PowerCHORD: constructing optimal experimental designs for biological rhythm discovery

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

January 17, 2025

Abstract

Equally spaced temporal sampling is the standard protocol for the study of biological rhythms. These equispaced designs perform well when calibrated to an oscillator's period yet can have systematic detection biases when applied to rhythms of unknown periodicity. Here, we present a broadly-applicable set of computational methods for seeking optimal measurement schedules for rhythm detection. Our PowerCHORD methods generate experimental designs by maximizing a closed-form expression for the statistical power of the cosinor model using a black-box optimization method (differential evolution), a brute-force search, or mixed-integer conic programming. Application of these three methods showed numerically that they improve upon equispaced designs under many experimental contexts. Our numerical results also revealed an intuitive approach for achieving optimal power for simultaneous investigation of circadian, circalunar, and circannual rhythms. Our findings suggest that timing optimization is an effective yet under-explored tool for improving biological rhythm discovery.

Contents

1 Introduction **2**

2 Background **2**

3 Results **4**
3.1 Calculation and optimization of cosinor power . . . 4
3.2 Irregular designs improve power across frequency windows . . . 5
3.3 PowerCHORD recovers power lost to timing constraints . . . 7
3.4 Globally optimal designs for discrete period uncertainty . . . 10

4 Discussion **12**

5 Methods **13**
5.1 Differential evolution . . . 13
5.2 Brute-force search . . . 13
5.3 Mixed-integer conic programming . . . 14

6 Code Availability **15**

7 Acknowledgements **15**

S1 Proofs of main results **16**
S1.1 Closed-form power expression 16
S1.2 Monotonicity of the noncentral F-distribution 19
S1.3 Worst-case power maximization 22
S1.4 Power analysis at the Nyquist rate 24
S1.5 Unions of equispaced designs 25

S2 Supplementary figures **25**

1 Introduction

Biological rhythms serve essential roles in living systems and arise from diverse mechanisms on scales ranging from individual cells to entire populations [1–4]. While familiar examples such as circadian rhythms, somitogenesis, and the cell-cycle have been studied for decades [5–7], new rhythms and layers of temporal organization continue to be discovered [8–12]. In the context of circadian studies, rhythm discovery has been facilitated in part by improvements to statistical methods [13–16], leading to evidence of new rhythmicity in existing datasets [17]. These improvements are now a major focus of rhythm discovery guidelines [18, 19], however the dependence of statistical performance on the underlying experimental design is less thoroughly explored. Hence, many studies make use of equispaced temporal sampling as the *de facto* design choice.

The preferred status of equispaced designs is justifiable by the strong assumptions of circadian studies. In particular, if the period of an oscillator is known ahead of time, designs with measurements equispaced along the oscillator's cycle achieve statistically optimal performance [20, 21]. Yet, in more general contexts, biological oscillations are known to occur with periods ranging from milliseconds to years and a study may need to consider potential cycles across these vast timescales. Equispaced collection along only one cycle may be unreliable for rhythm detection due to statistical power variability at nominal periods and acrophases (Supp. Fig. 1). Given the inadequacy of equispaced designs for exploring multiple cycles, there is a need for methods that enable optimal detection of cycles with unknown periodicity.

In this article, we present a method for constructing experimental designs that are optimal for detecting unknown rhythms. We obtain an efficient way to generate such designs by deriving a closed-form expression for the statistical power of the cosinor model in Section 3.1 and focusing on designs where the least detectable signal of interest has the highest possible statistical power. Using this formula, we derived a condition for when equispaced designs fail to be optimal for rhythm discovery and investigated its consequences numerically in Sections 3.2-3.4. In Section 3.2 we present optimized designs for rhythm discovery across a range of candidate periodicities while in Section 3.3, we optimize for experiments where the period is known but the measurement time windows are restricted. In Section 3.4 we consider studies with strong prior knowledge of the system, so that only a discrete list of periods is of interest and demonstrates that certain periods can be measured together simultaneously without trade-offs in power. Our analysis and optimization methods generate more efficient experimental designs for rhythm discovery, and the computational methods can be applied to additional experimental contexts using the PowerCHORD (*P*ower analysis and *C*osinor design optimization for *HO*moscedastic *R*hythm *D*etection) open source code repository.

2 Background

Harmonic regression is a popular statistical framework for studying systems with oscillatory features [22]. We perform harmonic regression using the single-frequency cosinor model. For a MESOR

(*M*idline *E*stimating *S*tatistic of *R*hythm) Y_0, amplitude A, acrophase ϕ, and frequency f, the cosinor model takes the form

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

in which $\varepsilon(t) \sim \mathcal{N}(0, \sigma)$ is homoscedastic Gaussian white noise. Assuming the frequency f is fixed, 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 times $\mathbf{t} = \{t_i\}_{i=1}^N \subset [0,1]^N$, where N is the sample size of the experiment, the optimal least squares solution $\hat{\boldsymbol{\beta}}$ for estimating the coefficients $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)$ appearing in Eq. (2) is

$$\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}(\mathbf{t}; f) = (X^T X)^{-1} X^T \mathbf{y}, \qquad X = X(\mathbf{t}; f) = \begin{bmatrix} \mathbf{1} & \cos(2\pi f \mathbf{t}) & \sin(2\pi f \mathbf{t}) \end{bmatrix}. \tag{3}$$

We assume $\mathbf{t}$ contains sufficiently many distinct measurement times for X to be of full column rank and Eq. (3) to be well-defined. For the cosinor model, the rank condition can be rephrased in terms of distinct phases of the cycle at which measurements are collected (Supp. Lemma S1.1). Estimates of the amplitude and acrophase of the signal can be computed from $\hat{\boldsymbol{\beta}}$, to obtain

$$\hat{A} = \sqrt{\hat{\beta}_1^2 + \hat{\beta}_2^2}, \quad \hat{\phi} = \operatorname{atan2}\left(\hat{\beta}_2, \hat{\beta}_1\right).$$

Using the cosinor model, rhythm detection can be formulated as a hypothesis test with null hypothesis $\beta_1 = 0 = \beta_2$ and alternative $\beta_1 \neq 0$ or $\beta_2 \neq 0$. The statistical power quantifies how reliably oscillations are detected by the experimental design and hypothesis test.

Definition 2.1 (Statistical power of the cosinor model)**.** Given a parametric model with parameters $\boldsymbol{\beta} \in \mathbb{R}^p$, data $Y \in \mathbb{R}^N$ measured at times $\mathbf{t} \in \mathbb{R}^N$, and a rejection region $R \subset \mathbb{R}^N$, the power of the hypothesis test is given by

$$\gamma(\mathbf{t}; \boldsymbol{\beta}) = \mathbb{P}_{\boldsymbol{\beta}}(Y \in R).$$

For a cosinor-based hypothesis test, the rejection region is $R = \{\mathbf{x} : \hat{F}(\mathbf{x}) \geq c\}$, in which $\hat{F}(\mathbf{x})$ is the F-statistic

$$\hat{F}(\mathbf{x}) = \frac{TSS - RSS}{RSS} \frac{N-p}{p-1},$$

TSS $= ||\mathbf{x} - \langle \mathbf{x} \rangle||^2$, RSS $= ||\mathbf{x} - X\hat{\boldsymbol{\beta}}||^2$, $\langle \mathbf{x} \rangle = \frac{1}{N}\sum_{i=1}^N x_i$, and $X = X(\mathbf{t}; f)$ is the cosinor design matrix and $\hat{\boldsymbol{\beta}}$ is the least-squares estimate of the parameters from Eq. (3).

Since the exact parameters of the signal are rarely well-known when the design is constructed, we seek designs that achieve high power across a range of parameter values. To this end, we quantify performance using the worst-case power of the design, meaning the lowest power across all signals of interest. Prioritization of worst-case rather than average power ensures that the power is above a known threshold for all relevant signals. A formal definition of worst-case power is given below. For the remainder of the paper we will simply use the terminology "optimal power designs" to refer to optimality with respect to a worst-case scenario.

Definition 2.2 (Optimal worst-case power)**.** Given a domain $\mathcal{B} \subset \mathbb{R}^p$ in parameter space and a design matrix $X \in \mathbb{R}^{N \times p}$, an experimental design $\mathbf{t}^*$ achieves optimal worst-case power with respect to $\mathcal{B}$ if it satisfies

$$\mathbf{t}^* = \underset{\mathbf{t} \in [0,1]^N}{\arg\max} \min_{\boldsymbol{\beta} \in \mathcal{B}} \gamma(\mathbf{t}; \boldsymbol{\beta}).$$

3 Results

3.1 Calculation and optimization of cosinor power

Power analysis is typically approached using Monte Carlo techniques. These techniques are transferable to power optimization, but introduce a significant computational burden. We took a more efficient route by deriving a closed-form expression for statistical power, stated in Theorem 3.1. The power expression is fast to evaluate, generalizes an earlier formula [23] which is only valid when measurements are equispaced (Supp. Fig. 2), and provides insight into the problem's mathematical structure. The reader is directed to Supp. Sec. S1 for proofs of this result and all subsequent theorems.

Theorem 3.1 (Power of the one-frequency cosinor model)**.** *Consider the one-frequency cosinor model*

$$y(t) = \beta_0 + \beta_1 \cos(2\pi f t) + \beta_2 \sin(2\pi f t) + \varepsilon(t), \quad \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

applied to data $\mathbf{y} = \{y_i\}_{i=1}^N$ *collected at distinct times* $\mathbf{t} = (t_i)_{i=1}^N$ *with* $N > 3$*. Suppose the following hypotheses are tested using an F-test*

- *null hypothesis* $H_0 : \beta_1 = 0 = \beta_2$*,*
- *alternative hypothesis* $H_1 : \beta_1 \neq 0$ *or* $\beta_2 \neq 0$*.*

Given parameters $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)$*, the power* γ *of this hypothesis test is given by*

$$\gamma(\mathbf{t}; \boldsymbol{\beta}, f, \sigma) = 1 - F_{\lambda(\mathbf{t};\boldsymbol{\beta},f,\sigma)} \left(F_0^{-1}\left(1-\alpha; 2, N-3\right); 2, N-3 \right),$$
$$\lambda(\mathbf{t}; \boldsymbol{\beta}, f, \sigma) = \frac{1}{\sigma^2} \boldsymbol{\beta}^T H^T (H (X^T X)^{-1} H^T)^{-1} H \boldsymbol{\beta},$$

in which $X = X(\mathbf{t}; f) = \begin{bmatrix} 1 & \cos(2\pi f \mathbf{t}) & \sin(2\pi f \mathbf{t}) \end{bmatrix}$ *is the design matrix,* $\alpha \in (0, 1)$ *is the type I error rate,* $F_\lambda(x; n_1, n_2)$ *is the noncentral F-distribution with* (n_1, n_2) *degrees of freedom and noncentrality parameter* λ*,* $F_0(x; n_1, n_2) = F(x; n_1, n_2)$ *is the F-distribution, and*

$$H = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

is the hypothesis matrix.

Using Theorem 3.1, power optimization can be equivalently formulated as an eigenvalue programming problem. The resemblance of this eigenvalue problem to the Elfving optimality criterion from optimal experimental design [20] leads to a simple condition for the optimality of equispaced designs. When the optimality condition holds, the power of equispaced designs is also guaranteed to be independent of the signal's acrophase (Supp. Fig. 3A). The eigenvalue problem and optimality condition are stated in Theorem 3.2. If an experiment cannot satisfy the optimality condition, optimal designs can be constructed numerically using Corollary 3.2.1.

Theorem 3.2 (Optimality condition for equispaced designs)**.** *Suppose a study aims to detect signals of a specific frequency and unknown acrophase. Maximizing the worst-case power across all acrophases is equivalent to the following eigenvalue optimization problem*

$$\mathbf{t}^* = \operatorname{argmax}_{\mathbf{t} \in [0,1]^N} \lambda_{\min} \left(B(\mathbf{t}; f)^{-1} \right) \tag{5}$$

in which

$$B(\mathbf{t}; f) = H \left(X \left(\mathbf{t}; f \right)^T X(\mathbf{t}; f) \right)^{-1} H^T. \tag{6}$$

If it is possible to collect $N > 3$ equispaced measurements per cycle, then equispaced designs maximize Eq. (5). Moreover, the power of such designs is constant as a function of acrophase with power corresponding to a noncentrality parameter

$$\lambda_{\text{opt}} = \frac{A^2 N}{2\sigma^2}.$$

Corollary 3.2.1. *If the frequency is assumed to lie in a specific window $f \in [f_{\min}, f_{\max}]$ and the acrophase is unknown, then maximizing the worst-case power is equivalent to the following eigenvalue programming problem*

$$\mathbf{t}^* = \underset{\mathbf{t}\in[0,1]^N}{\arg\max} \min_{f\in[f_{\min},f_{\max}]} \lambda_{\min}\left(B(\mathbf{t};f)^{-1}\right). \tag{7}$$

The value achieved by this optimal solution $\mathbf{t}^*$ is bounded by

$$\min_{f\in[f_{\min},f_{\max}]} \lambda_{\min}\left(B(\mathbf{t}^*;f)^{-1}\right) \leq \frac{A^2 N}{2\sigma^2}.$$

PowerCHORD generates designs by maximizing Eq. (7) using a brute-force search, differential evolution, or mixed-integer conic programming. Brute-force searches are computationally tractable if the sample size is low and the measurements are confined to an underlying grid (e.g. only sampling on the hour). For higher sample sizes, differential evolution provides an estimate of the power improvement and mixed-integer conic programming provides more refined information and better structured solutions. The three methods work directly with the eigenvalue problem and their results can be converted back into power using the following identity

$$\lambda^* = \frac{A^2}{\sigma^2}\lambda_{\min}\left(B(\mathbf{t};f)^{-1}\right), \tag{8}$$

in which $\lambda_{\min}\left(B(\mathbf{t};f)^{-1}\right)$ is the minimal eigenvalue from Eq. (5) and λ^* is the lowest value of the noncentrality parameter across all acrophases for signals of amplitude A and frequency f. Justification for Eq. (8) is given in the proof of Supp. Lemma S1.12. In our analysis, we set $A = \sigma$ so that we may refer to the noncentrality parameter and minimum eigenvalue $\lambda_{\min}\left(B(\mathbf{t};f)^{-1}\right)$ interchangeably. We also make use of the fact that the noncentral F distribution is a monotone function of its noncentrality parameter (Supp. Sec. S1.2) and therefore refer to the power of a design and its noncentrality parameter interchangeably.

