Heliyon 8 (2022) e12500

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Fast tuning of observer-based circadian phase estimator using biometric data

Chukwuemeka O. Ike [a,b,*], John T. Wen [a,b], Meeko M.K. Oishi [c], Lee K. Brown [c], A. Agung Julius [a,b]

[a] Lighting Enabled Systems and Applications (LESA) Engineering Research Center, Troy, NY, United States
[b] Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY, United States
[c] Department of Internal Medicine and School of Engineering, University of New Mexico, Albuquerque, NM, United States

ARTICLE INFO

Keywords:
Actigraphy
Circadian rhythm
Melatonin
Sleep-wake cycle
State observer
Wearable data

ABSTRACT

Circadian rhythms play a vital role in maintaining an individual's well-being, and they have been shown to be the product of the master oscillator in the suprachiasmatic nuclei (SCN) located in the brain. The SCN however, is inaccessible for assessment, so existing standards for circadian phase estimation often focus on the use of indirect measurements as proxies for the circadian state. These methods often suffer from severe delays due to invasive methods of sample collection, making online estimation impossible. In this paper, we propose a linear state observer as an elegant solution for continuous phase estimation. This observer-based filter is used in isolating the frequency components of input biometric signals, which are then taken to be the circadian state. We start the design process by fixing the observer's oscillatory frequency at 24 hours, and then we tune its gains using an evolutionary optimization algorithm to extract the target components from individuals' data. The resulting filter was able to provide phase estimates with an average absolute error within 1.5 hours on all test subjects, given their minute-to-minute actigraphy data collected in ambulatory conditions.

1. Introduction

A large number of biological processes within living organisms have been found to exhibit a cycle with an approximate 24-hour period. These processes are collectively referred to as circadian rhythms and include variations in heart rate, core body temperature, and metabolism, as well as a multitude of other processes. Circadian rhythm generators play a vital role in the overall well-being of the organism and are conserved throughout the evolutionary tree as far back as cyanobacteria, although the actual genetic machinery varies. In the animal kingdom (including humans), all cells contain identical genetic machinery capable of generating a circadian rhythm; indeed, peripheral circadian rhythm generators have been demonstrated in a wide variety of tissues including cardiomyocytes and vascular smooth muscle, where they exert their effects through the control of protein effectors. Under steady state conditions, these peripheral clocks are synchronized by a master circadian clock located in the suprachiasmatic nucleus (SCN) of the hypothalamus. The SCN is composed of approximately 20,000 neurons whose firing rate is governed by transcription and translation of a group of genes, the products of which are proteins that participate in a negative transcriptional feedback loop [1, 2]. Without external cues known as zeitgebers, the SCN exists in a free-running state that often varies from a strict 24-hour rhythm. Light is the strongest zeitgeber involved in entraining the SCN to a 24-hour rhythm, although other less effective cues also exist.

When the SCN and the peripheral clocks are entrained to a 24-hour rhythm synchronized to the outside environment, biological processes that depend on such synchronization are optimized. On the other hand, circadian disruptions can have adverse effects on a person's quality of life, with issues ranging from short-term problems such as cognitive impairment and digestive issues to chronic illnesses such as cardiovascular disease, diabetes, and certain cancers [3, 4]. One of the more important circadian rhythms consists of the drive for arousal known as Process C, which combines with the homeostatic Process S to determine the level of alertness or sleepiness of an individual at any given time of day; disturbances in wakefulness and sleepiness are ubiquitous in modern life and relief of these symptoms and their consequences are an important aspect of the study of chronobiology. These issues have spurred signif-

* Corresponding author at: Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY, United States.
E-mail address: ikec@rpi.edu (C.O. Ike).

https://doi.org/10.1016/j.heliyon.2022.e12500
Received 26 June 2022; Received in revised form 26 September 2022; Accepted 13 December 2022

icant research in the field of circadian control and estimation with the aim of developing techniques to assess and mitigate the effects of circadian disruption.

Multiple open-loop methods have been proposed to get the circadian system to a goal state. The bulk of these constitutes an optimal control approach to scheduling lighting in order to minimize the time to entrainment [5, 6, 7]. To control the rhythm in a feedback manner however, we need to be able to accurately assess the circadian state in real-time. The SCN, by virtue of its physical location, does not lend itself readily to direct measurements, so existing research has largely focused on using circadian signals as a proxy for the oscillator. A few of the signals most widely studied have been melatonin concentration (both in plasma and saliva), core body temperature, and activity levels [8]. Melatonin, due to its relative resistance to masking factors, has been the standard for circadian state estimation for decades. Particularly, studies have used the dim light melatonin onset (DLMO) - the time at which the melatonin concentration crosses a certain threshold in dim lighting - as a circadian phase marker since it was first proposed by Lewy et al. in [9]. This method has proven highly useful in clinical research, but DLMO can only be used in estimating the timing of phase markers or the phase shift between two points [10]. Moreover, melatonin values are not available in real-time but require laboratory processing. Consequently, there is still the need for continuous circadian phase estimation if the measurements are to be incorporated into an online feedback controller.

With the aim of eliminating the pitfalls of melatonin assessment and similarly invasive genetic assay techniques, recent research has focused on the use of easily accessible biometric data, most prominently activity levels, heart rate, skin temperature, and light exposure [11, 12, 13]. These measurements are used in estimating the circadian phase in a variety of ways - regression techniques, machine learning algorithms, and model-based approaches [14, 15]. Brown et al. frame the estimation problem as one of classification and use their method in estimating the DLMO phase marker with mean error of 1.3 hours [16]. Gil et al. use an auto-regressive moving average with exogenous input (ARMAX) model in estimating DLMO for 11 individuals [17]. Their work was particularly useful in showing that a combination of internal signals and external cues could improve the accuracy of the algorithm used. In [18], Cheng et al. combine actigraphy (rest/activity data monitored using a wrist-worn device) and light data in a model-based framework to predict the DLMO timing of fixed shift workers - a particularly difficult population to assess. Their method produced better correlations with in-lab DLMO measurements than agreement using sleep timing, providing a potential improvement to existing clinical practices in studying shift workers. Mott et al. in [19] use a particle filter in conjunction with the Kronauer model [14] in predicting the circadian phase marker. Woelders et al. [20] estimate the nadir of core body temperature (a phase marker) using measured ambient light and the Kronauer model. These methods largely still focus on the estimation of phase markers instead of continuous estimates. Yin et al. attempt to solve the continuous estimation problem in [10] by using an adaptive notch filter (ANF). Their approach is able to estimate the continuous circadian phase with appreciable accuracy, but the nonlinear system is rather complex and requires significant resources for tuning.

