold designs were generated for uniform frequency priors $f \sim \text{Unif}(f_{\min}, f_{\max})$ and various measurement budgets. Each circle represents the difference in worst-case power between a threshold design constructed by our method and the corresponding equispaced design. Comparing the gain for different values of $f_{\min}$ and $f_{\max}$ for each budget reveals an uncertainty threshold above which irregular sampling becomes worthwhile. As the budget increases, a greater amount of frequency uncertainty can be tolerated before irregular sampling becomes substantially beneficial. **(B)** Examples of threshold designs for a high amount of frequency uncertainty ($f_{\min} = 1$,$f_{\max} = 24$). Even at high budgets, the measurements in threshold designs tend to cluster in a sub-portion of the experiment. **(C)** Influence of frequency and acrophase on power for signals of a fixed amplitude ($A = 1$). The improvements in power achieved by the threshold design come at the expense of introducing weak power oscillations throughout the acrophase-frequency plane. At a high budget, these effects are negligible suggesting that threshold designs should only be considered when the budget is sufficiently large.

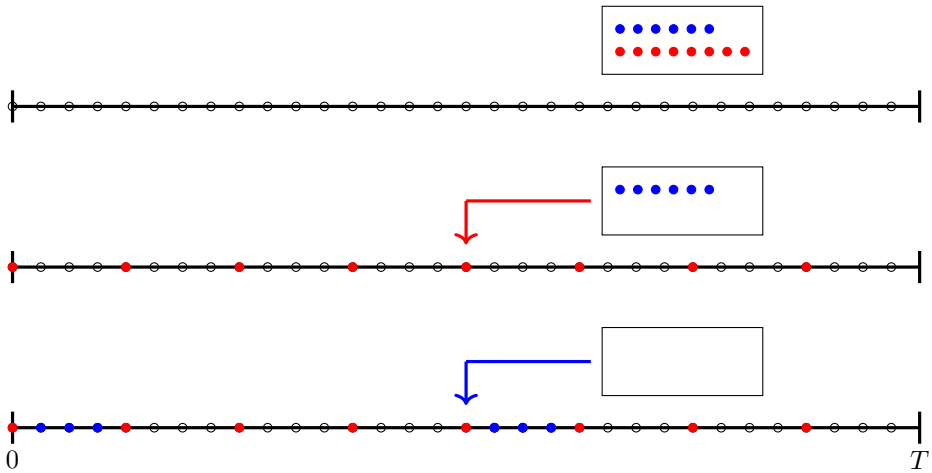

Figure 3: **Procedure to construct designs with equispaced constraints (Top)** The duration of the experiment is cut into a fine grid indicated by the open circles. The user splits their measurement budget into two groups, indicated here by color. **(Middle)** Measurements in the red group are assigned to a uniform grid over the entire duration of the experiment. **(Bottom)** The remaining measurements are assigned to the unoccupied times in the grid in such a way that the worst-case power is maximized. The result is a design where measurements are to be collected at both the blue and red time points.

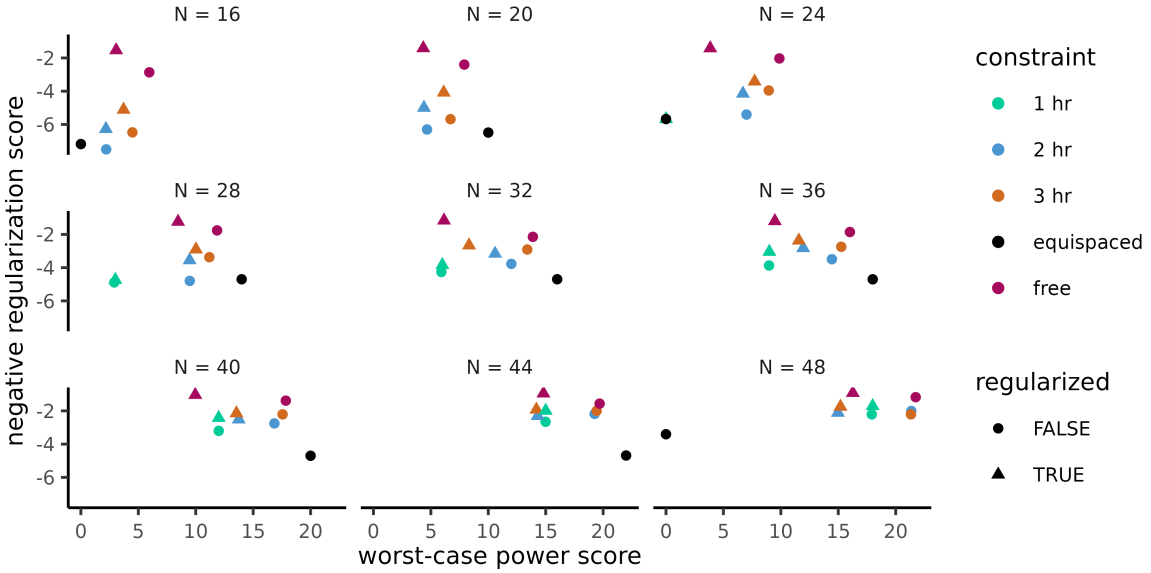

Figure 4: **Tradeoff between worst-case power and power variability** We investigate the tradeoff between worst-case power and power variability assuming a fixed amount of frequency uncertainty ($f_{\min} = 1$, $f_{\max} = 24$) for various measurement budgets and types of optimized design. Constraints mainly affect worst-case power and not the regularization score. The best trade-off is generally achieved by designs that do not adhere to equispaced measurement constraints