3.2 Irregular designs improve power across frequency windows

We assume the study seeks to achieve high power across a frequency window $f_{\min} \leq f \leq f_{\max}$. Using PowerCHORD's differential evolution method, we constructed irregular designs for various frequency windows and sample sizes, which improved power for most cases (Fig. 1A). The greatest improvements to power were observed when the frequency window included the Nyquist rate of the equispaced design. Optimization of other windows did not improve upon equispaced designs or only produced a slight change in the power. In comparison to equispaced designs, the power of irregular designs had much weaker power fluctuations around the Nyquist rate (Fig. 1B; Supp. Fig. 3B). Away from the Nyquist rate, the equispaced and irregular designs exhibited power fluctuations that diminish in intensity as the sample size increases. Hence, for a large enough sample size, irregular designs improve worst-case power while maintaining homogeneous power levels throughout the frequency-acrophase plane.

Some of the measurements in irregular designs are spaced farther apart than in equispaced designs of the same sample size. These measurement gaps could make the design undesirably sensitive to the timing of specific measurements. To investigate this sensitivity, we generated "jittered"

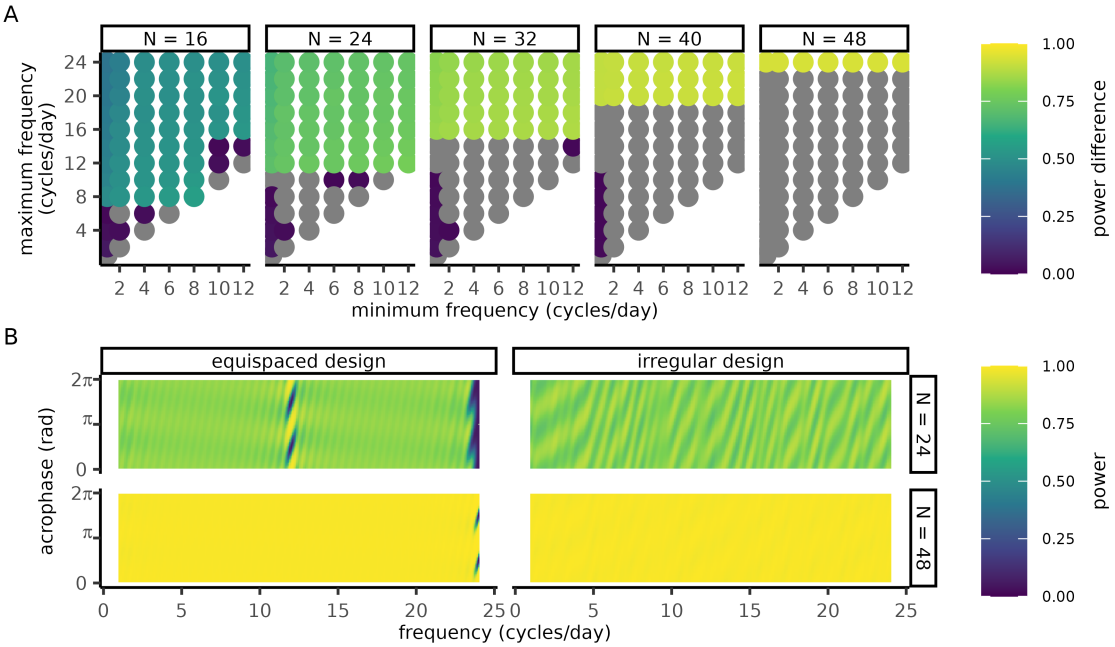


Figure 1: **Irregular sampling improves power in experiments with period uncertainty.** Irregular designs were optimized by differential evolution to detect signals whose frequency is in a window $f_{\min} \leq f \leq f_{\max}$. **(A)** Differences in power between equispaced designs and optimized irregular designs where the x and y axes represent the upper ($f_{\max}$) and lower ($f_{\min}$) limits of each frequency range of interest. The color of each circle represents the difference in worst-case power between an irregular and equispaced design with the same number of measurements (panels). Grey circles correspond to negative power differentials. **(B)** Heatmap of power as a function of frequency (x-axis) and acrophase (y-axis) where color represents the power. **(A-B)** Simulated cosinor parameters: amplitude $A = 1$ and noise strength $\sigma = 1$. Each differential evolution was run with 1hr of compute time with parameters $\mathrm{CR} = 5 \times 10^{-2}$, $N_{\mathrm{pop}} = 10^3$, $\varepsilon = 5 \times 10^{-2}$ (see Sec. 5.1). Details on power calculation at the Nyquist rate are given in Supp. Sec. S1.4.

irregular and equispaced designs by adding Gaussian white noise to each measurement time (Fig. 2A) and calculated the worst-case power of the increasingly perturbed designs. As the noise intensity increased, the power improvements of the irregular design declined as both designs converged in power as measurement times were effectively being sampled from the same distribution (Fig. 2B). For a sample size of 24, the power difference between the irregular and equispaced designs reached half of its full difference when the measurement error was 9 minutes, while for a larger sample size of 48 this difference diminishes more quickly at a measurement error of 5 minutes. Irregular designs always outperformed the eventual random design ($\Delta = 0.18, 0.14, 0.04$ for $N = 24, 32, 48$, respectively). Since timing error is likely to be on the order of minutes, our power improvements are robust to errors that occur on a realistic timescale.

Our irregular designs were optimized for cosinor analysis, yet the data they produce may be used for other purposes such as spectral analysis. To determine the applicability of our irregular designs in such a context, we performed periodogram analysis on simulated datasets. We made use of the Lomb-Scargle periodogram, a common method developed for the study of irregularly sampled data [24,25], and applied a hypothesis test derived from this method [26]. We generated independent simulated datasets for a range of frequencies, each made up of white noise and oscillatory signals with

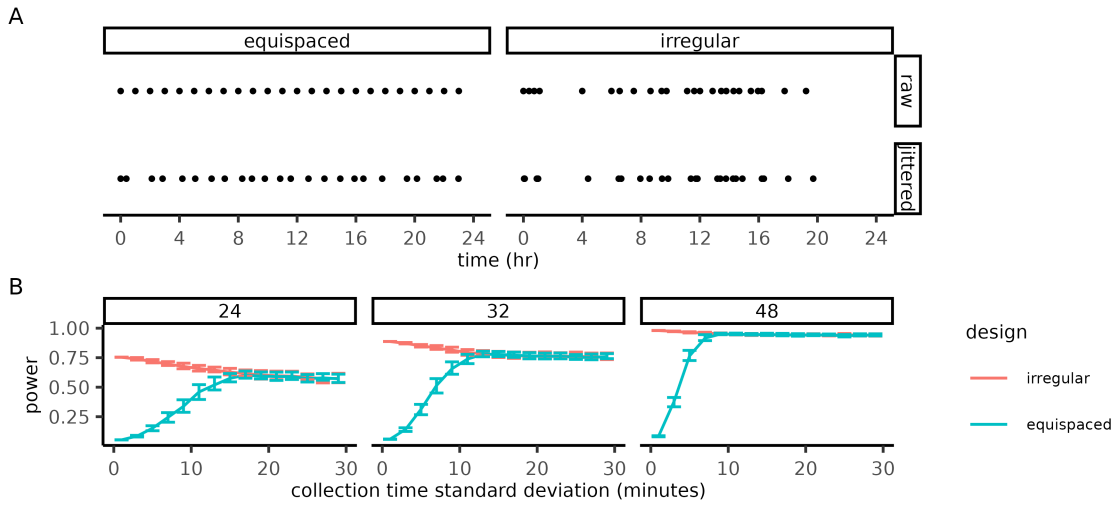


Figure 2: **Irregular designs are robust to perturbations in measurement timing. (A)** Measurement schedules before (top) and after (bottom) jittering. Jittered designs were generated by randomly perturbing the measurement times of equispaced and irregular designs with Gaussian white noise. **(B)** Worst-case power of jittered designs (x-axis) as a function of noise intensity (y-axis) for various sample sizes (panels). The bars indicate the interquartile range for ensembles ($n = 100$) jittered designs. Irregular designs were optimized for the frequency window $f_{\min} = 1$ and $f_{\max} = 24$ using differential evolution with same parameters as Fig. 1. Power was calculated assuming an amplitude $A = 1$.

uniformly random acrophases. We quantified performance using the area under an ROC (receiver operator characteristic) curve [27] because it can be calculated without knowing the false positive rate, and this rate could be affected by switching from cosinor to periodogram analysis. Relative to equispaced designs, irregular designs exhibited weaker fluctuations in their AUC (area under the curve) score at frequencies below the Nyquist rate (Fig. 3A). At the Nyquist rate, the irregular designs dramatically outperformed the equispaced design. Irregular designs performed well only at the frequencies included in their optimization (Fig. 3B), emphasizing the importance of choosing a realistic frequency range before optimizing an experimental design. Although they were constructed for cosinor-based analysis, PowerCHORD designs remain applicable in a broader statistical context.

3.3 PowerCHORD recovers power lost to timing constraints

Equispaced designs can provide optimal power in fixed-period experiments, but they may not always be the most convenient option. For instance, a circadian study with human participants may aim to minimize the number of times a participant is awoken during the night for sample collection. Since strong acute disruptions of sleep have been shown to impact various biological measurements [30,31], accurate characterization of naturally occurring oscillations should avoid disruptions to a 6-12hr window of a subject's regular sleep routine. Incorporating a rest-window reduces the power relative to an equispaced design, so we aimed to determine how much power could be recovered by optimizing the measurement times under the timing constraint.

To simplify the design space, we assumed that measurements could only be collected at times aligned with a half-hour grid. Consequently, it was possible to find optimal designs by running a brute-force search (Sec. 5.2). To gauge the benefit of timing optimization, we compared the optimal designs to "naive" designs in which all measurements are equispaced outside the rest-

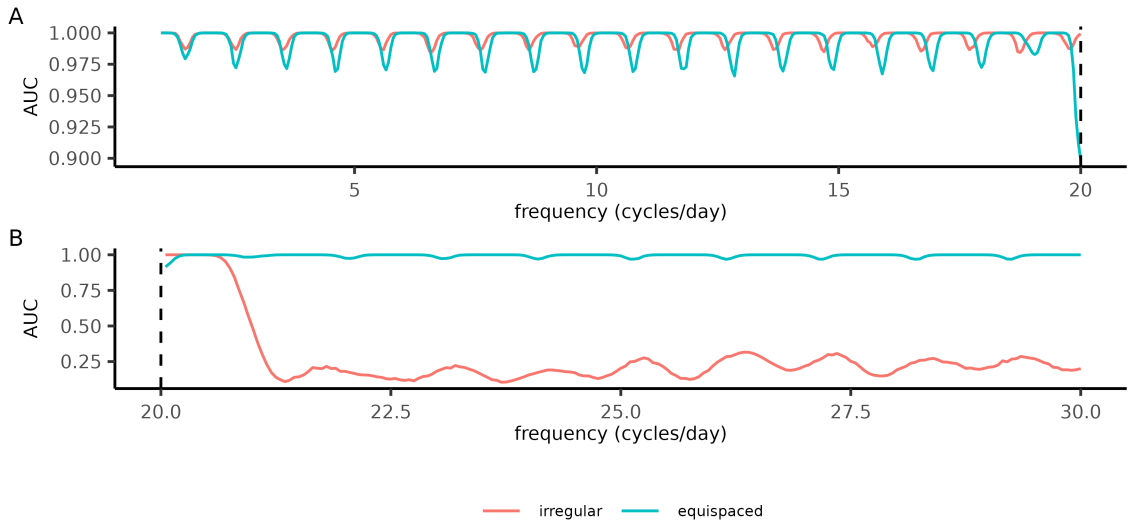


Figure 3: **Irregular designs improve simulated periodogram analysis at frequencies up to the Nyquist rate. (A)-(B)** Oscillations were detected using periodogram analysis with measurements ($N = 40$ samples) from either an equispaced or irregular design optimized for the frequencies $1 \leq f \leq 20$. Each signal in the dataset was assigned an oscillatory (amplitude $A = 2$, noise strength $\sigma = 1$) or non-oscillatory (amplitude $A = 0$, noise strength $\sigma = 1$) state with equal probability. The acrophase of the oscillatory signals was assigned uniformly at random ($\phi \sim \text{Unif}([0, 2\pi))$). **(A)** Performance of the irregular and equispaced designs at detecting oscillations across frequencies (x-axis) included in the optimization, summarized by the AUC score (y-axis) of a receiver operator characteristic curve with p-values generated from a Lomb-Scargle periodogram. The AUC score for each frequency was computed by testing for oscillations in a dataset of oscillatory and white-noise signals ($n = 10^4$ signals per dataset). The dashed line indicates the Nyquist rate of the equispaced design. **(B)** The same analysis as (A) but at frequencies above the Nyquist rate of the equispaced design. Periodogram analysis was performed using the `lomb` library [28] and AUC scores were computed using the `pROC` library [29]. Irregular designs were generated using the same differential evolution parameters as in Fig. 1.

window (Fig. 4A). Relative to the naive designs, the power of the constrained-optimal designs was considerably less sensitive to the acrophase of the signal (Fig. 4B). This difference in sensitivity can be observed in the peak-to-trough power variability of the two designs ($\Delta_{\text{naive}} = 22.93\%$, $\Delta_{\text{optimal}} = 0.35\%$; 8hr rest-window, amplitude $A = 2.5$). To estimate how much the constraints limit the power, we compared the noncentrality parameters of the constrained-optimal and naive designs to an equispaced design. Negligible power was lost in the constrained-optimal design when imposing the shortest window duration. As the window length increased, both the constrained-optimal and naive designs lost power (Fig. 4C). Still, the constrained-optimal designs out-performed the naive designs across the windows under consideration, suggesting that timing optimization can help studies accommodate timing constraints without unnecessarily reducing their power.