In this paper, we propose a method for estimating the continuous circadian phase using a linear state observer in conjunction with easily accessible biometric signals. By assuming a form on the composition of the biometric signal, we can design an observer for a multi-component harmonic oscillator, and then fit its gains to the user data using an evolutionary optimization algorithm. The optimization is structured such that the observer isolates target frequency components from the measured signals, allowing us reliably estimate the continuous circadian phase for individuals entrained to the light-dark cycle. Moreover, the online estimation and its computational efficiency open up the approach for use in a feedback control loop. We evaluate this observer-based filter (OBF) with actigraphy data collected for eight healthy subjects in ambulatory conditions. We compare its estimated phase shift with the DLMO values for each subject, along with the estimates from the ANF previously developed by our group and used in [10] for solving the same problem. We find that our approach provides comparable phase estimation with both methods at a fraction of the optimization run-time that was needed by the ANF, and without the use of invasive techniques required by the melatonin assessment.

2. Materials and methods

2.1. Experiment setup

2.1.1. Data collection

For this paper, we used the actigraphy data of eight healthy young adults. Five subjects were female and three were male, aged between 18 and 34 y (25.8 ± 6.6 y). All female subjects were pre-menopausal, though their menstrual cycles were not monitored during the study, as it's been shown that it has no effect on the circadian acrophase or period in non-controlled conditions [21, 22, 23].

While the data was collected over an 8-month period, the OBF was tested only on data over a 2-week period, ensuring weather and climate did not play a role. The actigraphy data was collected using an ActiGraph GT3X+ Monitor (Pensacola, FL) worn on the subject's non-dominant wrist. The device collected data at 1-minute intervals over the 2 weeks, providing near real-time information on the subjects. Ambient light intensity was also measured, though not considered in this paper.

The DLMO measurements were taken on the 7th and 14th days using saliva samples collected at 30-minute intervals with Salimetrics SalivaBio Oral Swabs (State College, PA). The samples were taken with the subjects in a supine position in dim light settings, starting approximately 5 hours before and ending 30 minutes after the average bedtime. Participants were instructed to accumulate saliva for 2 minutes, after which the samples were then stored at -20° Celsius. For analysis, the samples were thawed and centrifuged for 10 minutes at 2500 rpm.

All participants provided their informed written consent, and the experiments followed the principles in the Declaration of Helsinki from the World Medical Association. The experiments were monitored by the University of New Mexico (UNM) Health Sciences Center Human Research Protections office and approved by the UNM Institutional Review Board (IRB). The study's associated IRB number is 14-002.

2.1.2. Simulation environment

All numerical experiments were conducted using MATLAB R2021a and Simulink on a Dell workstation equipped with an Intel Core i7-3770 3.40 GHz processor and 16 GB RAM.

2.2. Observer-based filter

A state observer is a tool used in estimating the internal state of a target system given a model of its dynamics, and measurements of its inputs and outputs. When the target system is observable, the observer can be used to estimate the full state of the target, allowing us to solve many fundamental control problems, including feedback control. In the continuous-time case, we assume the state of an autonomous linear time-invariant (LTI) system satisfies

$$\dot{x}(t) = Ax(t)$$

$$y(t) = Cx(t)$$

where $x(t)$ is the system's state and $y(t)$ is the measurable output at time t, A represents the system dynamics, and C represents the output's relationship to the state. If this system is observable, then given the outputs, we can estimate the internal states using a state observer of the form

$$\dot{\hat{x}}(t) = (A - LC)\hat{x}(t) + Ly(t)$$

$$\hat{y}(t) = C\hat{x}(t)$$

where $\hat{x}(t)$ is the state estimate, $\hat{y}(t)$ is the estimated output, and L is the observer gain matrix. If the observer is appropriately designed to be asymptotically stable, the estimation error

$$e(t) = \hat{x}(t) - x(t)$$

converges to zero as $t \longrightarrow \infty$, implying that the estimate begins to mirror the true state given sufficient time. The speed of convergence can be increased by setting the poles of the continuous-time observer further to the left of the vertical axis in the complex-plane.

In this paper, we assume that the nominal circadian signal (e.g., actigraphy) is a periodic signal and as such, can be approximated arbitrarily closely using a Fourier series. The signal can then be represented as a sum of K harmonics and a bias term, in the form

$$y(t) = d + \sum_{i=1}^{K} a_i sin(i\omega^* t + \phi_i),$$

where ω^* is the fundamental frequency of the signal which we assume to be fixed at $\frac{2\pi}{24} rad/h$ (corresponding to a 24 hour period), d is the constant bias term, a_i is the amplitude, and ϕ_i is the phase offset of the i-th harmonic. The harmonic components of this sum, being sinusoidal, can each be generated by an appropriately designed linear harmonic oscillator. This fact allows us to assume the signal was generated by a harmonic oscillator model with multiple components, and to subsequently design a linear observer to approximate its output. Note that, in practice, d, a_i, and ϕ_i are time-varying. The filter that we develop is thus meant for tracking how these variables deviate from their nominal values over time.

We design an observer whose internal dynamics consists of K harmonically-related oscillators with fundamental frequency ω^*. By matching the output of this observer with the input signal, we are able to use the observer states in estimating the parameters a_i, ϕ_i, for $i \in \{1, \dots, K\}$, and d. The biometric signal thus forms the input to the observer, with which we estimate the internal state $\hat{x} \in \mathbb{R}^{2K+1}$, with $\hat{x}_{2i-1}$ and $\hat{x}_{2i}$ as the i-th harmonic and its derivative, respectively, and $\hat{x}_{2K+1}$ as the estimated bias term.

In this form, the OBF gains $L = \begin{bmatrix} L_1 & L_2 & \dots & L_{2K+1} \end{bmatrix}^T$ are the design variables of the algorithm and simply need to be tuned with an appropriate method to fit each individual's data. We select the system matrices

$$A = \begin{bmatrix} A_1 & 0 & \dots & 0 & 0 \\ 0 & A_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & A_K & 0 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}, \quad (1)$$

$$C = \begin{bmatrix} C_1 & C_2 & \dots & C_K & 1 \end{bmatrix}$$

with submatrices

$$A_i = \begin{bmatrix} 0 & 1 \\ -(i\omega^*)^2 & 0 \end{bmatrix},$$

$$C_i = \begin{bmatrix} 0 & \frac{2}{i\omega^*} \end{bmatrix},$$

$$i \in \{1, \dots, K\},$$

The choice of system dynamics A and output matrix C in Equation (1) represent our view of the biometric data as the sum of multiple harmonically-related components with a bias term, where K represents the target number of harmonics. The filter then yields the output

$$\hat{y}(t) = \hat{x}_{2K+1} + \sum_{i=1}^{K} \frac{2\hat{x}_{2i}}{i\omega^*}.$$

2.3. Gain optimization

To account for the variability in the human population, we tune the observer gains for every individual. To do so, we first quantify optimality using the cost function proposed in [10]

$$F(Y(\omega), \hat{Y}(\omega)) = J_{harmo} + J_{noise}, \quad (2)$$

with

$$J_{harmo} = \int_0^{\delta\omega} [Y(\omega) - \hat{Y}(\omega)]^2 \, d\omega + \sum_{n=1}^{K} \int_{n\omega^* - \delta\omega}^{n\omega^* + \delta\omega} [Y(\omega) - \hat{Y}(\omega)]^2 \, d\omega$$

and

$$J_{noise} = \sum_{n=0}^{K-1} \int_{n\omega^* + \delta\omega}^{(n+1)\omega^* - \delta\omega} \hat{Y}(\omega)^2 \, d\omega + \int_{K\omega^* + \delta\omega}^{+\infty} \hat{Y}(\omega)^2 \, d\omega,$$

where $\delta\omega$ represents the bandwidth around the frequency components and bias term that we want the filter to capture, J_{harmo} represents the square error around each component and the bias term, J_{noise} is the filtered output outside the desired components, K is the filter order being used, and $Y(\omega)$ and $\hat{Y}(\omega)$ are the Fourier transforms of the input biometric signal and the OBF output, respectively, defined

$$Y(\omega) \triangleq \int_0^{t_f} y(t) e^{-j\omega t} \, dt,$$

$$\hat{Y}(\omega) \triangleq \int_0^{t_f} \hat{y}(t) e^{-j\omega t} \, dt,$$

where t_f is the time of the final measurement.