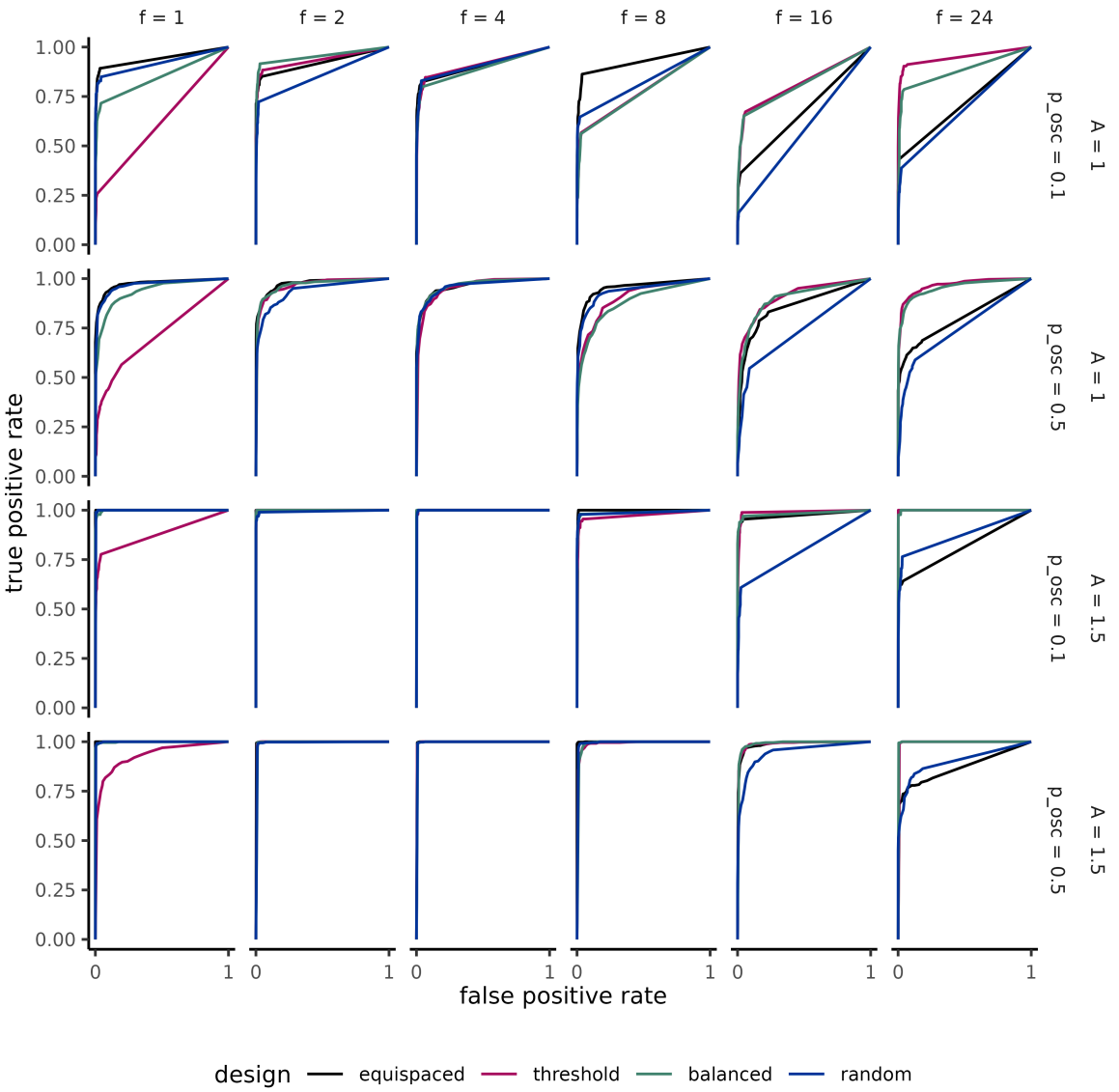

Figure 5: **Periodogram analysis with irregular sampling detects simulated oscillations** Each panel represents a distinct dataset made up of homoscedastic noise and cosinor signals as in Eq. (1) with acrophases distributed uniformly at random. The parameter p_{osc} dictates the probability that a randomly generated signal in the dataset will be oscillatory. Signals are classified as oscillatory if their periodogram produces a statistically significant p-value after FDR correction. The extent to which oscillations are reliably detected is influenced by their abundance (p_{osc}), as well as their frequency and amplitude. Detection is summarized visually using an receiver operator characteristic curve for each design. The equispaced design fails to detect high frequency oscillations, whereas the threshold design performs favorably at high frequencies but poorly at low frequencies. The balanced design performs comparably well at all frequencies of interest. For reference, a design points distributed uniformly at random is also included and performs generally worse than the optimized designs.