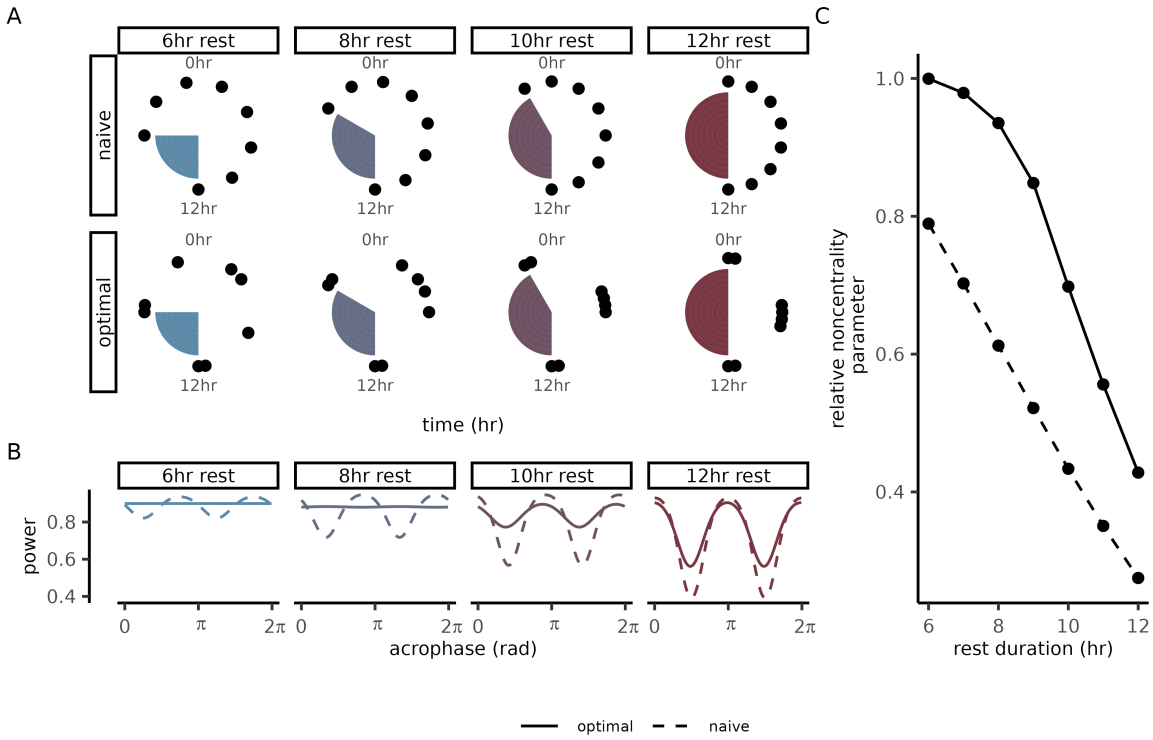


Figure 4: **Circadian studies can balance power with rest-window duration using timing optimization. (A)** Measurement times for optimal and naive designs plotted as phases of a 24hr cycle. Each panel corresponds to a different duration of the rest-window (shaded region), during which no samples can be collected. Optimal designs were generated using a brute-force search with $N_t = 48$ points in the temporal discretization (Sec. 5.2). Naive designs were constructed by distributing all measurements outside the rest window at equal time-intervals. **(B)** Power (y-axis) as a function of acrophase (x-axis) for various durations of the rest-window. The color of each curve represents rest-window duration and the line-type represents if the design is optimal or naive. **(C)** The noncentrality parameter (y-axis) as a function of rest-window duration (x-axis) for naive and optimal designs. Noncentrality parameters are reported relative to the noncentrality parameter of an unconstrained equispaced design with the same measurement budget ($N = 8$ samples).

3.4 Globally optimal designs for discrete period uncertainty

Brute-force searches quickly become intractable at high sample sizes (Supp. Fig. 4). Our third method is a mixed-integer conic program (Theorem 3.3) that continues to be informative about global optimality and remains applicable at higher sample sizes. After simplifying the design problem to only include a small number of frequencies (i.e. discrete period uncertainty), the conic program converged and provided numerical evidence of global optimality. In contrast to the previous two optimization methods, which relied only on the computational efficiency of our power formula, the construction of the conic program relies on the mathematical structure present in the formula (Sec. 5.3).

Theorem 3.3 (Mixed-integer conic program for power optimization)**.** *Worst-case power maximization with discrete time*

$$\begin{aligned} &\textit{maximize} && \min_{f\in[f_{\min},f_{\max}]} \lambda_{\min}\left(B\left(\boldsymbol{\mu};f,\boldsymbol{\tau}\right)^{-1}\right), \\ &\textit{subject to} && \boldsymbol{\mu}\in\{0,1\}^d, \qquad \sum_{i=1}^{d}\mu_i = N, \end{aligned}$$

is equivalent to the following mixed-integer conic programming problem

$$\begin{aligned} &\textit{maximize} && \eta, \\ &\textit{subject to} && X(f;\boldsymbol{\tau})^T \operatorname{diag}(\boldsymbol{\mu}) X(f;\boldsymbol{\tau}) - \eta I \succeq 0, \qquad f\in[f_{\min},f_{\max}], \\ & && \eta\in[0,N/2], \qquad \boldsymbol{\mu}\in\{0,1\}^d, \qquad \sum_{i=1}^{d}\mu_i = N, \end{aligned} \tag{9}$$

in which $X(f,\boldsymbol{\tau})$ is the design matrix of the one-frequency cosinor model evaluated at frequency f on the partition $\boldsymbol{\tau}$, and $B\left(\boldsymbol{\mu};f,\boldsymbol{\tau}\right)$ is given by

$$B(\boldsymbol{\mu};f,\boldsymbol{\tau}) = H\left(X(f,\boldsymbol{\tau})^T \operatorname{diag}(\boldsymbol{\mu}) X(f,\boldsymbol{\tau})\right)^{-1} H^T, \qquad H = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Optimal designs were generated by maximizing Eq. (9) using the CUTSDP method in YALMIP [32] together with Gurobi as a backend solver [33]. The widest frequency separation corresponded to resolving a pair of periods $T = 24$hr and $T = 2$hr with a sample size of $N = 12$ measurements. Interestingly, the optimal solution contained repeating patterns in its measurement schedule (Fig. 5A; Supp. Fig. 5A). While the equispaced design suffered from a low noncentrality parameter at the 2hr rhythm (Fig. 5B), the optimal design reached its theoretical maximum value (Theorem 3.2) at both frequencies of interest. This ability to resolve both frequencies without sacrificing power at either frequency was not observed in 10,000 randomly generated designs (Supp. Fig. 6A) or when additional frequencies were included in the design problem (Supp. Fig. 6B). Closer examination of the bifrequency optimal designs revealed a simple explanation for their ideal power balance; they can be split into equispaced designs when visualized as phases of each period in the design problem (Fig. 5C-D) and therefore inherit the optimality properties of equispaced designs (Supp. Sec. S1.5). Since optimal designs can be constructed by simply ensuring that they satisfy this equiphase property, we used this approach to construct optimal designs for detecting circadian, circalunar, and circannual rhythms in a 12 month experiment (Fig. 5E-F; Supp. Fig. 6C).

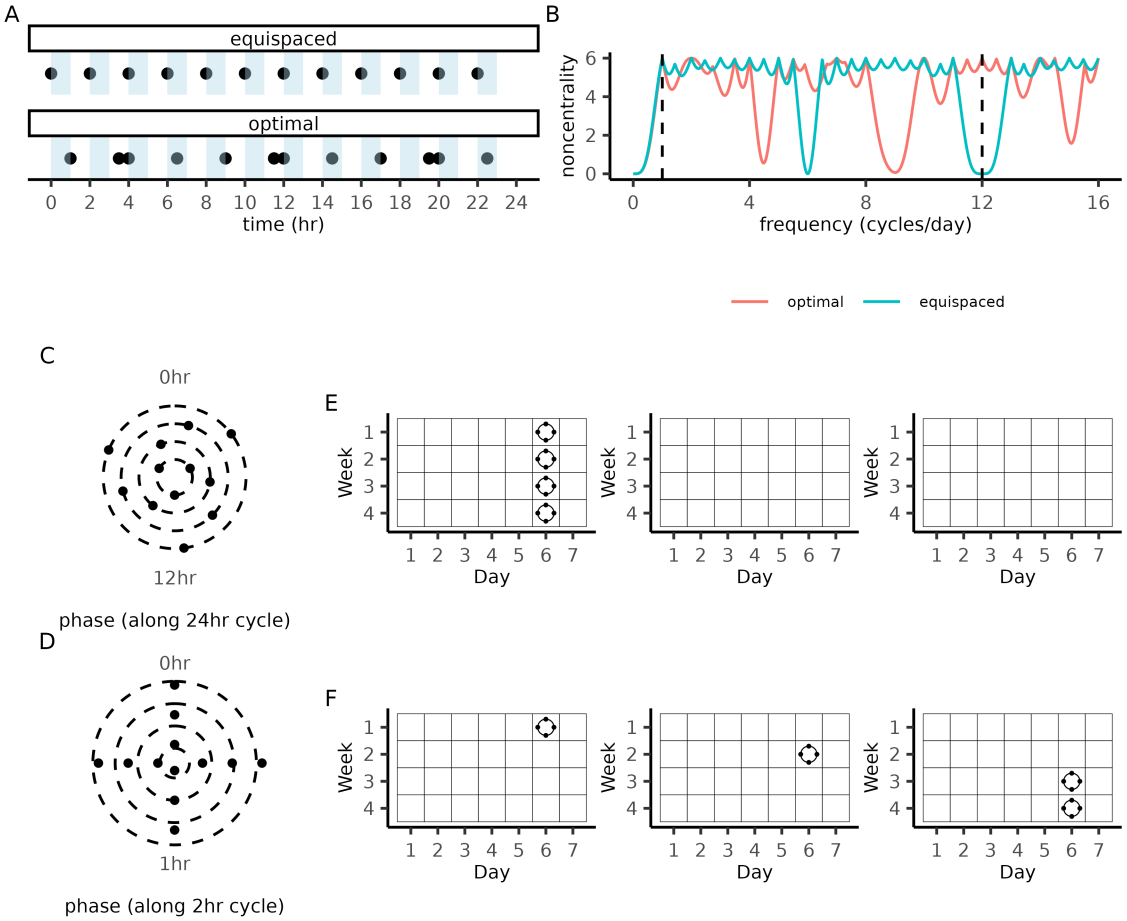


Figure 5: **Globally optimal designs for discrete period uncertainty. (A)** Measurement collection times of an equispaced design (top) and a bifrequency optimal design (bottom) for detecting 2hr and 24hr rhythms. **(B)** Power (y-axis) as a function of frequency (x-axis) for the equispaced and bifrequency optimal design. The two frequencies included in the optimization problem are indicated by the dashed vertical lines. **(C)** Bifrequency optimal design plotted as phases of a 24hr cycle. The values of the radial coordinate were chosen to emphasize that the design can be split into four equiphase designs, each containing three measurement times. **(D)** Bifrequency optimal design plotted as phases of the 2hr cycle. The radial position of each point represents its phase along the 24hr cycle. **(E)** Days marked with circles indicate that $n = 4$ equispaced circadian measurements are to be collected. Repeating this schedule four times over the course of a year, produces a trifrequency optimal design with measurements equispaced along circadian (24 hr), circalunar (28 day), and circannual ($12 \times 28 = 336$ day) cycles. **(F)** An alternative measurement schedule that distributes the measurements across all three months while maintaining the equiphase property.

4 Discussion

We presented a closed-form expression for the statistical power of cosinor based rhythm detection which holds for arbitrary measurement times. This formula enables exploration of the rhythm detection design space in a manner that was infeasible with Monte-Carlo power estimation. This formula will be useful for regular statistical practice to answer questions such as the phase-detection bias introduced by the loss of a sample, or in datasets which may not have been optimally designed (e.g. a postmortem dataset [34]). Using our closed-form power expression, we proved that any given periodicity can be optimally detected by an equispaced design with $N > 3$ measurements per cycle. This optimality condition implies that certain groups of periods can be investigated together without any trade-off in power. In particular, the power will be phase-independent provided that the measurements decompose into equispaced designs when viewed as phases of each period under consideration.

Application of PowerCHORD methods to various experimental scenarios improved the power relative to equispaced designs. When designs were optimized for a continuous range of periods, the improved power was robust to small changes in measurement timing and robust to the application of periodogram rather than cosinor analysis. Under the simplifying assumption that measurements are confined to an underlying grid, two of our three optimization methods (brute-force search and conic programming) produced globally optimal designs. The optimal designs produced by the brute-force searches exhibited a minimal loss in power relative to equispaced while removing the need for costly and logistically difficult overnight visits and avoiding sleep disruption biases. In the conic programming case, the globally optimal designs had noncentrality parameter $\lambda = N/2$ at all frequencies of interest and hence their performance was not limited by the restriction to an underlying grid. Examination of the bifrequency optimal designs revealed an "equiphase" property and led to simple optimal designs for simultaneous investigation of circadian, circalunar, and circannual rhythms.