The observer gains are then optimized using the evolutionary strategy [24] detailed in Algorithm 1 on the cost function in Equation (2). The optimization algorithm has μ initial members at each iteration, during which λ offspring are created by random combinations of ρ parents. The λ worst performing members of the population are then removed, along with their corresponding costs, and then the next iteration is started. During each iteration, if a population member is found to yield an unstable system, its cost is set to *int_max* (2147483647 on a 64-bit computer) to maximize the probability that it is removed from the population on that iteration. By doing this, we can also guarantee that the algorithm's output gain matrix will be stable provided the initial population includes at least one stable member.

2.3.1. Initial population

The initial population is created using a logarithmic sampling method between specified bounds. We start by creating a matrix $N \in \mathbb{R}^{\mu \times (2K+1)}$ of uniformly sampled points

$$n \sim \mathcal{U}(l, r),$$

where l is the left end and r is the right end of the uniformly sampled space. We then create the initial population $P \in \mathbb{R}^{\mu \times (2K+1)}$ with entries

$$p = \begin{cases} 10^{(n-m+LB)}, & n > m + 0.5 \\ 0, & m - 0.5 \le n \le m + 0.5 \\ -10^{(m-n+LB)}, & n < m - 0.5 \end{cases},$$

Algorithm 1: Gain Optimization.

Input: max_iterations, y, μ, λ, ρ

Initialize:
Create initial population as in 2.3.1
for $i = 1, \ldots, \mu$ **do**
  Create state space using population(i)
  if *created system is stable* **then**
    $\hat{y}(t) \longleftarrow OBF(y)$
    $Cost(i) \longleftarrow F(Y(\omega), \hat{Y}(\omega))$
  else
    $Cost(i) \longleftarrow int\_max$
  end
end

Iteratively improve population:
for $i = 1, \ldots, max\_iterations$ **do**
  for $j = \mu + 1, \ldots, \mu + \lambda$ **do**
    $Combination \longleftarrow \rho$ random population members
    population(j) $\longleftarrow mean(Combination)$
    Create state space using population(j)
    if *created system is stable* **then**
      $\hat{y}(t) \longleftarrow OBF(y)$
      $Cost(j) \longleftarrow F(Y(\omega), \hat{Y}(\omega))$
    else
      $Cost(j) \longleftarrow int\_max$
    end
  end
  Remove λ highest costs and corresponding population members
end
$L_{opt} \longleftarrow$ best member of final population
Result: Optimal gain matrix L_{opt}

where the midpoint $m = \frac{r+l}{2}$, and LB is the lower bound exponent. This gives an initial population with $\sim 45\%$ of its members negative, $\sim 45\%$ positive, and the rest set to zero. For example, with $LB = -5$, $l = 0$, and $r = 8$, we generate μ random gain vectors with elements uniformly distributed between $\pm 10^{-5}$ and $\pm 10^{-1}$.

2.4. Circadian phase estimation

Once the optimal gains have been obtained using the above procedure, we can then use them in estimating the continuous phase of the oscillator. We take the estimated circadian phase to be the argument of the first harmonic term

$$\theta(t) = \omega^* t + \phi_1(t) = -\tan^{-1}\left[\frac{x_2(t)}{\omega^* \cdot x_1(t)}\right]. \tag{3}$$

In our current experiments, we used the estimated phase shift between day 7 and 14 of the subject data in order to benchmark the filter's performance against the corresponding DLMO measurements. We also compare the OBF's estimate with that of the ANF. To calculate the phase shift for both filters, we use the formula

$$\Delta_{Filter} = \frac{1}{2\pi}\int_0^{24} \theta(t + 6 \times 24) - \theta(t + 13 \times 24)\, dt,$$

which is the difference between the average phase on day 7 and the average phase on day 14. A positive value here represents a phase delay, suggesting that circadian oscillation occured later in time on the 14th day than it did on the 7th. The phase shift gotten from DLMO data is calculated using the formula

$$\Delta_{DLMO} = T_{14} - T_7,$$

where T represents the time that the chosen melatonin threshold was reached. A positive value in this case also represents a phase delay suggesting that the DLMO happens later in time on the 14th day than it did on the 7th.

The entire process in this project is thus made up of the gain optimization followed by the phase estimation. Fig. 1 shows the usage envisioned for the filter, where the optimization can be run at set intervals, or when a certain amount of new data is available. The phase estimation can be carried out between optimizations and its results can then be used in a closed-loop control system.

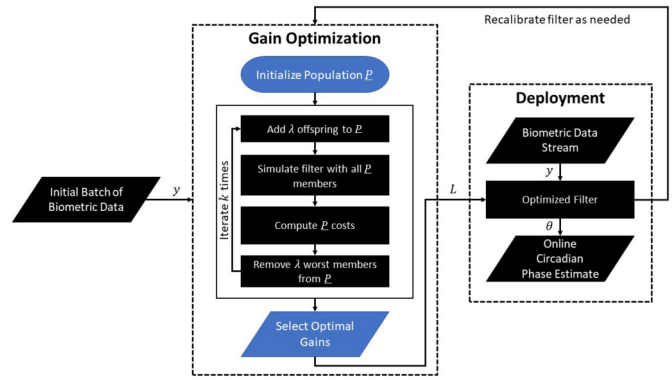

Fig. 1. Filter Optimization and Deployment Flowchart.