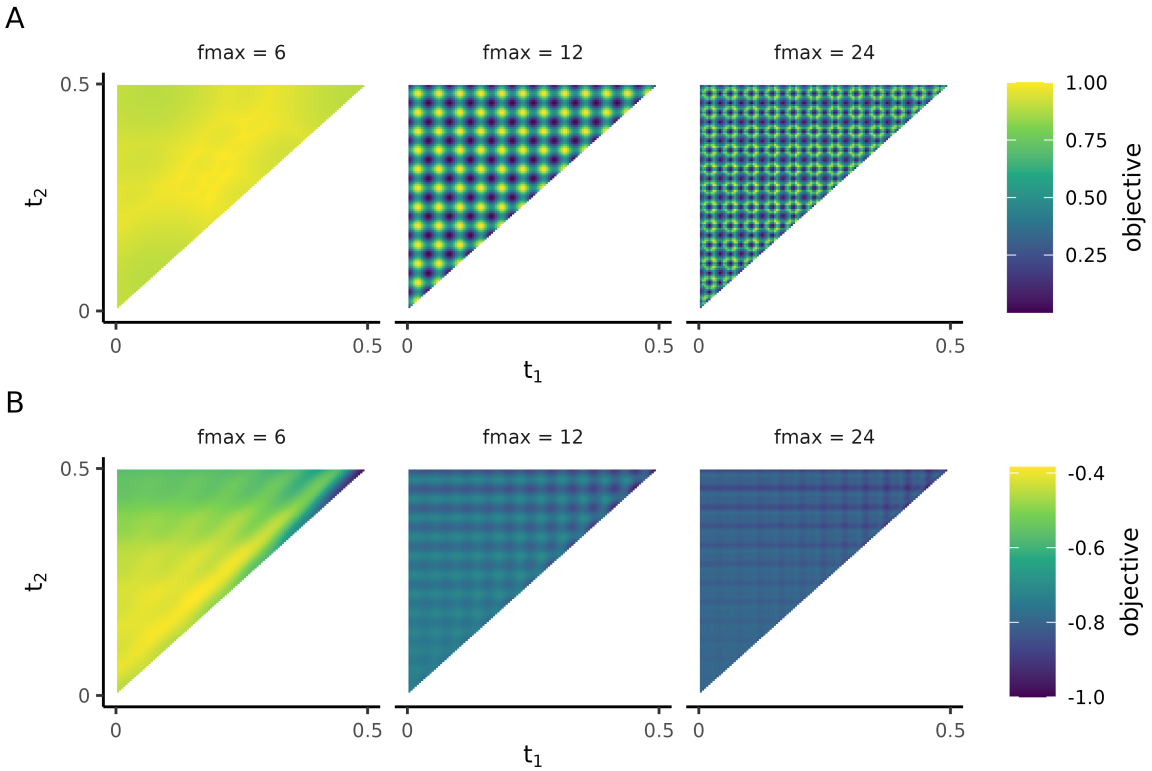

Figure 6: **Influence of regularization on fixed-phase worst-case power maximization (A)** An experiment aims to detect oscillations of known phase $\phi = 0$ and unknown frequency using $N = 10$ equally spaced measurements. Two additional measurements are to be added at times t_1 and t_2. (Left) If the frequency prior is tight, power is high and the additional two measurements have little influence on the objective. (Middle) As the upper limit of the frequency prior approaches the Nyquist rate, the extra two measurement times have a stronger influence. Importantly, the objective function has several local maxima, illustrating the non-uniqueness. (Right) At a higher level of frequency uncertainty, the objective function oscillates faster and becomes increasingly degenerate as a function of the two additional measurement times. **(B)** After adding regularization to the objective, there emerges a preference for earlier measurement times, reducing the degeneracy in the problem.