It may be possible to generate optimal designs in more challenging experimental contexts by improving the optimization methods in PowerCHORD. The brute-force searches could be accelerated to run at larger sample sizes using more efficient algorithms for parameterising the design space [35] and related methods in optimal experimental design [36–38] could be adapted to our problem. While our analysis was focused on how measurement timing influences worst-case power, there are many closely related questions that could be considered in future work. First, the number of biological and technical replicates could be treated as decision variables in the optimization problem [14, 39] to further improve power. Second, figures of merit relevant to the accuracy of parameter estimation or periodogram methods [26, 28, 40, 41] could be prioritized in the study design. Third, timing optimization could be applied to adaptive study design. In this setting, the frequency prior would be updated using Bayesian inference and influence the optimization of measurement times for the next experiment [42]. Beyond the scope of optimization, further improvements to rhythm detection can be achieved by higher quality data, innovative analytical methods and leveraging prior knowledge of the system [43].

In summary, we have demonstrated the benefits of integrating power optimization in the design of rhythm discovery experiments. Our methods are broadly applicable, and can be tailored to specific experimental applications using our open-source PowerCHORD repository. Such studies will be more efficient than traditional approaches for exploring ranges of periods, and may expedite the discovery of novel biological rhythms.

5 Methods

5.1 Differential evolution

Differential evolution algorithms initialize a population of candidate solutions at random. Successive generations are constructed by taking component-wise linear combinations of the parents according to an algorithm-specific update rule. While many variations on this core idea have been studied [44], we chose to implement a relatively simple version of the algorithm.

Given an objective function $J : \mathbb{R}^N \to \mathbb{R}$, our implementation of differential evolution requires four hyper-parameters: the population size N_{pop}, the differential weight $\varepsilon \in [0, 1]$, the crossover probability $CR \in [0, 1]$, and the number of iterations N_{iter}. At initialization, the population is represented by a matrix $x \in \mathbb{R}^{N_{\text{pop}} \times N}$, in which N is the sample size, whose entries are independently and identically distributed as $x_{ij} \sim \text{unif}(0, 1)$. To produce the next generation, the state of the i-th member of the population is updated by generating a random vector $u \sim \text{unif}(0, 1)$, sampling three indices $a, b, c \in \{1, \ldots, N_{\text{pop}}\}$ without replacement to obtain

$$x_{ij} = \begin{cases} y_{ij} & \text{if } u \geq CR \\ x_{ia} + \varepsilon\,(x_{ib} - x_{ic}) & \text{if } u < CR. \end{cases}$$

If $J(y_i) < J(x_i)$, then the new state is accepted. The system repeats this process until a total of N_{iter} generations have been produced. The highest scoring individual of the terminal population is then returned. A summary of the algorithm in pseudo-code is given below.

Algorithm 1 Differential evolution in PowerCHORD

```
Require: N ≥ 3                                          ▷ sample size
Require: N_pop ≥ 4, N_iter > 0, ε > 0, CR ∈ (0,1)       ▷ hyperparameters
Require: J : R^N → R                                    ▷ objective function
  x_ij ← unif(0,1)
  while iter ≤ N_iter do                                ▷ loop over generations
      for i ← 1 to N_pop do
          for j ← 1 to N do                             ▷ generate candidate y_i
              u ← unif(0,1)
              y_ij ← x_ij
              if u < CR then
                  {a,b,c} ⊂ {1,...,N_pop}               ▷ sample three distinct indices
                  y_ij ← x_ia + ε (y_ib − y_ic)
              end if
          end for
          if J(y_i) < J(x_i) then                       ▷ determine if y_i should replace x_i
              x_i ← y_i
          end if
      end for
      iter ← iter + 1
  end while
```

5.2 Brute-force search

Suppose measurements can only be collected at times belonging to a fixed partition

$$\boldsymbol{\tau} = \{k/N_t : k \in \{0, \ldots, N_t - 1\}\}$$

in which N_t is the coarseness of the partition. With measurement times confined to a fixed grid, designs are naturally represented by binary vectors $\boldsymbol{\mu} \in \{0,1\}^{N_t}$ which satisfy $\sum_i \mu_i = N$. Provided that N and N_t are not too large, it is feasible to search the entire space of binary vectors. Since any design $\boldsymbol{\mu} \in \{0,1\}^{N_t}$ will achieve equal performance to all of its cyclic translations, we need only consider equivalence classes of such binary vectors up to rotational symmetry. Following the convention from combinatorics, these equivalence classes are referred to as fixed-density binary necklaces. For a partition of coarseness N_t and sample size N, the number of equivalence classes of such necklaces is given by a sum over the factors of the greatest common divisor of the sample size and the partition coarseness

$$C(N, N_t) = \frac{1}{N_t} \sum_{j \mid \gcd(N, N_t)} \phi(j) \binom{N_t/j}{N/j}, \tag{10}$$

in which $\phi(\cdot)$ is Euler's totient function [45]. We recommend restricting $N_t \leq 72$ and $N \leq 10$ since this function grows rapidly with N and N_t (Supp Fig. 4). An efficient algorithm for generating a representative from each equivalence class is given in [45]. We use their C implementation to generate the designs and search the design database to identify representatives of equivalence classes of optimal solutions.

There is an additional reflection symmetry present in the problem which could further improve the performance of the brute-force search. Notice that the time-reversal of any design will still have equivalent power, hence designs could be considered equivalent up to rotational and reflectional symmetry. In the terminology of combinatorics, these larger equivalence classes are known as "fixed-density binary bracelets". Efficient algorithms for generating representatives of the bracelets have been proposed [35] and would improve the performance of PowerCHORD if they were included.

5.3 Mixed-integer conic programming

We provide a proof of Theorem 3.3 which justifies our use of mixed-integer conic programming. We simplify the optimization problem using the following lemma.

Lemma 5.1 (Schur positivity, [46, Section A.5.5])**.** *Let $Y \in \mathbb{R}^{n \times n}$ be a symmetric matrix partitioned as*

$$Y = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$

with $\det A \neq 0$ and let $S = C - BA^{-1}B^T$ be the Schur complement of Y. Then we have $Y \succ 0$ if and only if $A \succ 0$ and $S \succ 0$.

Proof of Theorem 3.3. Let $X = X(\boldsymbol{\tau}; f)$ and $\tilde{X} = \tilde{X}(\boldsymbol{\tau}; f)$. We seek a binary vector $\boldsymbol{\mu}^*$ that satisfies

$$\boldsymbol{\mu}^* = \underset{\boldsymbol{\mu} \in \{0,1\}^d, \sum_i \mu_i = N}{\arg\max} \; \min_{f \in [f_{\min}, f_{\max}]} \lambda_{\min} \left(B\left(\boldsymbol{\mu}; f, \boldsymbol{\tau}\right)^{-1} \right),$$

in which $B(\boldsymbol{\mu}; f, \boldsymbol{\tau})^{-1}$ can be expressed as

$$B(\boldsymbol{\mu}; f, \boldsymbol{\tau})^{-1} = \tilde{X}^T \operatorname{diag}(\boldsymbol{\mu}) \tilde{X} - \frac{1}{N} \mathbf{b}(\boldsymbol{\mu}; f, \boldsymbol{\tau}) \mathbf{b}(\boldsymbol{\mu}; f, \boldsymbol{\tau})^T.$$

with the latter equality justified by Lemma S1.13. Restricting frequency to a discrete set of frequencies of interest $\{f_1, \ldots, f_M\}$, we arrive at an optimization problem of the form

$$\begin{aligned} \max_{\eta \in \mathbb{R}, \boldsymbol{\mu} \in \{0,1\}^d} \eta \quad \text{s.t.} \quad & \sum_{i=1}^d \mu_i = N, \\ & \tilde{X}^T \operatorname{diag}(\boldsymbol{\mu}) \tilde{X} - \frac{1}{N} \mathbf{b}(\boldsymbol{\mu}; f, \boldsymbol{\tau}) \mathbf{b}(\boldsymbol{\mu}; f, \boldsymbol{\tau})^T - \eta I \succeq 0 \quad \text{for } f \in \{f_1, \ldots, f_M\}. \end{aligned} \tag{11}$$

The quadratic dependence on $\boldsymbol{\mu}$ in Eq. (11) can be reduced to a linear dependence by applying Lemma 5.1 with the matrix Y given by

$$Y = \begin{bmatrix} N & \mathbf{b}\left(\boldsymbol{\mu}; f, \boldsymbol{\tau}\right)^T \\ \mathbf{b}\left(\boldsymbol{\mu}; f, \boldsymbol{\tau}\right) & \tilde{X}(\boldsymbol{\tau}; f)(\boldsymbol{\tau}; f)^T \operatorname{diag}(\boldsymbol{\mu})\tilde{X} \end{bmatrix} = X(\boldsymbol{\tau}; f)^T \operatorname{diag}(\boldsymbol{\mu}) X(\boldsymbol{\tau}; f),$$

which gives

$$\max_{\eta \in \mathbb{R}, \boldsymbol{\mu} \in \{0,1\}^d} \eta \quad \text{s.t.} \quad \begin{aligned} &\sum_{i=1}^{d} \mu_i = N, \\ &X(\boldsymbol{\tau}; f)^T \operatorname{diag}(\boldsymbol{\mu}) X(\boldsymbol{\tau}; f) - \eta I \succeq 0 \quad \text{for } f \in \{f_1, \ldots, f_M\}. \end{aligned}$$

□

6 Code Availability

Implementations of PowerCHORD in R and MATLAB are available through the first author's GitHub. The power analysis features are included in both the R and MATLAB implementations. Only the MATLAB implementation includes GPU acceleration and provides access to the external YALMIP library necessary for mixed-integer conic programming.

7 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 Michael Tackenberg and John Limanto for providing thorough and insightful feedback on this manuscript.

Supplementary material

S1 Proofs of main results

S1.1 Closed-form power expression

The power formula in Theorem 3.1 follows from computing the distribution of the F-statistic under the null and alternative models. We show these computations in detail since self-contained proofs are unavailable in the literature. The first three lemmas are concerned with elementary properties of the design matrix X and the hypothesis matrix H.

Lemma S1.1. *Let $X = X(\mathbf{t}; f) \in \mathbb{R}^{N\times 3}$ be the design matrix of the cosinor model and suppose at least three components of the vector $\mathbf{z} = e^{2\pi i f \mathbf{t}} \in \mathbb{C}^N$ are distinct, then X has full column rank.*

Proof. Without loss of generality, assume the first three components z_1, z_2, z_3 of $\mathbf{z} \in \mathbb{C}^N$ are distinct. Let $\theta_1, \theta_2, \theta_3 \in [0, 2\pi)$ be the corresponding angles $\theta_j = \mathrm{Arg}(z_j)$. To verify that X is of full column rank, it suffices to show that the submatrix

$$M = \begin{bmatrix} 1 & \cos(\theta_1) & \sin(\theta_1) \\ 1 & \cos(\theta_2) & \sin(\theta_2) \\ 1 & \cos(\theta_3) & \sin(\theta_3) \end{bmatrix}$$

has nonzero determinant. Assume without loss of generality that $\theta_3 = 0$ and compute the determinant of M to obtain

$$\det M = \sin(\theta_1) - \sin(\theta_2) - \sin(\theta_1)\cos(\theta_2) + \cos(\theta_1)\sin(\theta_2). \tag{12}$$

Suppose towards contradiction that $\det M = 0$ and rearrange Eq. (12) to show

$$\frac{\sin(\theta_1)}{1-\cos(\theta_1)} = \frac{\sin(\theta_2)}{1-\cos(\theta_2)}. \tag{13}$$

The restriction of the function $f(x) = \frac{\sin(x)}{1-\cos(x)}$ to $x \in [0, 2\pi)$ is strictly monotone and therefore injective. Hence Eq. (13) implies $\theta_1 = \theta_2$, a contradiction. □

Lemma S1.2. *Let $X \in \mathbb{R}^{n\times p}$ with $\mathrm{rank}(X) = p \leq n$ then $X^T X$ is symmetric positive definite.*

Proof. Since $X^T X$ is symmetric, there exists an orthonormal eigenbasis $\mathbf{v}_1, \ldots, \mathbf{v}_p$ and real eigenvalues $\lambda_1, \ldots, \lambda_p$ such that $X^T X \mathbf{v}_i = \lambda_i \mathbf{v}_i$ for $i = 1, \ldots, p$. Since $\mathrm{rank}(X) = p$, we know $X\mathbf{v}_i \neq 0$ and so

$$0 < ||X\mathbf{v}_i||_2^2 = \mathbf{v}_i^T X^T X \mathbf{v}_i = \lambda_i ||\mathbf{v}_i||_2^2,$$

which forces $\lambda_i > 0$ for $i = 1, \ldots, p$. □

Lemma S1.3. *Suppose $A \in \mathbb{R}^{p\times p}$ is symmetric positive definite and $H \in \mathbb{R}^{q\times p}$ with $\mathrm{rank}(H) = q$. If $q \leq p$ then $B = HAH^T$ is symmetric positive definite.*

Proof. Symmetry is apparent, so it suffices to show that $\mathbf{x}^T HAH^T \mathbf{x} > 0$ for any $\mathbf{x} \in \mathbb{R}^q$. Since $\mathrm{rank}(H) = q$, we know $H^T \mathbf{x} \neq 0$ for all nonzero $\mathbf{x} \in \mathbb{R}^q$. Positivity of HAH^T now follows from the positivity of A, since $(\mathbf{x}^T H)A(H^T \mathbf{x}) > 0$ for all $\mathbf{x} \in \mathbb{R}^q$. □

The next lemma and its corollary will be useful for computing the null and alternative distributions of the F statistic. In particular, it is useful to recognise that the definition of the F-statistic most frequently applied in experiments can also be written in a form more suitable for mathematical proofs.