2.5. Adaptive notch filter

The adaptive notch filter was initially proposed by Mojiri et al. [25] to extract a sinusoid with variable frequency from a noisy signal. It was then modified for multi-harmonic signals with constant bias by Zhang et al. [26, 27] and then used in [28] for multiple inputs with the same fundamental frequency. The continuous-time ANF dynamics are given by

$$\dot{x}(t) = A_{ANF}(\hat{\omega})x + B_{ANF}(\hat{\omega})(y - \tilde{y}),$$
$$\dot{\hat{\omega}}(t) = -\gamma_\omega x_1 \hat{\omega}^2 (y - \tilde{y}),$$
$$\tilde{y}(t) = C_{ANF}(\hat{\omega})x,$$

where $\hat{\omega}$ is the ANF's continuous estimate of the fundamental frequency, $x \in \mathbb{R}^{2K+1}$ represents the ANF states, with x_{2i-1} and x_{2i} as the i-th harmonic and its derivative, respectively, and $\tilde{y}$ is the ANF's filtered output. The state space matrices are given as

$$A_{ANF}(\hat{\omega}) = \begin{bmatrix} A_1 & 0 & \ldots & 0 & 0 \\ 0 & A_2 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & A_K & 0 \\ 0 & 0 & \ldots & 0 & 0 \end{bmatrix},$$
$$B_{ANF}(\hat{\omega}) = \begin{bmatrix} B_1 & B_2 & \ldots & B_K & \gamma_d \end{bmatrix}^T,$$
$$C_{ANF}(\hat{\omega}) = \begin{bmatrix} C_1 & C_2 & \ldots & C_K & 1 \end{bmatrix}$$

with submatrices

$$A_i(\hat{\omega}) = \begin{bmatrix} 0 & 1 \\ -(i\hat{\omega})^2 & 0 \end{bmatrix},$$
$$B_i(\hat{\omega}) = \begin{bmatrix} 0 & (i\hat{\omega})^2 \end{bmatrix}^T,$$
$$C_i(\hat{\omega}) = \begin{bmatrix} 0 & \frac{2\zeta}{i\hat{\omega}} \end{bmatrix},$$
$$i \in \{1, 2, \ldots, K\}$$

In this form, there are three parameters which can be tuned to fit a subject's data – the damping factor ζ, the frequency adaptation γ_ω, and the bias adaptation γ_d. Yin et al. proposed a tuning algorithm in [10] to fit these parameters to each individual subject's data, yielding promising results on the actigraphy dataset.

Of note, however, is that the frequency adaptation and the attendant mutability of the estimated frequency $\hat{\omega}$ render this as a nonlinear time-varying system, making it more complex to optimize and analyze. The OBF eliminates the frequency adaptation based on the argument that doing so materially affects neither the cost of the achieved optimal filter, nor the resulting phase estimation in entrained settings. We know

from existing research that the average person's free-running circadian period is distributed tightly around 24.18 hours [29]. However, in ambulatory conditions where subjects are allowed to follow a daily routine and are exposed to the day-night cycle, it has been shown that the resulting entrained rhythm exhibits a period closer to 24 hours [30]. This is the known effect of zeitgebers, of which light is the strongest, that work to synchronize the biological clock to the 24-hour cycle. This fact allows us fix the period of our model at 24 hours and to view slight inter- and intra-individual variations in period as noise components in the output signal. From this perspective, the adaptability of the ANF opens it up to learning the noise parameters in the signal which then hampers its ability to accurately assess the circadian phase. Moreover, we found that in practice and in [10], the ANF frequency adaptation was consistently optimized to near zero, effectively freezing the signal period at 24 hours. The OBF eliminates this adaptation, and we see in the next section that it has better estimation power than the ANF on real actigraphy data, and comparable power for synthetic data with no phase corruption.

3. Results

In order to evaluate the accuracy and efficiency of the OBF, we look at the optimized filters' phase shift estimates, the optimization runtimes, and the final optimal costs. These numbers are all compared with those produced by the ANF after running it with identical optimizations. For the results discussed in this paper, each optimization ran for 50 iterations, starting with $\mu = 100$ random members of the population, $\rho = 2$ parents, and $\lambda = 50$ offspring at each iteration. We focus primarily on subjects 3 and 10 from [10], as 3 has the most detail across the filter orders studied in that paper, and subject 10 has complete DLMO information. The phase shifts and runtimes for subjects 4 to 9 are also shown in the appropriate sections.

First, we note Fig. 2 which was generated with optimized versions of both filters on subject 10's data. Fig. 2A shows the original input signal, along with the filter outputs. We see that the OBF is able to extract a clean periodic signal that closely follows the output of the ANF, which serves as an initial qualitative indication that the filter follows the desired behavior. In Fig. 2B, we show the continuous estimated phase offset $\phi_1(t)$ from both filters. This value was calculated by solving Equation (3) for $\phi_1(t)$ yielding

$$\phi_1(t) = \theta(t) - \omega^*(t)t. \tag{4}$$

As the ANF continuously estimates the frequency $\hat{\omega}(t)$, we used that output estimate in calculating the ANF's phase offset. However, as the OBF does not provide any estimate for the frequency, $\omega^*(t)$ was used, yielding the slightly different offset seen in the graph,

$$\phi_{OBF}(t) = \phi_1(t) + \omega_{diff}(t)t,$$

where $\omega_{diff}(t) = \omega^*(t) - \hat{\omega}(t)$ is the variation that occurs in that frequency due to disturbances. This has the effect that the OBF offset estimate does not differentiate deviations in the signal frequency from true occurrences of phase offset like the ANF can. As will be shown in subsequent subsections, however, this subtle difference between the filters has no effect on the OBF's ability to estimate the true phase shift.

3.1. Phase shift estimation accuracy

The most important performance metric for our algorithm is the accuracy of the phase shift estimates, as this is the major concern of our problem space. We focused primarily on whether the proposed OBF algorithm is able to provide results that are at least as good as those from the ANF. To tease out this information, we look at the estimated phase shifts from both methods and compare them with the phase shifts as calculated from the DLMO data. To visually compare the algorithms, we use Fig. 3, which shows the results for orders 1-5 on subject 3 and

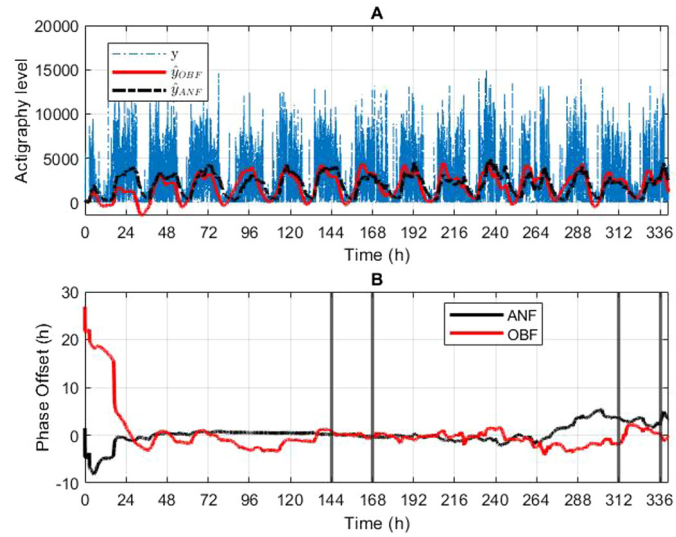

Fig. 2. (A) Actigraphy Data with Filtered Outputs from the OBF (Red) and the ANF (Black). (B) Online phase offset estimates from the OBF (Red) and ANF (Black).