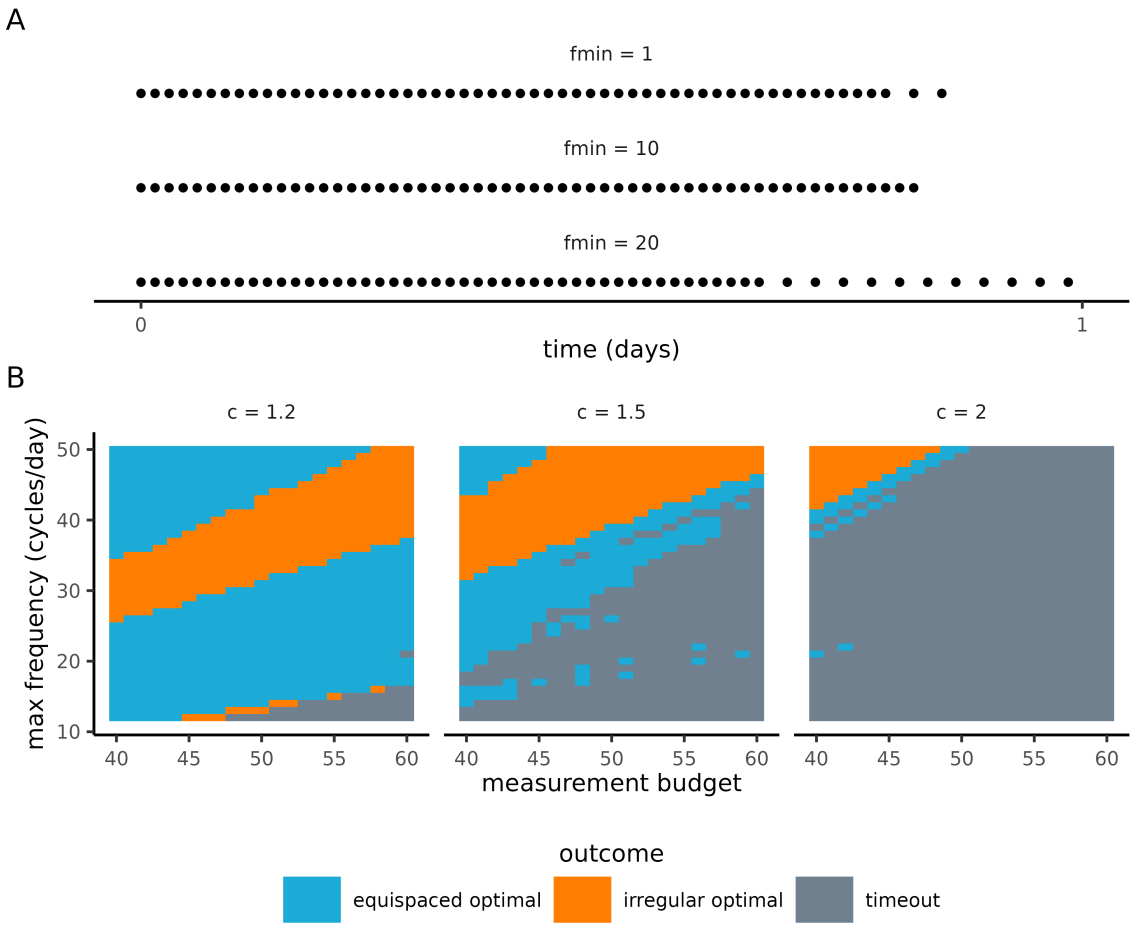

Figure 7: **Regularization favors irregular measurements.** The pure regularization problem corresponds to minimizing the mean squared norm of the frequency-acrophase gradient of the non-centrality parameter. This problem is parameterized by the frequency window under consideration, measurement budget, and spatial resolution of the underlying partition. The number of points in the partition is given by $N_{\text{fine}} = \lfloor cN \rfloor$ for a constant $c > 0$. **(A)** Optimal designs for various choices of $f_{\min}$ with all other parameters held fixed ($f_{\max} = 45$, $N = 56$, $N_{\text{fine}} = 67$). Solutions are similar in structure: measurements are collected as rapidly as possible near the start of the experiment and then, depending on the value of $f_{\min}$, a portion of the budget is allocated to slightly sparser measurements. **(B)** To investigate how discretization affects the optimality, we studied different ratios between measurement budget and discretization ($f_{\min} = 1$). Depending on the budget, discretization, and frequency prior, the MIQCP solver can show that a particular type of sampling is optimal or fail to prove optimality within the provided computation time. Comparing the panels reveals a discretization-dependent region in the frequency-budget plane where irregular sampling is optimal.

4 Discussion

Our work has identified conditions where irregular measurement schedules significantly improve the statistical power of rhythm detection in the presence of parameter uncertainty. We provide an efficient numerical method for optimizing the power in such settings by reformulating worst-case power optimization as an MIQCP. We add structure to the optimization problem through regularization and equispaced constraints. These features simplify the designs without significantly worsening their performance, making them more applicable to realistic experiments. Application of our designs to a periodogram-based hypothesis test demonstrates that they remain applicable in statistical frameworks beyond the cosinor model. Focusing on the regularization-only problem, we obtained numerical evidence for optimality of irregular designs at various degrees of frequency uncertainty.

Several immediate extensions of our method could be considered in future work. First, many studies are constrained by windows of time in which measurements cannot be collected. Such constraints could be added to our method without altering its structure. An alternative approach accommodates such constraints using weighted least-squares [18] and could potentially be combined with our framework to allow for simultaneous optimization of measurement times and weights. Second, statistical power is not the only figure of merit relevant to reliable design of rhythm detection experiments. Depending on the goals of the study, it may be important to understand how the design of the experiment influences the accuracy of period estimation. Various periodogram and spectral techniques are available, many of which have a long history of application to irregularly sampled data [14, 19, 13, 20]. Explicitly treating spectral accuracy as an additional figure of merit could further improve usefulness. Third, our method could be applied to adaptive experimental design where the prior distributions are updated using Bayesian inference. Closed form expressions for the regularization kernel would no longer be available, but nothing would be fundamentally changed in making the method adaptive. If the method were extended to other priors, it would be worthwhile to compare the uniform priors considered here to discrete priors made up of Dirac masses centered at integer frequencies. This would still be analytically tractable and could lead to qualitative differences in the optimal solutions of the problem.

The assumption of homoscedastic Gaussian white noise and the cosinor-based hypothesis test are central to our method but also limit its applicability. Extending the method to more general statistical frameworks and understanding its relation to methods in nonlinear optimal experimental design [21] both provide promising directions for future research. Another meaningful extension to the method could come from treating the number of replicates at each time-point as a decision variable in the optimization problem. This is common in optimal experimental design and could provide additional benefits to the statistical power. Through continued investigation in settings including and beyond power analysis, we hope that irregular sampling techniques may join the vast array of tools available to experimental biologists.

5 Methods

The following propositions are used in Prop. 3.2 to verify that the noncentral F-distribution depends monotonically on the noncentrality parameter .