Lemma S1.4. *Given a design matrix $X \in \mathbb{R}^{n\times p}$ with* $\operatorname{rank}(X) = p$ *and a hypothesis matrix $H \in \mathbb{R}^{q\times p}$ with $q \le p \le n$, define the following linear subspaces*

- $V_u = \{X\boldsymbol{\beta} : \boldsymbol{\beta} \in \mathbb{R}^p\} = \operatorname{Col}(X)$,
- $V_r = \{X\boldsymbol{\beta} : \boldsymbol{\beta} \in \mathbb{R}^p, H\boldsymbol{\beta} = 0\}$,
- $V_\ell = \{X\boldsymbol{\beta} : \boldsymbol{\beta} \in \mathbb{R}^p, H\boldsymbol{\beta} \neq 0\}$.

Let P_u, P_r, and P_ℓ be the orthogonal projectors onto the respective subspaces. The projector P_ℓ satisfies

$$P_\ell = P_u - P_r = Q(Q^TQ)^{-1}Q^T \tag{14}$$

in which $Q = X(X^TX)^{-1}H^T$.

Proof. Since V_u can be expressed as a direct sum $V_u = V_\ell \oplus V_r$, we clearly have $P_\ell = P_u - P_r$. To verify the second equality, it suffices to show that the columns of Q span V_ℓ. To confirm this, we verify the equivalent condition $\operatorname{Null}(Q^T) = V_\ell^\perp$. We may write $V_\ell^\perp$ as $V_\ell^\perp = V_u^\perp \oplus V_r$. Any $\mathbf{y} \in V_r$ can be expressed as $\mathbf{y} = X\boldsymbol{\beta}$ so $Q^T\mathbf{y} = H\boldsymbol{\beta} = 0$. Similarly $\mathbf{y} \in V_u^\perp$ must satisfy $P_u\mathbf{y} = X(X^TX)^{-1}X^T\mathbf{y} = 0$ and hence $Q^T\mathbf{y} = 0$ and we see $\operatorname{Null}(Q^T) \supset V_u^\perp \oplus V_r$. For the reverse containment, notice that any $\mathbf{z} \in \mathbb{R}^n \setminus \left(V_r \oplus V_u^\perp\right) = V_\ell$, must satisfy $\mathbf{z} = X\boldsymbol{\beta}$ and $H\mathbf{z} \neq 0$. It follows that $\mathbf{z} \notin \operatorname{Null}(Q^T)$ because

$$Q^T\mathbf{z} = H(X^TX)^{-1}X^TX\boldsymbol{\beta} = H\boldsymbol{\beta} \neq 0.$$

Since $\operatorname{Null}(Q^T) = V_u^\perp \oplus V_r$, the rank-nullity theorem implies that $\operatorname{Col}(Q) = V_\ell$ and the second equality of Eq. (14) follows. □

Corollary S1.4.1. *Let $X = X(\mathbf{t}; f)$ be the design matrix of the cosinor model. Given data $\mathbf{y} \in \mathbb{R}^n$, the total sum of squares*

$$\mathrm{TSS} = ||\mathbf{y} - \langle\mathbf{y}\rangle||^2 = ||(I - P_r)\mathbf{y}||^2,$$

and the residual sum of squares

$$\mathrm{RSS} = ||\mathbf{y} - X\hat{\boldsymbol{\beta}}||^2 = ||(I - P_u)\mathbf{y}||^2,$$

in which $\hat{\boldsymbol{\beta}} = (X^TX)^{-1}X^T\mathbf{y}$ is the least squares estimator and X is the one-frequency cosinor design matrix, the following three definitions of the F-statistic are equivalent

$$\begin{aligned}
\hat{F}_1 &= \frac{TSS - RSS}{RSS}\frac{N-p}{p-1}, \\
\hat{F}_2 &= \frac{\mathbf{y}^T\left(P_u - P_r\right)\mathbf{y}}{\mathbf{y}^T\left(I - P_u\right)\mathbf{y}}\frac{N-p}{p-1}, \\
\hat{F}_3 &= \frac{\left(H\hat{\boldsymbol{\beta}}\right)^T\left(H\left(X^TX\right)^{-1}H^T\right)^{-1}\left(H\hat{\boldsymbol{\beta}}\right)}{\mathbf{y}^T\left(I - P_u\right)\mathbf{y}}\frac{N-p}{p-1},
\end{aligned} \tag{15}$$

in which $H = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is the rhythm-detection hypothesis matrix.

The linear algebraic tools assembled so far allow us to deduce the null and alternative distributions of the F-statistic by simply applying Cochran's theorem.

Theorem S1.5 (Cochran (1934), [47])**.** *If $\mathbf{x} \in \mathbb{R}^n \sim \mathcal{N}(\mathbf{0}, I)$ and a family of $n \times n$ matrices $\{A_i\}_{i=1}^m$ satisfy*

1. $A_i = A_i^T$ *for* $i = 1, \ldots, m$,

2. $\text{rank}(A_i) = r_i$ *for* $i = 1, \ldots, m$,

3. $\sum_{i=1}^m A_i = I$,

then each $\mathbf{x}^T A_i \mathbf{x} \sim \chi^2_{r_i}$ *independently if and only if* $\sum_{i=1}^m r_i = n$.

Proof. See Chapter 2 of [48]. □

Lemma S1.6 (Null distribution of F-statistic)**.** *Suppose data* $\mathbf{y} \in \mathbb{R}^n$ *is of the form* $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon} \in \mathbb{R}^N$ *with* $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I)$ *and the null hypothesis* $H\boldsymbol{\beta} = 0$ *holds true, then the F-statistic from Cor. S1.4.1 satisfies*

$$\hat{F}(\mathbf{y}) \sim F(\cdot; 2, N-3).$$

Proof. We use the definition of the F-statistic in Eq. (15). Since the projection matrices $(I - P_u)$ and $(P_u - P_r)$ vanish on constant vectors, we may assume without loss of generality that $\boldsymbol{\mu} = \mathbf{0}$. Since the projection matrices are idempotent and symmetric, their ranks are given by the dimension of the subspaces onto which they project

$$\text{rank}(P_u - P_r) = p - 1 = 2, \quad \text{rank}(I - P_u) = N - p = N - 3.$$

It follows from Theorem S1.5 that the following two quadratic forms have independent chi-squared distributions

$$\begin{aligned} \mathbf{y}^T (P_u - P_r) \mathbf{y} &\sim \chi^2(2) \\ \mathbf{y}^T (I - P_u) \mathbf{y} &\sim \chi^2(N-3), \end{aligned}$$

and hence their ratio has an $F(\cdot; 2, N-3)$ distribution after scaling by the degrees of freedom. □

Lemma S1.7. *Given projection matrices* P_A *and* P_B *that satisfy* $P_A P_B = 0$ *and* $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \sigma^2 I)$, *the random variables* $||P_A \mathbf{x}||_2^2$ *and* $||P_B \mathbf{x}||_2^2$ *are independent.*

Proof. Since P_A and P_B are commuting projectors they can be simultaneously diagonalized. Working in their common eigenbasis, their characteristic function $\chi_{AB}(||P_A\mathbf{x}||_2^2, ||P_B\mathbf{x}||_2^2)$ clearly satisfies

$$\chi(||P_A\mathbf{x}||_2^2, ||P_B\mathbf{x}||_2^2) = \chi_A(||P_A\mathbf{x}||_2^2)\chi_B(||P_B\mathbf{x}||_2^2)$$

which implies independence. □

Lemma S1.8 (Alternative distribution of F-statistic)**.** *Suppose data* $\mathbf{y} \in \mathbb{R}^N$ *is of the form* $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon} \in \mathbb{R}^N$ *with* $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I)$ *and the alternative hypothesis* $H\boldsymbol{\beta} \neq 0$ *holds true, then the F-statistic from Cor. S1.4.1 has a noncentral F-distribution*

$$\hat{F}(\mathbf{y}) \sim F_\lambda(\cdot | 2, N-3)$$

with noncentrality parameter $\lambda = \frac{1}{\sigma^2} (H\boldsymbol{\beta})^T B^{-1} H\boldsymbol{\beta}$ *in which* $B = H(X^T X)^{-1} H^T$.

Proof. The vector appearing in the numerator of the F-statistic in Eq. (15) satisfies

$$H\hat{\boldsymbol{\beta}} \sim \mathcal{N}(H\boldsymbol{\beta}, \sigma^2 H(X^T X)^{-1} H^T).$$

Since X and H satisfy the non-degeneracy conditions of Lemmas S1.1-S1.3, B is symmetric positive definite and therefore has a Cholesky factorization. Let $B^{-1/2}$ be a Cholesky factor of B^{-1}, and

since $\frac{1}{\sigma}B^{-1/2}H\hat{\boldsymbol{\beta}} \sim \mathcal{N}(B^{-1/2}H\boldsymbol{\beta}, I)$, the numerator of Eq. (15) can be rescaled to have a noncentral chi-squared distribution

$$\frac{1}{\sigma^2}(H\hat{\boldsymbol{\beta}})^T B^{-1}(H\hat{\boldsymbol{\beta}}) \sim \chi^2_\lambda(2)$$

with noncentrality parameter $\lambda = \frac{1}{\sigma^2}(H\boldsymbol{\beta})^T B^{-1} H\boldsymbol{\beta}$. Since $(I - P_u)X\boldsymbol{\beta} = 0$ for any $\boldsymbol{\beta} \in \mathbb{R}^p$, the denominator of Eq. (15) satisfies

$$\frac{1}{\sigma^2}\left(\mathbf{y} - X\boldsymbol{\beta}\right)^T \left(I - P_u\right)\left(\mathbf{y} - X\boldsymbol{\beta}\right) = \boldsymbol{\varepsilon}^T(I - P_u)\boldsymbol{\varepsilon} \sim \chi^2(N-3).$$

Finally Lemma S1.7 ensures that the numerator and denominator of Eq. (15) are independent and the claim follows. □

Proof of Theorem 3.1. The power γ of a hypothesis test is given by

$$\gamma(\boldsymbol{\beta}) = \mathbb{P}_{\boldsymbol{\beta}}(\hat{F}(\mathbf{y}) \in R)$$

in which $R \subset \mathbb{R}^N$ is the rejection region, $\mathbf{y} \in \mathbb{R}^N$ is the data and $\hat{F}$ is the F-statistic. Lemma S1.6 implies $R = \{\mathbf{y} \in \mathbb{R}^N : \hat{F}(\mathbf{y}) \geq F^{-1}(1-\alpha|2, N-3)\}$ and we may conclude from Lemma S1.8 that $\mathbb{P}_{\boldsymbol{\beta}}(x \geq c) = 1 - F_\lambda(c; 2, N-3)$ for any $c > 0$. Hence we have

$$\gamma(\boldsymbol{\beta}) = 1 - F_\lambda(F_0^{-1}(1-\alpha; 2, N-3); 2, N-3)$$

with $\lambda = \frac{1}{\sigma^2}(H\boldsymbol{\beta})^T B^{-1} H\boldsymbol{\beta}$. □

S1.2 Monotonicity of the noncentral F-distribution

Throughout our work, we treat power maximziation and maximization of the noncentrality parameter as equivalent optimization problems. This equivalence is justified because the F-distribution is a monotone function of its noncentrality parameter.

Proposition S1.1 (Monotonicity of the noncentral F distribution)**.** *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.*

We prove Prop. S1.1 by expressing the F-distribution as an infinite series, justifying term-wise differentiation, and verifying that the derivative of the distribution with respect to its noncentrality parameter has a definite sign.

Lemma S1.9 (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*

$$F(x; n_1, n_2, \lambda) = \sum_{k=0}^{\infty}\left(\frac{\left(\frac{\lambda}{2}\right)^k}{k!}e^{-\lambda/2}\right) I\left(\frac{n_1 x}{n_2 + n_1 x}; \frac{n_1}{2} + k, \frac{n_2}{2}\right), \tag{16}$$

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

$$\begin{aligned} B(z; a, b) &:= \int_0^z t^{a-1}(1-t)^{b-1}\ \mathrm{d}t, \\ I(z; a, b) &:= B(z; a, b)/B(1; a, b). \end{aligned} \tag{17}$$

Proof. A derivation of this formula is given in [49]. □

Lemma S1.10. *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. (17) can be expressed as

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

in which $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}\mathrm{d}t$ is the gamma function [50]. Use integration by parts and the identity $\Gamma(z+1) = z\Gamma(z)$ to obtain

$$I(z; a+1, b) = I(z; a, b) - \frac{z^a(1-z)^b}{aB(1; a, b)}. \tag{18}$$

Since $z \in [0,1]$ the final term in Eq. (18) is of definite sign, hence

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

To obtain the asymptotic result, rearrange Eq. (18) and integrate by parts to obtain

$$\begin{aligned}
\frac{I(z; a+1, b)}{I(z; a, b)} &= 1 - \frac{z^a(1-z)^b}{a\int_0^z x^{a-1}(1-x)^{b-1}\mathrm{d}x} \\
&= 1 - \frac{z^a(1-z)^b}{z^a(1-z)^{b-1} + (b-1)\int_0^z x^a(1-x)^{b-2}\mathrm{d}x} \\
&= 1 - \frac{1-z}{1 + \frac{(b-1)}{(1-z)^{b-1}}\int_0^z \left(\frac{x}{z}\right)^a (1-x)^{b-2}\mathrm{d}x}.
\end{aligned} \tag{19}$$

Since $b > 0$, $x < z$ and $z \in [0,1]$, we have

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

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

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

So we conclude from Eq. (19) that $\lim_{a\to\infty} I(z; a+1, b)/I(z; a, b) = z$. □

Lemma S1.11. *Let $T_k(z, n_1, n_2, \lambda)$ be the terms of the series expansion in Eq. (16) 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 for $k = 0$, we have $|\partial_\lambda T_0| \le \frac{1}{2}T_0$ and for $k > 0$

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

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. S1.10. □

Proof of Proposition S1.1. Use Prop. S1.9 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 Lemma S1.10 and Lemma S1.11 to justify term-wise differentiation of the series. Differentiate term-wise and apply the same reasoning as in the proof of Lemma S1.11 to find

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

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

S1.3 Worst-case power maximization

We prove Theorem 3.2 in two steps. We first show that the worst-case value of the noncentrality parameter is related to an eigenvalue of the matrix $B(\mathbf{t}; f)$ defined in Eq. (6). We then compute this eigenvalue for equispaced designs and show that it is optimal using a classical result from optimal experimental design.