10; and Fig. 4, which shows the performance of the third order filters on subjects 4 (4A), 5 (4B), 6 (4C), 7 (4D), 8 (4E), and 9 (4F). The DLMO values were not available for subjects 5 and 9, but their graphs are included here for completeness. We see that in the case of subject 3, the OBF estimates (Fig. 3A) appear as close to those given by the DLMO data as the ANF estimates (Fig. 3B), but with a tighter variance than the ANF values. For subject 10, the OBF estimates (Fig. 3C) are closer than the ANF's (Fig. 3D), with significantly lower run-to-run variance. To further quantify the filters' performance on these two subjects, we performed a one-tailed t-test on the absolute deviations of the filters' estimates from the mean of the DLMO estimates. Specifically, we tested the null hypothesis that the means of the deviations are equal, against the alternative hypothesis that the mean deviation of the ANF estimates is greater than that of the OBF. For subjects 3 and 10, the OBF outperformed the ANF in 70% of the cases, suggesting that the OBF is a solid option on these subjects' data. These results are shown in Table 1. Table 2 shows the average absolute deviation of the third order filter estimates from the average DLMO estimates for all subjects with valid DLMO data, along with the t-test results with the same hypotheses from above. We chose to focus on the third order filter as [10] found that the ANF did not see a material improvement in estimation power past the third order, and the OBF possesses a similar property as is seen in Fig. 3. We see that for subjects 3, 4, and 10, the OBF yields lower average estimation error with smaller run-to-run variance than the ANF. The t-test further confirms that the ANF deviations are greater with statistical significance. However, for subjects 6, 7, and 8, the ANF appears to yield lower average error rates. To gather a more complete picture of this, we conducted a second t-test on the results for subjects 6, 7, and 8, with the null hypothesis that the deviations were equal against the alternative that the OBF deviations were greater. The results are shown in Table 3, and we see that the OBF deviations were greater with statistical significance for only subject 8, thus showing that the OBF performs on par or better than the ANF for 83% of the subjects. These results are thus consistent with the preceding graphs in showing that the OBF is indeed a more efficient alternative to the ANF in entrained conditions.

3.2. Optimization runtimes

While optimization is currently done offline at a set cadence, in practical settings, it may need to be done more often based on the user's daily biometric data patterns (e.g., changes in lifestyle, work schedule, or wearable sensors). Moreover, the algorithms will be most useful

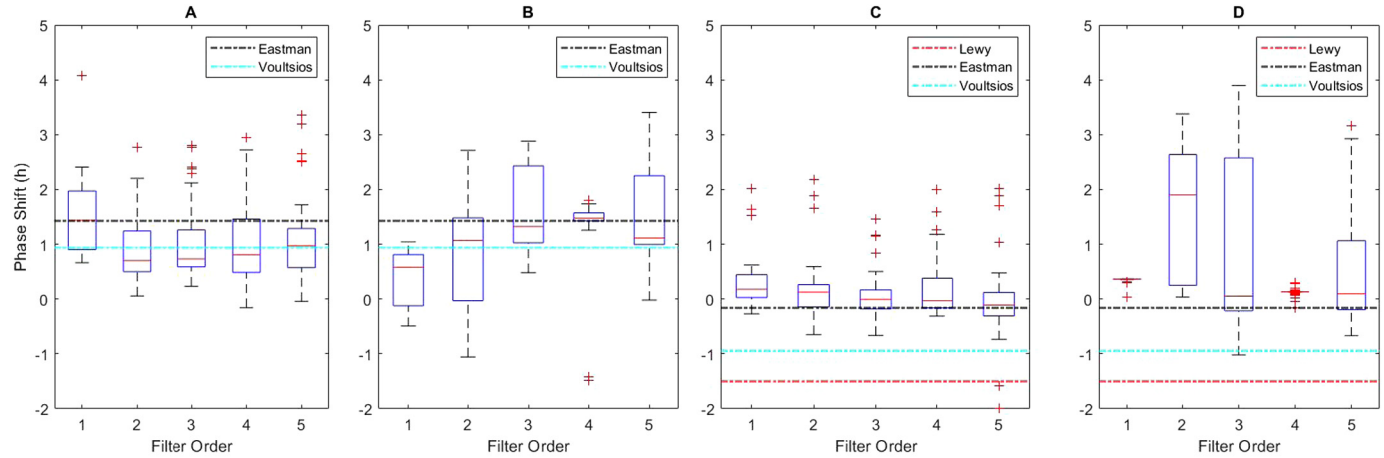

Fig. 3. Phase Shift Estimates from Actigraphy Data. (A) OBF Estimates for Subject 3. (B) ANF Estimates for Subject 3. (C) OBF Estimates for Subject 10. (D) ANF Estimates for Subject 10. Note: Box-plot variations are due to the random initialization of the filters' parameters for each optimization.

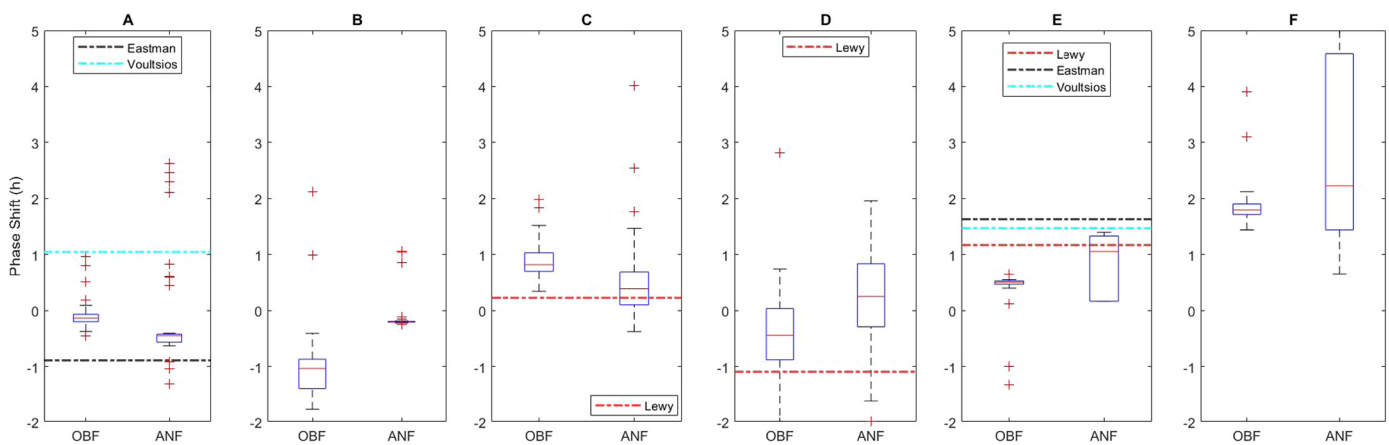

Fig. 4. Phase Shift Estimates for (A) Subject 4, (B) Subject 5, (C) Subject 6, (D) Subject 7, (E) Subject 8, and (F) Subject 9. Note: Box-plot variations are due to the random initialization of the filters' parameters for each optimization.

Table 1. Results from One-Tailed T-Tests on the Absolute Deviations of the Phase Shifts for Subjects 3 and 10 from Average DLMO Values.