Proposition 5.1 (Series representation of noncentral F-distribution)**.** *The cumulative distribution function $F(x; n_1, n_2, \lambda)$ for a noncentral F-distributed random variable with degrees of freedom n_1, n_2 and noncentrality parameter λ can be expressed as a series*

$$F(x; n_1, n_2, \lambda) = \sum_{k=0}^{\infty} \left(\frac{\left(\frac{\lambda}{2}\right)^k}{k!} e^{-\lambda/2} \right) I\left(\frac{n_1 x}{n_2 + n_1 x}; \frac{n_1}{2} + k, \frac{n_2}{2} \right), \tag{20}$$

in which $I(z;a,b)$ is the regularized incomplete beta function, defined by

$$B(z;a,b) := \int_0^z t^{a-1}(1-t)^{b-1}\ \mathrm{d}t,$$
$$I(z;a,b) := B(z;a,b)/B(1;a,b). \tag{21}$$

Proof. A derivation of this formula is given in [22]. □

Proposition 5.2. *For $z \in [0,1]$ and $a > 0, b > 0$, the regularized incomplete beta functions obey*

- $I(z;a+1,b) \leq I(z,a,b)$,
- $\lim_{a\to\infty} I(z;a+1,b)/I(z;a,b) = z$.

Proof. The coefficient $B(1,a,b)$ in Eq. (21) can be expressed as

$$B(1,a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$$

in which $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}\mathrm{d}t$ is the gamma function [23]. Use integration by parts and the identity $\Gamma(z+1) = z\Gamma(z)$ to obtain

$$I(z;a+1,b) = I(z;a,b) - \frac{z^a(1-z)^b}{aB(1;a,b)}. \tag{22}$$

Since $z \in [0,1]$ the second term in Eq. (22) is of definite sign, hence

$$I(z;a+1,b) \leq I(z;a,b).$$

To obtain the asymptotic result, rearrange Eq. (22) and integrate by parts to obtain

$$\begin{aligned}\frac{I(z;a+1,b)}{I(z;a,b)} &= 1 - \frac{z^a(1-z)^b}{a\int_0^z x^{a-1}(1-x)^{b-1}\mathrm{d}x} \\ &= 1 - \frac{z^a(1-z)^b}{z^a(1-z)^{b-1} + (b-1)\int_0^z x^a(1-x)^{b-2}\mathrm{d}x} \\ &= 1 - \frac{1-z}{1 + \frac{(b-1)}{(1-z)^{b-1}}\int_0^z \left(\frac{x}{z}\right)^a (1-x)^{b-2}\mathrm{d}x}.\end{aligned}$$

Since $\left(\frac{x}{z}\right)^a < 1$ for $x < z$, and for $b > 0$,$z \in [0,1]$, the following integral is finite

$$\int_0^z (1-x)^{b-2}\mathrm{d}z = \begin{cases} -\log(1-z) & b = 1 \\ \frac{1-(1-z)^{b-1}}{b-1} & \text{otherwise} \end{cases},$$

we can evaluate the desired limit using the dominated convergence theorem to obtain

$$\lim_{a\to\infty} \int_0^z \left(\frac{x}{z}\right)^a (1-x)^{b-2}\mathrm{d}x = 0$$

and so $\lim_{a\to\infty} I(z;a+1,b)/I(z;a,b) = z$. □

Proposition 5.3. *Let $T_k(z,n_1,n_2,\lambda)$ be the terms of the series expansion in Eq. (20) in which $z = \frac{n_1 x}{n_1 x + n_2}$ and $\lambda > 0$. The terms $T_k(z,n_1,n_2,\lambda)$ satisfy the following:*

- $\sum_{k=0}^\infty T_k(z;n_1,n_2,\lambda)$ *converges,*

- $\sum_{k=0}^{\infty} \partial_\lambda T_k(z, n_1, n_2, \lambda)$ *converges uniformly for all* $\lambda > 0$.

Proof. Convergence of $\sum_{k=0}^{\infty} T_k(z, n_1, n_2, \lambda)$ follows from the ratio test

$$\frac{T_{k+1}(z, n_1, n_2, \lambda)}{T_k(z, n_1, n_2, \lambda)} = \frac{I_{k+1}}{I_k} \frac{\lambda}{2(k+1)} \leq \frac{\lambda}{2(k+1)},$$

which implies $\lim_{k\to\infty} \left| \frac{T_{k+1}(z,n_1,n_2,\lambda)}{T_k(z,n_1,n_2,\lambda)} \right| = 0$ for any $\lambda > 0$. Uniform convergence of the derivatives can be verified using the Weierstrass M-test. To construct a majorant for $T_k(z, n_1, n_2, \lambda)$, notice that $\partial_\lambda T_k$ has critical points at $\lambda = 2k$ and $\lambda = 0$. We know that $T_k(z, n_1, n_2, 0) = 0$ and $\lim_{\lambda\to\infty} T_k(z, n_1, n_2, \lambda) = 0$ so

$$g_k(z, n_1, n_2) := T_k(z, n_1, n_2, 2k) = I_k \frac{k^k}{k!} e^{-k}$$

must be the global maximum of $T_k(z, n_1, n_2, \lambda)$ on the interval $\lambda \in [0, \infty)$. This majorant also leads to a bound on the derivative, since $|\partial_\lambda T_0| \leq \frac{1}{2} T_0$ and for $k > 0$

$$\begin{aligned}
|\partial_\lambda T_k| &= \left| -\frac{1}{2} T_k + e^{-\lambda/2} \frac{\lambda^{k-1}}{(k-1)! 2^k} I\left(z; \frac{n_1}{2} + k, \frac{n_2}{2}\right) \right| \\
&\leq \frac{1}{2} T_k + \frac{1}{2} e^{-\lambda/2} \frac{\lambda^{k-1}}{(k-1)! 2^{k-1}} I\left(z; \frac{n_1}{2} + k, \frac{n_2}{2}\right) \\
&\leq \frac{1}{2} T_k + \frac{1}{2} e^{-\lambda/2} \frac{\lambda^{k-1}}{(k-1)! 2^{k-1}} I\left(z; \frac{n_1}{2} + k - 1, \frac{n_2}{2}\right) \\
&= \frac{1}{2} \left(T_k + T_{k-1}\right).
\end{aligned}$$