Lemma S1.12. *Let H be the hypothesis matrix defined in Eq. (4) and let $B = B(\mathbf{t}; f)$ be the matrix given in Eq. (6). The lowest value of the noncentrality parameter across all signals of a given amplitude and frequency*

$$\lambda^* := \min_{\beta_1^2+\beta_2^2=A^2} \frac{1}{\sigma^2}\boldsymbol{\beta}^T H^T B^{-1} H\boldsymbol{\beta}$$

is given by

$$\lambda^* = \frac{A^2}{\sigma^2}\lambda_{\min}(B^{-1}),$$

in which $\lambda_{\min}(\cdot)$ is the smallest eigenvalue of a symmetric matrix.

Proof. Start by simplifying the definition of λ^* as

$$\begin{aligned}\lambda^* &= \min_{\beta_1^2+\beta_2^2=A^2} \frac{1}{\sigma^2}\boldsymbol{\beta}^T H^T B^{-1} H\boldsymbol{\beta} = \min_{\beta_1^2+\beta_2^2=A^2} \frac{1}{\sigma^2}\begin{bmatrix}\beta_1 & \beta_2\end{bmatrix} B^{-1}\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix} \\ &= \frac{A^2}{\sigma^2}\min_{\beta_1^2+\beta_2^2=1}\begin{bmatrix}\beta_1 & \beta_2\end{bmatrix} B^{-1}\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}. \end{aligned} \tag{20}$$

We can set $\frac{A}{\sigma} = 1$ without loss of generality since it appears homogeneously in Eq. (20). Lemma S1.2 ensures that $(X^TX)^{-1}$ is symmetric positive definite when X is the design matrix for the cosinor model. Since the rhythm detection hypothesis matrix has full row rank, Lemma S1.3 ensures that B is symmetric positive definite. It follows that B^{-1} has a Cholesky factorization $B^{-1} = LL^T$ for some $L \in \mathbb{R}^{2\times 2}$. We may rewrite Eq. (20) in terms of L and obtain

$$\lambda^* = \min_{\beta_1^2+\beta_2^2=1}\left\|L\begin{bmatrix}\beta_1 & \beta_2\end{bmatrix}^T\right\|^2 = \sigma_{\min}(L)^2.$$

The last equality is justified by the max-min principle for singular values

$$\sigma_i(L) = \max_{\dim(U)=i}\ \min_{x\in U, ||x||=1} ||Lx||,$$

with $\sigma_2 \leq \sigma_1$, so $i = 2$ gives the smallest singular value. The result follows by recalling that B^{-1} has real eigenvalues and thus $\sigma_{\min}(L)^2 = \lambda_{\min}(B^{-1})$. □

The next two lemmas allow are used in computing the minimum eigenvalue of $B(\mathbf{t}; f)^{-1}$ for equispaced designs.

Lemma S1.13. *Let $B = B(\mathbf{t}; f)$ be the matrix $B = H(X^TX)^{-1}H^T$, in which $X = X(\mathbf{t}; f)$ is the design matrix for the one-frequency cosinor model and H is the hypothesis matrix. The inverse of B can be expressed as*

$$B^{-1} = \tilde{X}^T\tilde{X} - \frac{1}{N}\mathbf{b}\mathbf{b}^T, \tag{21}$$

in which $\tilde{X} = \tilde{X}(\mathbf{t}; f) = \begin{bmatrix}\cos(2\pi f\mathbf{t}) & \sin(2\pi f\mathbf{t})\end{bmatrix}$ is the design matrix for the mean-free cosinor model $Y(t) = \beta_1\cos(2\pi ft) + \beta_2\sin(2\pi ft) + \varepsilon(t)$ and $\mathbf{b}$ is the vector of time-averages

$$\mathbf{b} = \mathbf{b}(\mathbf{t}; f) = \begin{bmatrix}\sum_{i=1}^N \cos(2\pi ft_i) & \sum_{i=1}^N \sin(2\pi ft_i)\end{bmatrix}^T. \tag{22}$$

Proof. Compute the inverse of $X^T X$ using the Schur complement formula and act with H to find that the only remaining term agrees with the right hand side of Eq. (21). □

Lemma S1.14. *If $\mathbf{t} \in \mathbb{R}^N$ is equispaced over a cycle of frequency f and $N \geq 3$ then*

1. $\sum_{i=1}^N \cos(2\pi f t_j) = 0 = \sum_{j=1}^N \sin(2\pi f t_j)$,
2. $X(\mathbf{t}; f)^T X(\mathbf{t}; f) = \mathrm{diag}(N, N/2, N/2)$.

Proof. Since the measurements are equispaced over a cycle of frequency f, we may assume without loss of generality that they are located at the roots of unity $z_k = e^{2\pi i k/N}$ for $k = 0, \ldots, N-1$. The time averages can be evaluated using a finite geometric series to find

$$\sum_{j=1}^N \cos(2\pi f t_j) + i \sin(2\pi f t_j) = \sum_{k=0}^{N-1} z_k = \sum_{k=0}^{N-1} \left(e^{2\pi i/N}\right)^k = \frac{1 - e^{2\pi i}}{1 - e^{2\pi i/N}} = 0.$$

The same argument applied to the entries of the $X^T X$ reveals that $X^T X = \mathrm{diag}(N, N/2, N/2)$. □

It follows from Lemma S1.14 that any equispaced $\mathbf{t} \in \mathbb{R}^N$ with $N > 3$ will produce a design matrix that satisfies

$$\lambda_{\min}(B(\mathbf{t}; f)^{-1}) = \lambda_{\min}(\tilde{X}^T \tilde{X}) = \frac{N}{2}. \tag{23}$$

We verify that the value $N/2$ on the right hand side of Eq. (23) is globally optimal by using an equivalence theorem from optimal experimental design. These classical theorems are reviewed in [51]. For our purposes, it is convenient to use a more modern statement of the result, which we adapt from [21].

Theorem S1.15 (Elfving optimality condition, [21, Theorem 7.22])**.** *Let Ξ be the design space*

$$\Xi = \{(w_i, t_i)_{i=1}^d : d \in \mathbb{N}, \quad 0 \leq w_i \leq 1, \quad \sum_{i=1}^N w_i = 1, \quad t_i \in [0,1] \quad \text{for } i = 1, \ldots, d\}$$

and let $\mathcal{M}(\Xi)$ be the corresponding space of moment matrices

$$\mathcal{M}(\Xi) = \sum_{i=1}^N w_i \mathbf{f}(t_i) \mathbf{f}(t_i)^T$$

in which $\mathbf{f}(t) = \begin{bmatrix}\cos(2\pi f t) & \sin(2\pi f t)\end{bmatrix}^T$ is the mean-free cosinor model. A moment matrix $M^* \in \mathcal{M}(\Xi)$ is Elfving optimal in the sense that

$$\lambda_{\min}(M^*) = \max_{M \in \mathcal{M}(\Xi)} \lambda_{\min}(M)$$

if and only if there exists a positive semidefinite matrix $E \in \mathbb{R}^{2\times 2}$ such that $\mathrm{tr}\,(E) = 1$ and

$$\mathbf{f}(t)^T E \mathbf{f}(t) \leq \lambda_{\min}(M) \quad \textit{for all } t \in [0,1].$$

Proof of Theorem 3.2. Power and noncentrality parameter maximization for a given acrophase, frequency, and amplitude are equivalent by Prop. S1.1. Maximizing the noncentrality parameter across all acrophases is equivalent to maximizing the minimum eigenvalue of $B(\mathbf{t}; f)^{-1}$ by Lemma S1.12. It remains to show that equispaced designs with $N > 3$ measurements provide optimal solutions to the eigenvalue programming problem and that their power is acrophase-independent.

First use Lemma S1.13 to rewrite the eigenvalue problem as

$$\lambda_{\min}(B^{-1}) = \inf_{||\mathbf{x}||=1} \mathbf{x}^T B^{-1}\mathbf{x} = \inf_{||\mathbf{x}||=1} \mathbf{x}^T \left(\tilde{X}^T\tilde{X} - \frac{1}{N}\mathbf{b}\mathbf{b}^T\right)\mathbf{x}.$$

We clearly have $\mathbf{x}^T \left(\tilde{X}^T\tilde{X} - \frac{1}{N}\mathbf{b}\mathbf{b}^T\right)\mathbf{x} \leq \mathbf{x}^T\tilde{X}^T\tilde{X}\mathbf{x}$ for any $\mathbf{x} \in \mathbb{R}^2$, and consequently

$$\lambda_{\min}(B^{-1}) = \mathbf{x}^T \left(\tilde{X}^T\tilde{X} - \frac{1}{N}\mathbf{b}\mathbf{b}^T\right)\mathbf{x} \leq \lambda_{\min}(\tilde{X}^T\tilde{X}).$$

If $\mathbf{t}$ is equispaced, then $\mathbf{b} = \mathbf{0}$ by Lemma S1.14 and so $B^{-1} = \tilde{X}^T\tilde{X}$ and we find

$$\lambda_{\min}(B^{-1}) = \lambda_{\min}(\tilde{X}^T\tilde{X}) = \frac{N}{2}.$$

To apply Theorem S1.15 and complete the verification of global optimality, notice that every equispaced design $\mathbf{t}$ corresponds to a standardized design $\xi(w_i, t_i)_{i=1}^N \in \mathcal{M}(\Xi)$ with $w_i = 1/N$ with design matrix $M(\xi) = \mathrm{diag}(1/2, 1/2)$. Notice that we have

$$\mathbf{f}(t)^T M(\xi)\mathbf{f}(t) = f_1^2(t) + f_2^2(t) = 1 = \mathrm{tr}\,(M(\xi))$$

which verifies global optimality by Theorem S1.15. Finally, explicit computation confirms that the noncentrality of an equispaced design satisfies

$$\lambda(\mathbf{t}; \boldsymbol{\beta}, f, \sigma) = \frac{A^2 N}{2\sigma^2} \tag{24}$$

where $A^2 = \beta_1^2 + \beta_2^2$ is the amplitude. Notice that Eq. (24) does not depend on the acrophase of the signal, as required.

□

S1.4 Power analysis at the Nyquist rate

For a sample size N and equispaced design $\mathbf{t}_N$, power analysis at the Nyquist rate $f = \frac{N}{2}$ requires some additional consideration. Even if the sample size satisfies $N > 3$, the measurements will not cover sufficiently many distinct phases for the design matrix to be full rank. Hence, the methods developed so far in this section are not immediately applicable for computing the worst-case power of an equispaced design at the Nyquist rate. The following proposition allows us to circumvent this issue.

Proposition S1.2 (Noncentrality parameter at the Nyquist rate)**.** *For $N > 3$, the noncentrality parameter of an equispaced design $\mathbf{t}_N$ satisfies*

$$\lim_{f \to \frac{N}{2}} \lambda_{\min}(B(\mathbf{t}_N; f)^{-1}) = 0$$

and hence the worst-case power is given by $\gamma = \alpha$, in which α is the type-I error rate.

Proof. Evaluate the entries of $B\left(\mathbf{t}_N; \frac{N}{2}\right)^{-1}$ using Eq. (21) and a finite geometric series to obtain

$$B\left(\mathbf{t}_N; \frac{N}{2}\right)^{-1} = \begin{cases} \begin{bmatrix} N & 0 \\ 0 & 0 \end{bmatrix} & \text{if } N \text{ is even,} \\ \begin{bmatrix} N - \frac{1}{N} & 0 \\ 0 & 0 \end{bmatrix} & \text{if } N \text{ is odd.} \end{cases}$$

Since eigenvalues are continuous with respect to entrywise matrix convergence (see for instance [52, Proposition 2.4.9.2]), it follows that

$$\lim_{f \to \frac{N}{2}} \lambda_{\min} \left(B \left(\mathbf{t}_N; \frac{N}{2} \right)^{-1} \right) = 0.$$

When the noncentrality parameter $\lambda = 0$, the power expression simplifies to

$$\gamma \left(\mathbf{t}_N; \boldsymbol{\beta}, \frac{N}{2}, \sigma \right) = 1 - F_0 \left(F_0^{-1} \left(1 - \alpha; 2, N - 3 \right); 2, N - 3 \right) = \alpha.$$

□

S1.5 Unions of equispaced designs

Consider a family of equispaced designs $\mathcal{T} = \{\mathbf{t}^{(1)}, \ldots, \mathbf{t}^{(K)}\}$ with each $\mathbf{t}^{(i)} \in [0,1]^{N_i}$ and suppose the design matrix of each individual equispaced $\mathbf{t}^{(i)}$ satisfies the non-degeneracy condition from Lemma S1.1. Our goal here is to show that the union of these designs

$$\mathbf{t} = \begin{bmatrix} \mathbf{t}^{(1)} \\ \vdots \\ \mathbf{t}^{(K)} \end{bmatrix}$$

retains the optimality properties of the individual designs. The key observation comes from Lemma S1.13 and noticing that the time-average vector $\mathbf{b}(\mathbf{t}; f)$ defined in Eq. (22) vanishes since