Order	Subject 3			Subject 10		
	h[a]	p[b]	ci[c]	h[a]	p[b]	ci[c]
1	1	0.0498	$[1.7147 \times 10^{-4}, \infty]$	0	0.2274	$[-0.0688, \infty]$
2	1	0.0073	$[0.3521, \infty]$	1	5.2643×10^{-10}	$[1.0614, \infty]$
3	1	0.0112	$[0.3059, \infty]$	1	6.7813×10^{-4}	$[0.5078, \infty]$
4	0	0.2235	$[-0.3135, \infty]$	0	0.8086	$[-0.2225, \infty]$
5	1	0.0384	$[0.0522, \infty]$	1	0.0020	$[0.3213, \infty]$

[a] Hypothesis test results - 0 indicates mean deviations are statistically same, 1 indicates ANF deviations are higher.
[b] Significance level - smaller values cast doubt on validity of null hypothesis.
[c] Confidence interval where the results hold.

Table 2. Average Absolute Deviation (in minutes) from the Average DLMO Values By Subject.

Subject	OBF	ANF	h[a]	p[b]
3	37.2 (±20.4)	101.4 (±173.4)	1	0.0112
4	15 (±10.2)	57.6 (±71.9)	1	1.9696×10^{-4}
6	39.6 (±20.4)	32.4 (±41.4)	0	0.8447
7	79.8 (±132)	82.2 (±43.8)	0	0.4507
8	60.6 (±22.8)	36.6 (±32.4)	0	0.9999
10	56.4 (±26.4)	117.6 (±112.8)	1	6.7813×10^{-4}

[a] Hypothesis test results - 0 indicates mean deviations are statistically same, 1 indicates ANF deviations are higher.
[b] Significance level - smaller values cast doubt on validity of null hypothesis.

Table 3. Results of Additional 2-Sample T-Test By Subject.

Subject	h[a]	p[b]
6	0	0.1553
7	0	0.5493
8	1	0.0001

[a] Hypothesis test results - 0 indicates mean deviations are statistically same, 1 indicates OBF deviations are higher.
[b] Significance level - smaller values cast doubt on validity of null hypothesis.

when deployed on computationally constrained devices - most likely smartphones and activity trackers. In both cases, the algorithm efficiency is an important metric, as greater efficiency is tied to greater potential usefulness. To this end, we use the optimization runtimes as a proxy for evaluating each algorithm's overall complexity and efficiency. In Fig. 5, we see the boxplots of the runtimes for all 40 optimizations for orders 1-5 of both filters on subject 3, Figs. 5A and 5B, and subject 10, Figs. 5C and 5D, respectively. Fig. 6 shows the optimization runtimes for order 3 filters on subjects 4 (6A), 5 (6B), 6 (6C), 7 (6D), 8 (6E), and 9 (6F). Note that at 50 iterations, with 100 initial population members and 50 offspring, each optimization involved an upper bound of 2,600 simulations (fewer if unstable systems were encountered). We can easily see that the OBF gains a significant speedup over the ANF in both the median and worst case (on the order of 2.9x). The average OBF optimization runtime for all subjects was calculated to be 301.83 seconds, while the average ANF runtime was 866.82 seconds. This is a direct result of the comparative simplicity of the OBF's dynamics over the ANF's. Where the OBF is a linear, easily discretizable system, the

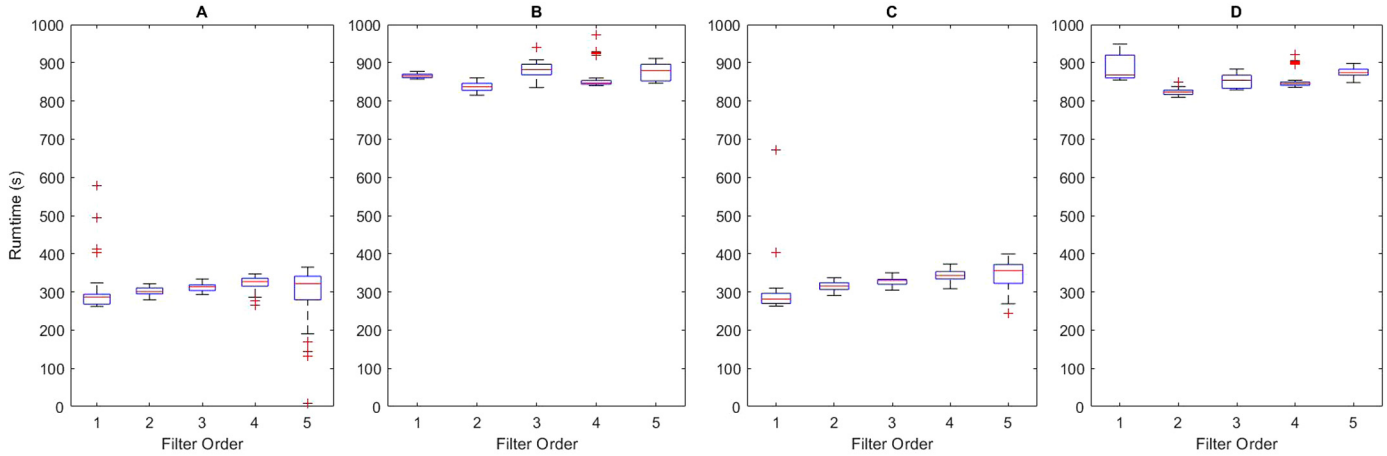

Fig. 5. Optimization Runtimes. (A) OBF Runtimes for Subject 3. (B) ANF Runtimes for Subject 3. (C) OBF Runtimes for Subject 10. (D) ANF Runtimes for Subject 10. Note: Box-plot variations are due to the random initialization of the filters' parameters for each optimization.

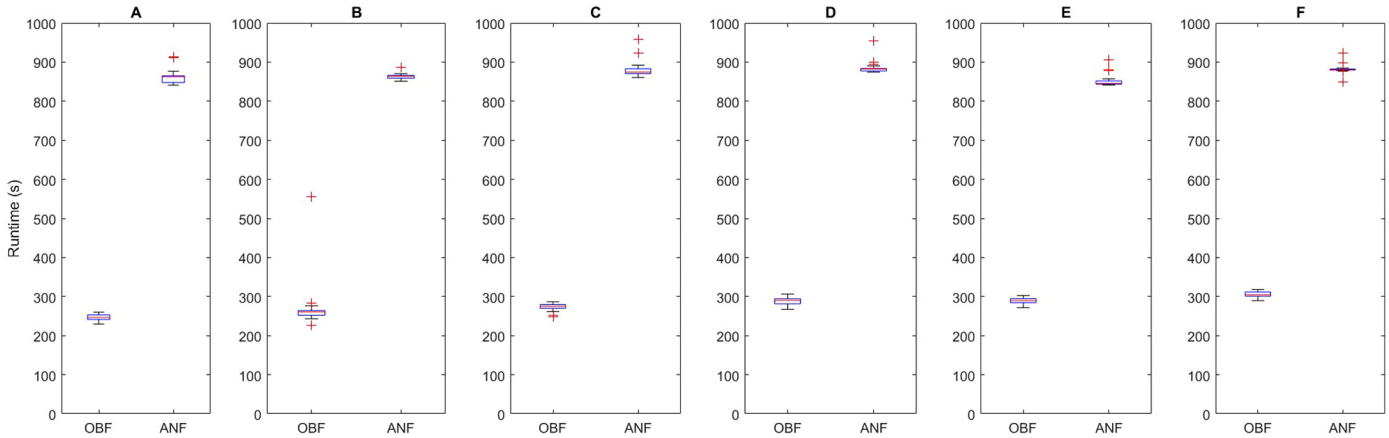

Fig. 6. Optimization Runtimes for (A) Subject 4, (B) Subject 5, (C) Subject 6, (D) Subject 7, (E) Subject 8, and (F) Subject 9. Note: Box-plot variations are due to the random initialization of the filters' parameters for each optimization.