To prove uniform convergence of $\sum_k \partial_\lambda T_k$, it remains to show that $\lim_{k\to\infty} \left| \frac{g_{k+1}}{g_k} \right| < 1$. Indeed, we have

$$\frac{g_{k+1}}{g_k} = \frac{I_{k+1}}{I_k} \frac{(k+1)^{k+1}}{k^k} \frac{k!}{(k+1)!} e^{-1} = \frac{I_{k+1}}{I_k} \left(\frac{k+1}{k}\right)^k e^{-1} = \frac{I_{k+1}}{I_k} e^{-1} \left(1 + \frac{1}{k}\right)^k.$$

We know $\lim_{k\to\infty} \left(1 + \frac{1}{k}\right)^k = e$ and so

$$\lim_{k\to\infty} \left| \frac{g_{k+1}}{g_k} \right| = \lim_{k\to\infty} \frac{I_{k+1}}{I_k} = z < 1,$$

with the last equality justified by the second claim of Prop. 5.3. □

The following proposition is useful for verifying convexity of the reformulation of power optimization in terms of the de-meaned Fisher information matrix.

Proposition 5.4. *For any* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ *the matrix*

$$C = \mathbf{u}\mathbf{u}^T + \mathbf{w}\mathbf{w}^T - \left(\mathbf{u}\mathbf{w}^T + \mathbf{w}\mathbf{u}^T\right) + 4\mathbf{v}\mathbf{v}^T \tag{23}$$

is symmetric and positive semidefinite.

Proof. From Eq. (23) it is clear that $C = C^T$, notice that C can be expressed as a positive combination of rank 1 matrices

$$C = (\mathbf{u} - \mathbf{w})(\mathbf{u} - \mathbf{w})^T + 4\mathbf{v}\mathbf{v}^T,$$

and thus C is positive semidefinite. □

The main goal of our method is to maximize the statistical power across a range of parameters. This direct approach can produce low-amplitude power variations in the frequency-phase plane. The next proposition shows how we can use L^2 regularization to remove this unwanted variability without altering the convexity of the problem.

Proposition 5.5. *Given*

- *a partition* $\boldsymbol{\tau} \subset [0,1]$ *of size* K *and weight function* $\boldsymbol{\mu} \in \{0,1\}^K$,
- *amplitude* A, *noise strength* σ, *frequency* f, *and acrophase* ϕ,
- *a prior distribution* $\rho : \mathbb{R}_{>0} \times [0, 2\pi) \to \mathbb{R}$ *on* (f, ϕ),

let $\lambda(\boldsymbol{\mu})$ *be the associated noncentrality parameter*

$$\lambda(\boldsymbol{\mu}) = \lambda(\boldsymbol{\mu}; A, \sigma, f, \phi, \tau) = \frac{A^2}{\sigma^2} \sum_{i=1}^{K} \mu_i \cos^2(2\pi f \tau_i - \phi) = \boldsymbol{\mu}^T \boldsymbol{\eta}(\boldsymbol{\tau})$$

in which and $\boldsymbol{\eta}(\boldsymbol{\tau}) = \frac{A^2}{\sigma^2} \cos^2(2\pi f \boldsymbol{\tau} - \phi)$. *The* L^2 *norm of the phase-frequency gradient of* $\lambda(\boldsymbol{\mu})$ *can be expressed as*

$$||\nabla_{f,\phi} \lambda||^2_{L^2_{f,\phi}} = \boldsymbol{\mu}^T R \boldsymbol{\mu}$$

in which $R \in \mathbb{R}^{K \times K}$ *is the symmetric positive semidefinite matrix given by*

$$R = \iint \left(\partial_f \boldsymbol{\eta}(\boldsymbol{\tau}) \partial_f \boldsymbol{\eta}(\boldsymbol{\tau})^T + \partial_\phi \boldsymbol{\eta}(\boldsymbol{\tau}) \partial_\phi \boldsymbol{\eta}(\boldsymbol{\tau})^T \right) \rho(f, \phi) \, \mathrm{d}f \mathrm{d}\phi, \tag{24}$$