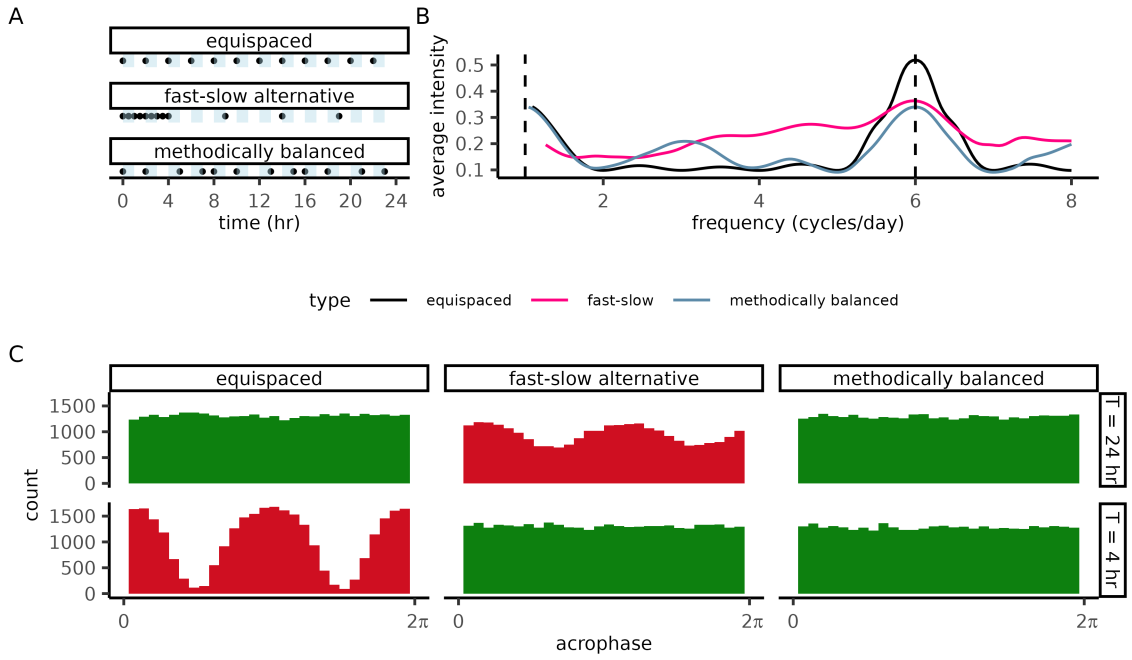
$$\mathbf{b}(\mathbf{t}; f) = \begin{bmatrix} \sum_{i=1}^{K} \sum_{j=1}^{N_i} \cos(2\pi f t_j^{(i)}) \\ \sum_{i=1}^{K} \sum_{j=1}^{N_i} \sin(2\pi f t_j^{(i)}) \end{bmatrix} = \sum_{i=1}^{K} \mathbf{b}(\mathbf{t}^{(i)}; f) = \mathbf{0}.$$

Hence, using Eq. (21), matrix $B(\mathbf{t}, f) = H \left(X(\mathbf{t}; f)^T X(\mathbf{t}; f) \right)^{-1} H^T$ simplifies to

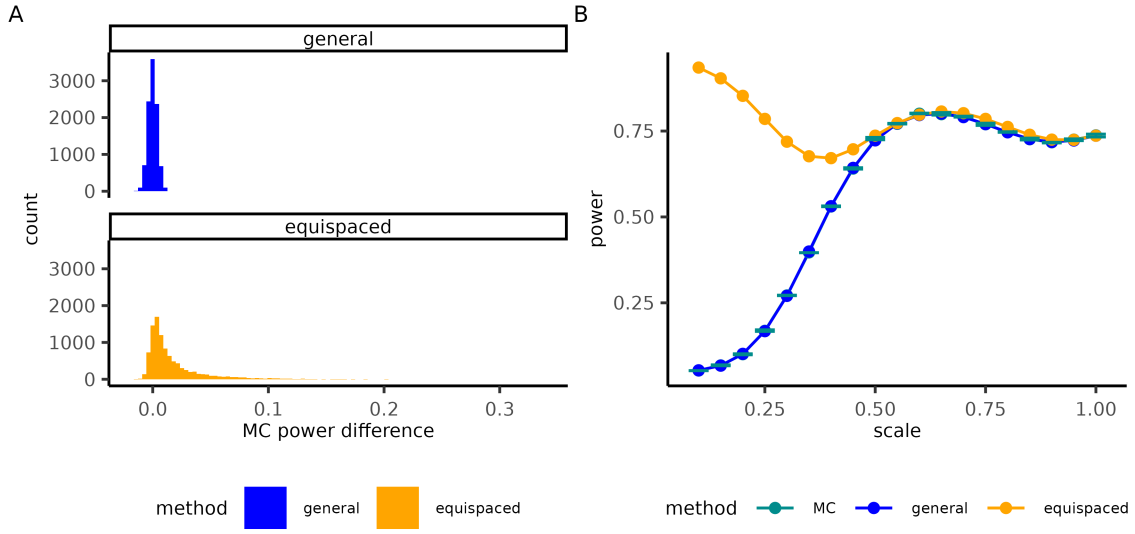
$$B(\mathbf{t}; f)^{-1} = \tilde{X}(\mathbf{t}; f)^T \tilde{X}(\mathbf{t}; f) = \mathrm{diag}(N/2, N/2) \tag{25}$$

in which $N = \sum_{i=1}^{K} N_i$. It follows from Eq. (25) that the phase-independent power and optimality properties of the individual designs are still applicable in the full design $\mathbf{t}$.

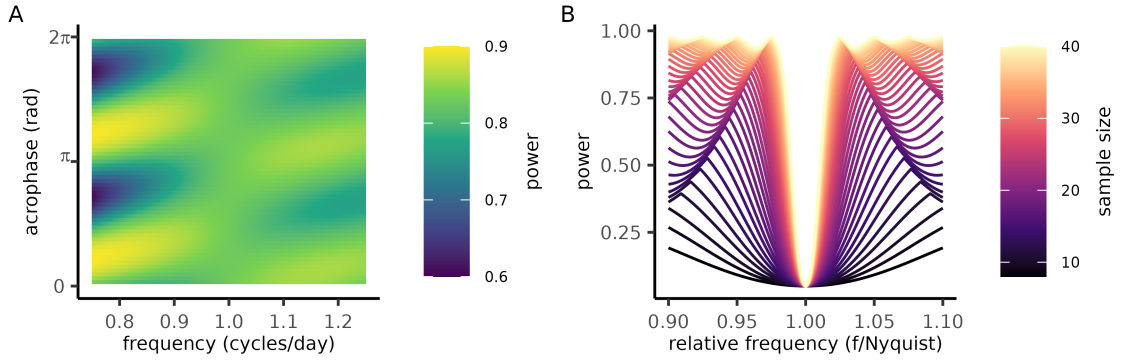
S2 Supplementary figures



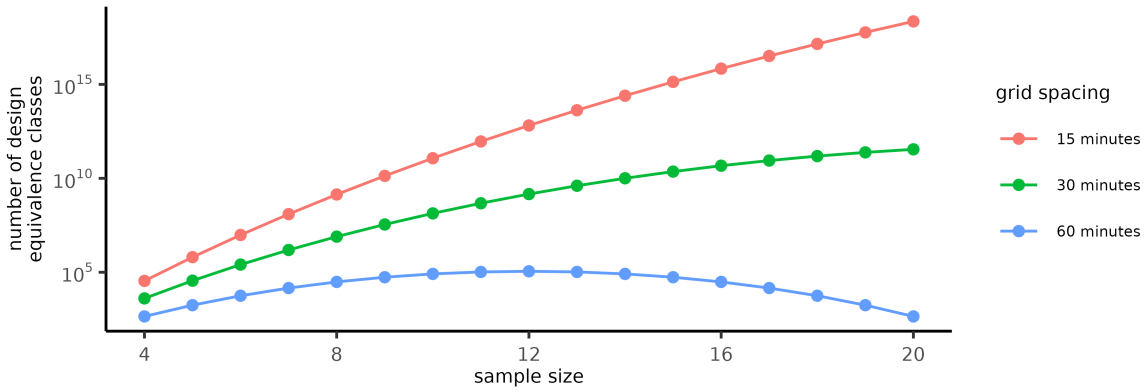
Supplementary Figure 1: **An irregular design detects 24hr and 4hr rhythms at all acrophases.** Each oscillator ($n = 10^5$) in the simulated dataset was assigned a 24hr or 4hr period and a uniformly random acrophase $\phi \in [0, 2\pi)$. Measurements are simulated with Gaussian white noise at each measurement time. **(A)** Measurement schedules for (top) a traditional equispaced design, a (middle) fast-slow irregular design, and (bottom) a methodically constructed irregular design. The shaded bars represent 2hr increments and dots indicate sample collection ($N = 12$ samples for each design). **(B)** The average intensity of a Lomb-Scargle periodogram for each design. The true periods in the system are marked by the dashed vertical lines. **(C)** True acrophases of statistically significant oscillators ($p < 0.05$) detected by cosinor analysis at each of the true periods. Distributions with phase-dependent detection are shown in red to emphasize that the distribution's variability is due to a statistical artifact. Simulation parameters: amplitude $A = \sqrt{2}$, noise strength $\sigma = 1$, acrophase $\phi \sim \text{Unif}(0, 2\pi)$, period $T = 4\text{hr}, 24\text{hr}$. The methodically-constructed design was generated using mixed-integer conic programming in PowerCHORD.



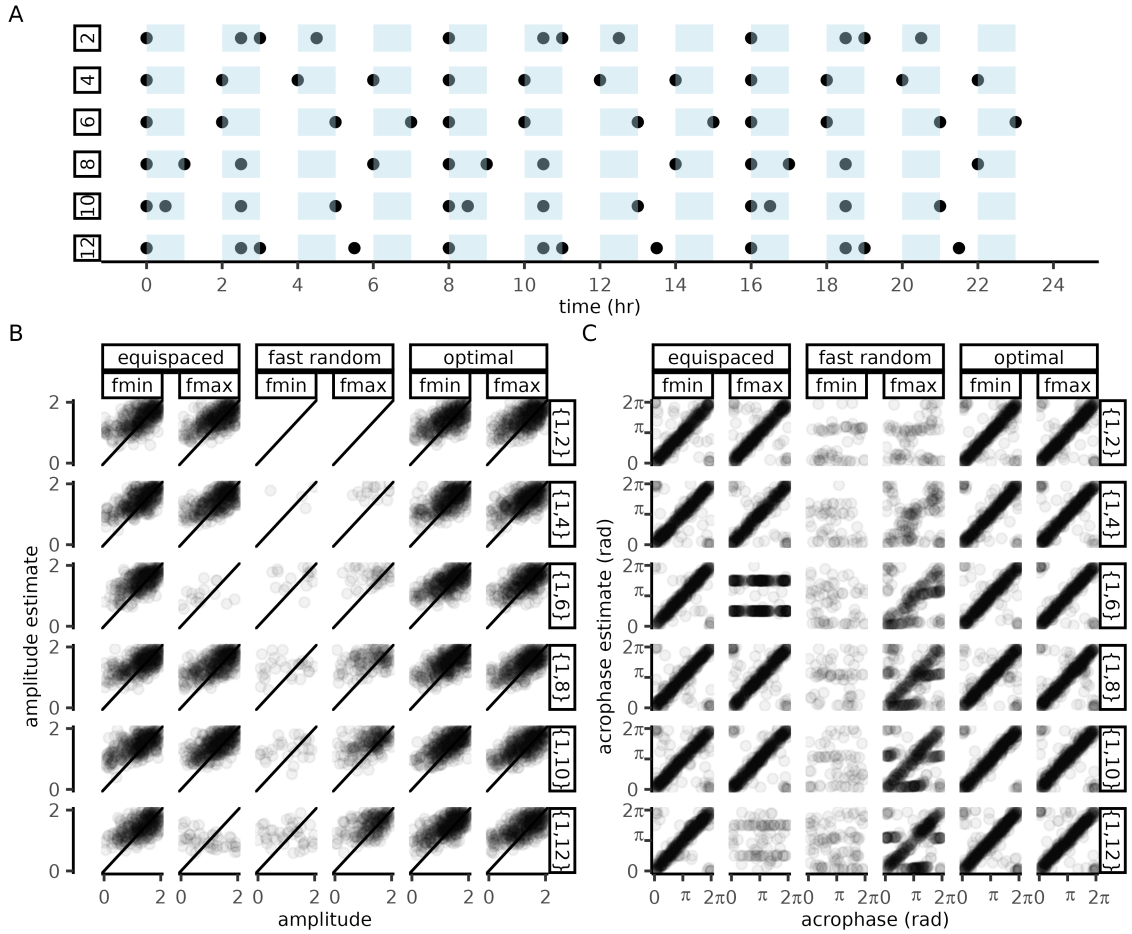
Supplementary Figure 2: **The general power formula is necessary for accurate power analysis. (A)** Comparison of our power formula and the equispaced formula to Monte Carlo estimates for randomly generated designs in which $t \sim \text{unif}([0,1])$ for each measurement time t. On average, the equispaced formula tended to over-estimate the power of such designs. Parameters: sample size $N = 8$, amplitude $A = 2$, frequency $f = 1$,acrophase $\phi = \pi$, noise strength $\sigma = 1$. **(B)** We computed the power of designs $\mathbf{t}_{N,\kappa} = \kappa \mathbf{t}_N$, where $\mathbf{t}_N$ is an N measurement equispaced design and $\kappa > 0$ is a scale factor. As κ shrinks, the irregularity in the design becomes more pronounced and the equispaced formula diverges from the Monte Carlo power estimates. Parameters: $N = 24$, $A = 1$, $f = 1$,$\phi = 0$