ANF is nonlinear and was simulated in approximate continuous-time. We thus see that the OBF is able to outperform the ANF at a fraction of the adjustment runtime. This further bolsters the case for the OBF as a tool for phase estimation in entrained settings.

3.3. Optimal costs

The last object of comparison is the optimal costs obtained from the optimizations of both methods. We would like to see that the OBF optimization provides comparable optimal costs to those obtained when optimizing the ANF parameters. Fig. 7 shows the distributions of the optimal costs obtained on subject 3, Figs. 7A and 7B, and subject 10, Figs. 7C and 7D, respectively. We see that across orders, the median values of the OBF are comparable with those of the ANF. We do note that the variance of the OBF costs increases with filter order. We attribute this to the increasing dimension of the observer gain vector, which yields an increase in degrees of freedom that need optimization. To confirm this, we tested a population size of $\mu = 200$ for the optimizations, and found that the change mitigated the increase in variance to more closely follow the ANF's distributions. This had the trade-off of slightly increasing the optimization runtime, though the algorithm still ran much faster than the ANF. We do note, however, that even with the greater cost variance shown here, the OBF phase shifts were statistically better than those of the ANF, suggesting that the costs only represent a coarse evaluation of the filter's performance. To this end, we opted to retain the optimization population at $\mu = 100$.

Table 4. Results from One-Tailed T-Tests on the Optimal Costs for Subjects 3 and 10.

Order	Subject 3		Subject 10	
	h[a]	p[b]	h[a]	p[b]
1	1	6.9177×10^{-9}	0	0.9782
2	1	2.9230×10^{-14}	1	2.2651×10^{-33}
3	1	2.1160×10^{-7}	1	1.0024×10^{-11}
4	1	5.3230×10^{-27}	1	1.8861×10^{-10}
5	0	0.8558	0	0.6705

[a] Hypothesis test results - 0 indicates mean costs are statistically same, 1 indicates ANF costs are higher.

[b] Significance level - smaller values cast doubt on validity of null hypothesis.

We also conducted a one-tailed t-test on the optimal costs for both methods. In this case, we tested the populations for sameness against the alternative hypothesis that the means of the ANF costs are higher than those of the OBF. The results are presented in Table 4. We see that in 70% of the tested cases, the ANF optimal costs are statistically higher than those of the OBF. This yields another argument in favor of the OBF. In general, the OBF performs better than the ANF with our specified cost function, indicating that the filter more closely follows our specifications on the desired spectral content of the output signal.

3.4. Synthetic data testing

To further evaluate the OBF's performance in a controlled setting, we tested it on a simulated signal with known phase shift. We used the

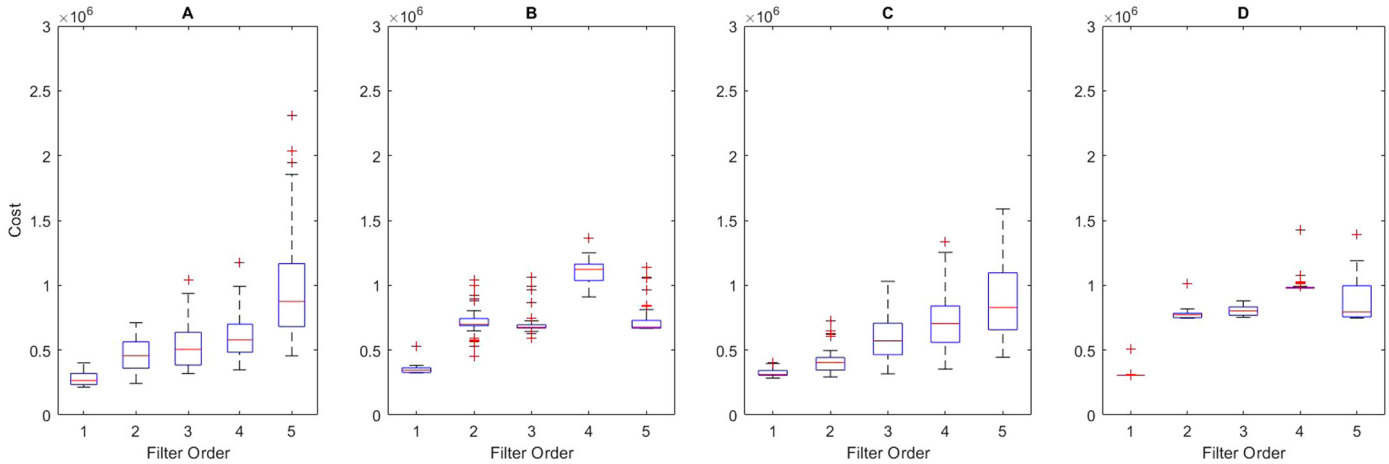

Fig. 7. Optimal Costs. (A) OBF Costs for Subject 3. (B) ANF Costs for Subject 3. (C) OBF Costs for Subject 10. (D) ANF Costs for Subject 10. Note: Box-plot variations are due to the random initialization of the filters' parameters for each optimization.

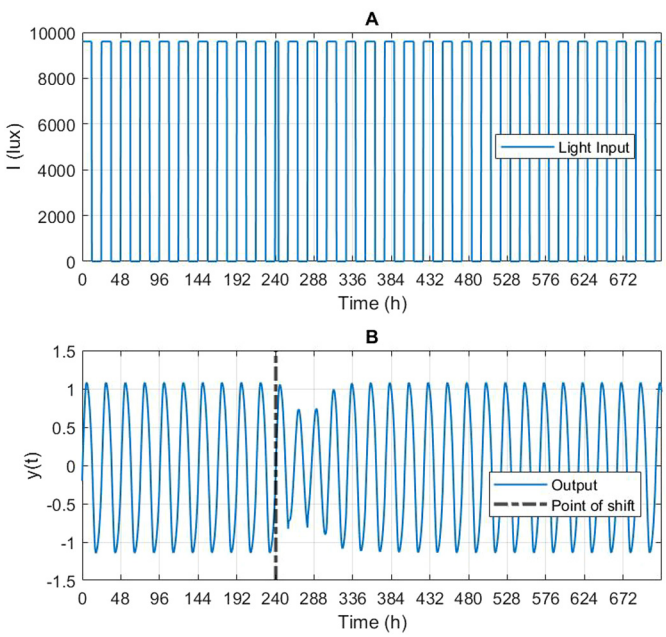

Fig. 8. (A) Light Input and corresponding (B) JFK Model Output.