Proof. Since $\lambda(\boldsymbol{\mu}) = \boldsymbol{\mu}^T \boldsymbol{\eta}(\boldsymbol{\tau})$, the gradient can be written as

$$\nabla_{f,\phi} \lambda(\boldsymbol{\mu}) = \begin{pmatrix} \boldsymbol{\mu}^T \partial_f \boldsymbol{\eta}(\boldsymbol{\tau}) \\ \boldsymbol{\mu}^T \partial_\phi \boldsymbol{\eta}(\boldsymbol{\tau}) \end{pmatrix},$$

from which it follows that

$$\begin{aligned} ||\nabla_{f,\phi} \lambda(\boldsymbol{\mu})||^2_{L^2_{f,\phi}} &= \iint \left(\nabla_{f,\phi} \lambda(\boldsymbol{\mu}) \right)^T \left(\nabla_{f,\phi} \lambda(\boldsymbol{\mu}) \right) \rho(f, \phi) \mathrm{d}f \mathrm{d}\phi \\ &= \boldsymbol{\mu}^T \left(\iint \partial_f \boldsymbol{\eta}(\boldsymbol{\tau}) \partial_f \boldsymbol{\eta}(\boldsymbol{\tau})^T + \partial_\phi \boldsymbol{\eta}(\boldsymbol{\tau}) \partial_\phi \boldsymbol{\eta}(\boldsymbol{\tau})^T \rho(f, \phi) \mathrm{d}f \mathrm{d}\phi \right) \boldsymbol{\mu}. \end{aligned} \tag{25}$$

Symmetry of Q is apparent from its definition, and the requirement $\mathbf{x}^T Q \mathbf{x} \geq 0$ for any $\mathbf{x} \in \mathbb{R}^K$ follows from linearity of the integral in Eq. (25). □

Corollary 5.0.1 (Sinc kernel for uniform prior)**.** *If the prior distribution is uniform*

$$\rho(f, \phi) = \begin{cases} ((f_{\max} - f_{\min}) 2\pi)^{-1} & \text{if} f \in [f_{\min}, \leq f_{\max}], \\ 0 & \text{otherwise}, \end{cases}$$

then Eq. (24) *simplifies to*

$$R_{ij} = \begin{cases} \left(4\tau_i \tau_j \pi^2 + 1\right) \left(\frac{\operatorname{sinc}(4\pi f_{\max}(\tau_i - \tau_j))}{\pi f_{\max}} - \frac{\operatorname{sinc}(4\pi f_{\min}(\tau_i - \tau_j))}{\pi f_{\min}} \right), & i \neq j \\ \frac{\pi(4\pi^2 \tau_i^2 + 1)}{f_{\max} - f_{\min}}, & i = j \end{cases}. \tag{26}$$

Throughout our numerical results, we assume that $\rho(f, \phi)$ is a uniform prior. In addition to simplifying the expression for the regularization term in Cor. 5.0.1, the compact support of the uniform prior allows us to perform routine manipulations of nested optimization problems in the sense given below.

Proposition 5.6. *For $X \times Y$ compact and $f : X \times Y \to \mathbb{R}$ continuous*

$$\inf_{x \in X} \inf_{y \in Y} f(x, y) = \inf_{(x,y) \in X \times Y} f(x, y).$$

Proof. By the extreme value theorem, we can define a function $y^* : X \to Y$ such that

$$\inf_{y \in Y} f(x, y) = f(x, y^*(x)).$$

Since $\{(x, y^*(x)) : x \in X\} \subset X \times Y$, it follows that

$$\inf_{x \in X} \inf_{y \in Y} f(x, y) = \inf_{x \in X} f(x, y^*(x)) \geq \inf_{(x,y) \in X \times Y} f(x, y).$$

To obtain the reverse inequality, again use the extreme value theorem to find $\hat{x}, \hat{y}$ such that

$$\inf_{(x,y) \in X \times Y} f(x, y) = f(\hat{x}, \hat{y}) \geq f(\hat{x}, y^*(\hat{x})) \geq \inf_{x \in X} \inf_{y \in Y} f(x, y).$$

□

The techniques discussed so far produce an integer-constrained optimization problem with convex relaxation. Solving this problem numerically requires a final simplification, whereby the cost function is simplified at the expense of adding additional constraints to the problem. This approach is an application of "disciplined convex programming" [24], an algorithmic framework for posing convex programming problems in a format suitable for numerical investigation. For our problem, the cost function is simple enough that the reduction can instead be carried out by hand using the techniques below.

Definition 5.1 (Epigraph)**.** For $f : X \to \mathbb{R}$, in which $X \subset \mathbb{R}^n$, the epigraph of f is

$$\text{epi} f = \{(\mathbf{x}, s) \in X \times \mathbb{R} : f(\mathbf{x}) \leq s\}$$

Proposition 5.7 (Graph representation)**.** *For any $x \in X \subset \mathbb{R}^n$ and $f : X \to \mathbb{R}$*

$$f(x) = \inf\{s \in \mathbb{R} : (\mathbf{x}, s) \in \text{epi} f\}.$$

Lemma 5.1 (Reduction to an MIQCP)**.** *For a family of continuous functions $\{f_i\}_{i=1}^N$ with $f_i : X \to \mathbb{R}$ and X compact, the optimization problem*

$$\mathbf{x}^* = \arg\min_{\mathbf{x} \in X} \max_{1 \leq i \leq N} f_i(\mathbf{x})$$

is equivalent to

$$(\mathbf{x}^*, s^*) = \arg\min_{(\mathbf{x}, s) \in X \times \mathbb{R}} \{s : s \geq f_1(x), \ldots, s \geq f_N(x)\} \tag{27}$$