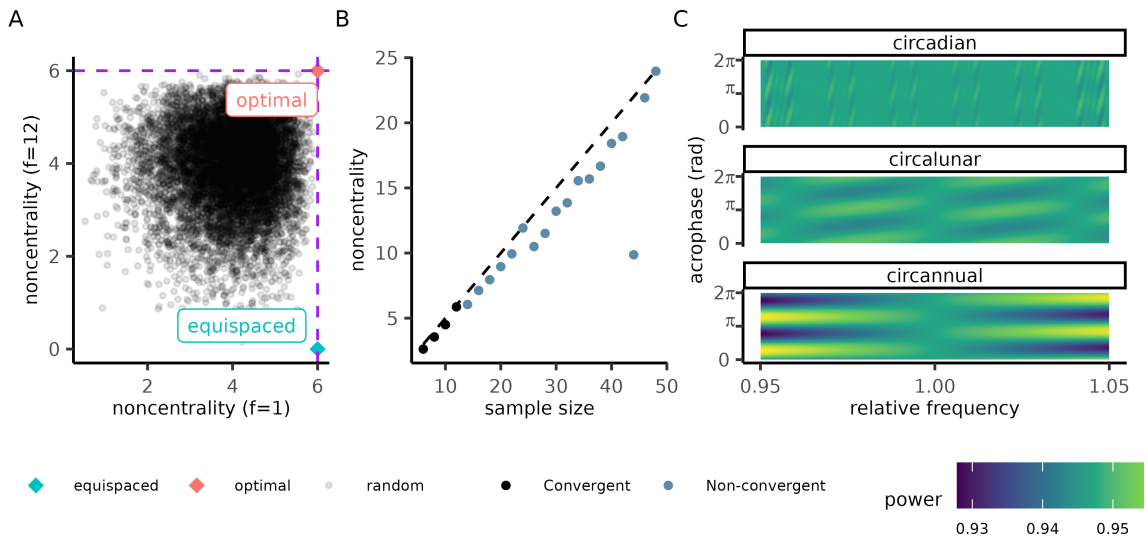
Supplementary Figure 3: **Phase-dependence of equispaced power near critical frequencies. (A)** The power (color) of an equispaced design ($N = 24$ samples) evaluated at each frequency (x-axis) and acrophase (y-axis). The power is independent of phase when the frequency reaches $f = 1$ because the design is equiphase at this frequency. **(B)** The worst-case power of equispaced designs (sample size $8 \leq N \leq 40$) as a function of frequency, with frequency scaled relative to Nyquist rate of each design ($f_{\text{rel}} = f/f_{\text{Nyq}}$). Parameters: amplitude $A = 1$, noise strength $\sigma = 1$.



Supplementary Figure 4: **The number of experimental design equivalence classes grows rapidly with sample size.** For a given sample size (x-axis) and grid spacing (color), the number of design equivalence classes (y-axis) can be calculated using Eq. (10). Designs are in the same equivalence class if they can be transformed into one another by a cyclic shift (i.e. $t \to t + k/N_t$ mod 1, for some $1 \leq k \leq N_t$ assuming measurements are in the interval $[0, 1]$ and confined to a grid of spacing $1/N_t$).



Supplementary Figure 5: **Bifrequency optimal designs exhibit low bias at both intended frequencies.** We computed bifrequency optimal designs for frequency priors $\nu = (1, f)$ with $f \in \{2, 4, 6, 8, 10, 12\}$ and a sample size $N = 12$. **(A)** Repetitive patterns appear in the measurement times of the bifrequency optimal designs. **(B)** Each design was tested for bias by comparing the true values of the amplitude and acrophase to their cosinor estimates after filtering for statistically significance. With the exception of signals at integer multiples of the Nyquist rate, equispaced and optimal designs performed similarly. At the Nyquist multiples, the optimal designs exhibited much less bias than equispaced designs. A randomly generated design ($t \sim \text{unif}(0, 1/12)$) with measurements confined to a short timescale was included as a reference. The random design performed poorly at low frequencies and improved as the higher frequency approaches the scale on which its points are distributed.



Supplementary Figure 6: **Bifrequency and trifrequency optimal designs at critical frequencies. (A)** Non-centrality parameters for the two periodicities of interest (f=1 and f=12) for the optimal design (red dot) an equispaced design (blue dot) and an ensemble ($n = 10^4$) of randomly generated designs. The theoretical maximum value of the non-centrality parameter ($\lambda = N/2$; Theorem 3.2) is indicated by dashed lines. **(B)** Designs were generated to maximize power at the first $N/2$ harmonics ($f \in (1, \ldots, N/2)$) for each sample size N (x-axis). The performance of each design is summarized by the lowest value of its non-centrality parameter (y-axis) across all harmonics included in the optimization. Color indicates convergence of the conic program within 1hr of computation time. The optimal noncentrality parameter in a single frequency design for each sample size ($\lambda = N/2$) is shown for reference (dashed line). For sample sizes $1 \leq N < 12$ measurements were confined to a 36 point grid, for $12 \leq N < 24$ a 48 point grid, and for $24 < N$ a 96 point grid. **(C)** The trifrequency optimal design with all measurements confined to the first month achieves phase independent power at 24 hr (circadian), 28 day (circalunar), and $12 \times 28 = 336$ day (circannual) periods. Parameters: Amplitude $A = 1/\sqrt{2}$.

References

[1] Albert Goldbeter. *Biochemical Oscillations and Cellular Rhythms: The Molecular Bases of Periodic and Chaotic Behaviour*. Cambridge University Press, 1996.

[2] Yangxiaolu Cao, Allison Lopatkin, and Lingchong You. Elements of biological oscillations in time and space. *Nature Structural & Molecular Biology*, 23(12):1030–1034, 2016.

[3] Albert Goldbeter. Dissipative structures in biological systems: bistability, oscillations, spatial patterns and waves. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2124):20170376, 2018.

[4] Béla Novák and John J Tyson. Design principles of biochemical oscillators. *Nature Reviews Molecular Cell Biology*, 9(12):981–991, 2008.

[5] Alina Patke, Michael W Young, and Sofia Axelrod. Molecular mechanisms and physiological importance of circadian rhythms. *Nature Reviews Molecular Cell Biology*, 21(2):67–84, 2020.

[6] Yuchuan Miao and Olivier Pourquié. Cellular and molecular control of vertebrate somitogenesis. *Nature Reviews Molecular Cell Biology*, pages 1–17, 2024.

[7] Catherine Oikonomou and Frederick R Cross. Frequency control of cell cycle oscillators. *Current Opinion in Genetics & Development*, 20(6):605–612, 2010.

[8] Lamya Ghenim, Cédric Allier, Patricia Obeid, Lionel Hervé, Jean-Yves Fortin, Maxim Balakirev, and Xavier Gidrol. A new ultradian rhythm in mammalian cell dry mass observed by holography. *Scientific Reports*, 11(1):1290, 2021.

[9] Isaiah J Ting, Andreas Psomas, Debra J Skene, and Daan R Van der Veen. Reduced glucose concentration enhances ultradian rhythms in Pdcd5 promoter activity in vitro. *Frontiers in Physiology*, 14:1244497, 2023.

[10] Kamil Pradel, Gniewosz Drwiega, Lukasz Chrobok, and Tomasz Blasiak. Racing and pacing in the reward system: a multi-clock circadian control over dopaminergic signalling. *Frontiers in Physiology*, 13:932378, 2022.

[11] Gabriel Oh, Sasha Ebrahimi, Matthew Carlucci, Aiping Zhang, Akhil Nair, Daniel E. Groot, Viviane Labrie, Peixin Jia, Edward S. Oh, Richie H. Jeremian, others Miki Susic, Tenjin C. Shrestha, Martin R. Ralph, Juozas Gordevičius, Karolis Koncevičius, and Art Petronis. Cytosine modifications exhibit circadian oscillations that are involved in epigenetic diversity and aging. *Nature Communications*, 9, 2018.

[12] Michael E. Hughes, Luciano DiTacchio, Kevin R. Hayes, Christopher Vollmers, S. Pulivarthy, Julie E. Baggs, Satchidananda Panda, and John B. Hogenesch. Harmonics of circadian gene transcription in mammals. *PLOS Genetics*, 5(4):1–12, 04 2009.

[13] Michael E Hughes, John B Hogenesch, and Karl Kornacker. JTK_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *Journal of Biological Rhythms*, 25(5):372–380, 2010.

[14] Alan L Hutchison, Mark Maienschein-Cline, Andrew H Chiang, SM Ali Tabei, Herman Gudjonson, Neil Bahroos, Ravi Allada, and Aaron R Dinner. Improved statistical methods enable greater sensitivity in rhythm detection for genome-wide data. *PLOS Computational Biology*, 11(3):e1004094, 2015.

[15] Paul F Thaben and Pål O Westermark. Detecting rhythms in time series with RAIN. *Journal of Biological Rhythms*, 29(6):391–400, 2014.

[16] Forest Agostinelli, Nicholas Ceglia, Babak Shahbaba, Paolo Sassone-Corsi, and Pierre Baldi. What time is it? Deep learning approaches for circadian rhythms. *Bioinformatics*, 32(12):i8–i17, 2016.

[17] Bokai Zhu, Qiang Zhang, Yinghong Pan, Emily M Mace, Brian York, Athanasios C Antoulas, Clifford C Dacso, and Bert W O'Malley. A cell-autonomous mammalian 12 hr clock coordinates metabolic and stress rhythms. *Cell Metabolism*, 25(6):1305–1319, 2017.

[18] Michael E Hughes, Katherine C Abruzzi, Ravi Allada, Ron Anafi, Alaaddin Bulak Arpat, Gad Asher, Pierre Baldi, Charissa De Bekker, Deborah Bell-Pedersen, Justin Blau, et al. Guidelines for genome-scale analysis of biological rhythms. *Journal of Biological Rhythms*, 32(5):380–393, 2017.

[19] Gang Wu, Marc D Ruben, Yinyeng Lee, Jiajia Li, Michael E Hughes, and John B Hogenesch. Genome-wide studies of time of day in the brain: Design and analysis. *Brain Science Advances*, 6(2):92–105, 2020.

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

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

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

[23] 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.

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

[25] 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.

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

[27] Ricardo Cao and Ignacio López-de Ullibarri. *ROC Curves for the Statistical Analysis of Microarray Data*, pages 245–253. Springer New York, New York, NY, 2019.

[28] 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.

[29] Xavier Robin, Natacha Turck, Alexandre Hainard, Natalia Tiberti, Frédérique Lisacek, Jean-Charles Sanchez, and Markus Müller. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, 12:77, 2011.

[30] Marie E Gaine, Snehajyoti Chatterjee, and Ted Abel. Sleep deprivation and the epigenome. *Frontiers in Neural Circuits*, 12:14, 2018.

[31] Simon N Archer, Emma E Laing, Carla S Möller-Levet, Daan R van der Veen, Giselda Bucca, Alpar S Lazar, Nayantara Santhi, Ana Slak, Renata Kabiljo, Malcolm von Schantz, et al. Mistimed sleep disrupts circadian regulation of the human transcriptome. *Proceedings of the National Academy of Sciences*, 111(6):E682–E691, 2014.

[32] J. Löfberg. YALMIP : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

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

[34] Marianne L Seney, Kelly Cahill, John F Enwright III, Ryan W Logan, Zhiguang Huo, Wei Zong, George Tseng, and Colleen A McClung. Diurnal rhythms in gene expression in the prefrontal cortex in schizophrenia. *Nature Communications*, 10(1):3355, 2019.

[35] Z. Alamgir S. Karim and S. M. Husnine. Generating fixed density bracelets of arbitrary base. *International Journal of Computer Mathematics*, 91(3):434–446, 2014.

[36] 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.

[37] Neythen J Treloar, Nathan Braniff, Brian Ingalls, and Chris P Barnes. Deep reinforcement learning for optimal experimental design in biology. *PLOS Computational Biology*, 18(11):e1010695, 2022.

[38] Samuel Bandara, Johannes P Schlöder, Roland Eils, Hans Georg Bock, and Tobias Meyer. Optimal experimental design for parameter estimation of a cell signaling model. *PLOS Computational Biology*, 5(11):e1000558, 2009.

[39] David Laloum and Marc Robinson-Rechavi. Methods detecting rhythmic gene expression are biologically relevant only for strong signal. *PLOS Computational Biology*, 16(3):e1007666, 2020.

[40] 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.

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

[42] Juliane Liepe, Sarah Filippi, Michał Komorowski, and Michael PH Stumpf. Maximizing the information content of experiments in systems biology. *PLOS Computational Biology*, 9(1):e1002888, 2013.

[43] Athanasios C Antoulas, Bokai Zhu, Qiang Zhang, Brian York, Bert W O'Malley, and Clifford C Dacso. A novel mathematical method for disclosing oscillations in gene transcription: A comparative study. *PLOS One*, 13(9):e0198503, 2018.

[44] Swagatam Das and Ponnuthurai Nagaratnam Suganthan. Differential evolution: A survey of the state-of-the-art. *IEEE Transactions on Evolutionary Computation*, 15(1):4–31, 2011.

[45] Frank Ruskey and Joe Sawada. An efficient algorithm for generating necklaces with fixed density. *SIAM Journal on Computing*, 29(2):671–684, 1999.

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

[47] William G Cochran. The distribution of quadratic forms in a normal system, with applications to the analysis of covariance. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 30, pages 178–191. Cambridge University Press, 1934.

[48] Shayle R Searle. *Linear Models*, volume 65. John Wiley & Sons, 1997.

[49] Norman L Johnson, Samuel Kotz, and Narayanaswamy Balakrishnan. *Continuous Univariate Distributions, Volume 2*, volume 289. John Wiley & Sons, 1995.

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

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

[52] Roger A Horn and Charles R Johnson. *Matrix Analysis*. Cambridge university press, second edition, 2012.