Proof. Apply the graph representation to the function $(x_1, \ldots, x_N) \to \max(x_1, \ldots, x_N)$ to obtain

$$\max_{1 \leq i \leq N} f_i(\mathbf{x}) = \min_{s \in \mathbb{R}} \{s : s \geq f_i(\mathbf{x})\}$$

the claim follows by combining the nested minimization problems using Prop. 5.6. □

Corollary 5.1.1. *If $f_i(\mathbf{x}) = \mathbf{x}^T A_i \mathbf{x}$ and $\{A_1, \ldots, A_N\} \subset \mathbb{R}^{n \times n}$ are symmetric positive semidefinite matrices, then Eq. (27) is an MIQCP with convex relaxation.*

6 Acknowledgements

TS is supported by an NSERC Canada Graduate scholarship. This work was supported by the Krembil Foundation, Toronto, Canada, the Future Biomedicine Charity Fund, Vilnius, Lithuania, and by Research Council of Lithuania under the Programme "University Excellence Initiatives" of the Ministry of Education, Science and Sports of the Republic of Lithuania (No. 12-001-01-01-01 "Improving the Research and Study Environment"; project No S-A-UEI-23-10). AP is a Marius Jakulis Jason Foundation scholar. ARS acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC): RGPIN-2019-06946.

This research was enabled in part by support provided by Compute Ontario and the Digital Research Alliance of Canada. We give thanks to John Limanto for reviewing and providing feedback on this manuscript.

References

[1] Albert Goldbeter et al. Biochemical oscillations and cellular rhythms. *Biochemical Oscillations and Cellular Rhythms*, 1997.

[2] Valerii Vadimovich Fedorov. *Theory of Optimal Experiments*. New York: Academic Press, 1972.

[3] Friedrich Pukelsheim. *Optimal design of experiments*. SIAM, 2006.

[4] Wei Zong, Marianne L Seney, Kyle D Ketchesin, Michael T Gorczyca, Andrew C Liu, Karyn A Esser, George C Tseng, Colleen A McClung, and Zhiguang Huo. Experimental design and power calculation in omics circadian rhythmicity detection using the cosinor model. *Statistics in Medicine*, 42(18):3236–3258, 2023.

[5] CP Kitsos, DM Titterington, and B Torsney. An optimal design problem in rhythmometry. *Biometrics*, pages 657–671, 1988.

[6] Turner Silverthorne, Matthew Carlucci, Arturas Petronis, and Adam R Stinchcombe. Power optimization via convex programming for homoscedastic rhythm detection (Power-CHORD). https://github.com/t-silverthorne/paper_oscDetectionLimitedBudget, 2024.

[7] Germaine Cornelissen. Cosinor-based rhythmometry. *Theoretical Biology and Medical Modelling*, 11(1):1–24, 2014.

[8] Nathan Braniff. *Optimal Experimental Design Applied to Models of Microbial Gene Regulation*. Phd thesis, University of Waterloo, 2020.

[9] Valerii Fedorov. Optimal experimental design. *WIREs Computational Statistics*, 2(5):581–589, 2010.

[10] Malvin H Kalos and Paula A Whitlock. *Monte Carlo methods*. John Wiley & Sons, 2009.

[11] Nicholas R Lomb. Least-squares frequency analysis of unequally spaced data. *Astrophysics and space science*, 39:447–462, 1976.

[12] Jeffrey D Scargle. Studies in astronomical time series analysis. II-statistical aspects of spectral analysis of unevenly spaced data. *Astrophysical Journal, Part 1, vol. 263, Dec. 15, 1982, p. 835-853.*, 263:835–853, 1982.

[13] Roman V Baluev. Assessing the statistical significance of periodogram peaks. *Monthly Notices of the Royal Astronomical Society*, 385(3):1279–1285, 2008.

[14] T Ruf. The Lomb-Scargle periodogram in biological rhythm research: analysis of incomplete and unequally spaced time-series. *Biological Rhythm Research*, 30(2):178–201, 1999. See the lomb R library.

[15] Stephen P Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[16] Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2023.

[17] Chris Coey, Miles Lubin, and Juan Pablo Vielma. Outer approximation with conic certificates for mixed-integer convex problems. *Mathematical Programming Computation*, 12(2):249–293, 2020.

[18] Michael Gorczyca and Justice Sefas. On weighted trigonometric regression for suboptimal designs in circadian biology studies. *arXiv preprint arXiv:2403.14452*, 2024.

[19] Petre Stoica, Jian Li, and Hao He. Spectral analysis of nonuniformly sampled data: A new approach versus the periodogram. *IEEE Transactions on Signal Processing*, 57(3):843–858, 2008.

[20] Leslie Greengard and June-Yub Lee. Accelerating the nonuniform fast Fourier transform. *SIAM review*, 46(3):443–454, 2004.

[21] Julie Zhou Jane J. Ye and Wenjie Zhou. Computing A-optimal and E-optimal designs for regression models via semidefinite programming. *Communications in Statistics - Simulation and Computation*, 46(3):2011–2024, 2017.

[22] Norman L Johnson, Samuel Kotz, and Narayanaswamy Balakrishnan. *Continuous univariate distributions, volume 2*, volume 289. John Wiley & Sons, 1995.

[23] Milton Abramowitz and Irene A Stegun. *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. National Bureau of Standards, 1964.

[24] Michael Grant, Stephen Boyd, and Yinyu Ye. *Disciplined Convex Programming*, pages 155–210. Springer US, Boston, MA, 2006.