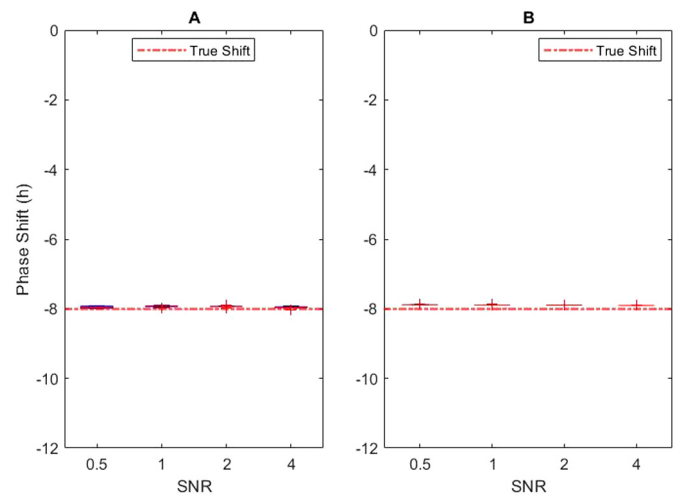

Fig. 9. Phase Shift Estimates from the (A) OBF and (B) ANF.

Jewett-Forger-Kronauer (JFK) model [14] in creating the 30-day signal shown in Fig. 8B. It is the result of a 10-day period of entrainment followed by a slam shift of 8 hours in the light input as seen in Fig. 8A, after which the oscillator is allowed to return to its limit cycle and run for an additional 20 days. We further corrupted the model output with 4 levels of white Gaussian amplitude noise specified by the signal-to-noise ratios $\mathrm{SNR} = 0.5, 1, 2, 4$, or $\mathrm{SNR}_{\mathrm{dB}} = -3$ dB, 0 dB, 6 dB, and 12 dB, respectively. We then ran 10 optimizations for the 3rd order ANF and OBF on these signals and compared the filters' performance. Specifically, we focused on the accuracy and precision of their phase shift estimates, as well as the obtained optimal costs.

In Fig. 9, we see the phase shift estimates from the OBF (9A) and the ANF (9B). On this simple signal, both filters yielded highly consistent results centered on the true shift. The degree of similarity between the estimates serves as the final confirmation that the OBF can replace the ANF in situations where the subject is entrained to the natural light-dark cycle (24 hours).

To further show this, Fig. 10A shows the noisy signal with SNR 0.5 overlayed with outputs from both optimized filters. Even with the extreme noise, both filters are able to consistently extract the clean original signal. The OBF possesses the same level of robustness to white Gaussian amplitude noise even in this extreme case. Fig. 10B shows the estimated phase offset from both filters calculated as in Equation (4), and we see that their performance on the signal remains nearly identical.

Finally, Fig. 11 shows that the OBF (11A) performs slightly better than the ANF (11B) across orders, with improvement on the order of 10%. However, as can be seen from the phase shifts, both filters are within 6 minutes of the true shift.

4. Conclusion

In this paper, we proposed a linear state observer as a tool for estimating an individual's continuous circadian phase from biometric data. In cases of circadian alignment, we expected that the OBF would perform at least as well as the ANF algorithm proposed in [10]. We experimentally validated these claims using real actigraphy data and further tested the algorithm on synthetically generated data. In both cases, we found that the OBF provided accurate estimates with a fraction of the computational cost of the ANF. In the real case, the OBF performed better than the ANF, while in the synthetic case, it performed identically to the ANF. These results lead us to believe that the OBF is the superior algorithm, particularly in situations where efficiency of implementation is important.

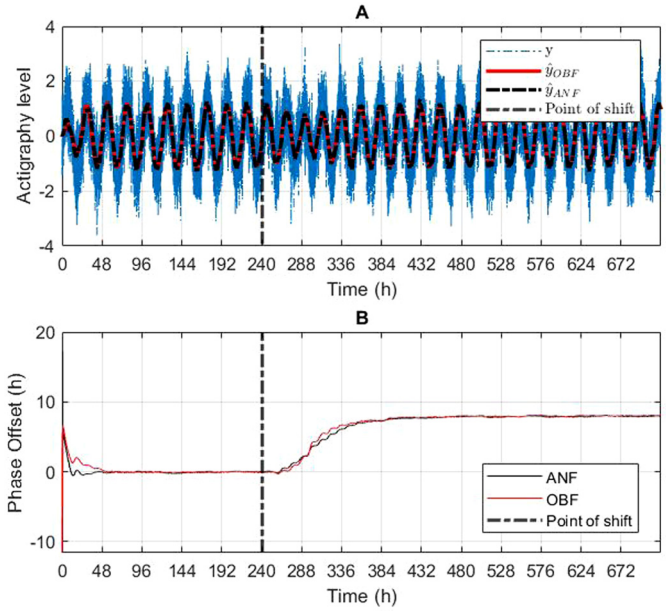

Fig. 10. (A) Synthetic Wave with 0.5 SNR with Outputs from the OBF (Red) and ANF (Black). (B) Phase offset estimates from the OBF (Red) and ANF (Black).

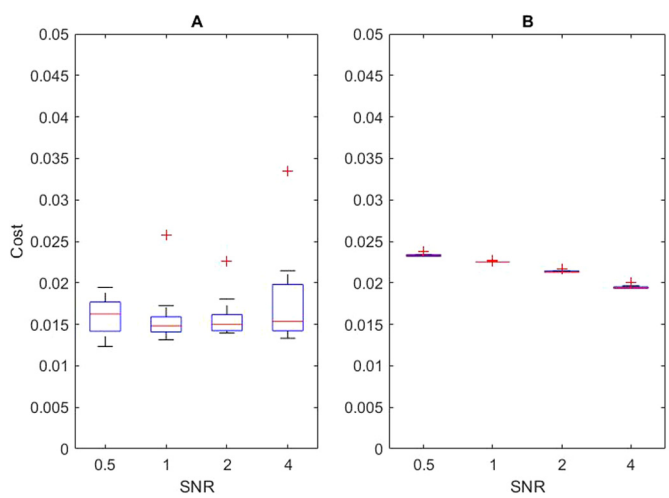

Fig. 11. Optimal Costs from the (A) OBF and (B) ANF. Note: Box-plot variations are due to the random initialization of the filters' parameters for each optimization.

The linearity of the system also opens up the algorithm to more widely researched methods and theory. In future work, we would like to further validate the OBF with more interesting data. The dataset used in this study was from individuals who followed a normal daily pattern and were not subjected to any drastic changes. We intend on testing the OBF on data from rotating shift workers, to assess the algorithm's robustness in such situations.

Moreover, we would like to assess the possibility of using multiple input signals to the algorithm. Existing approaches have shown improved performance with a combination of measurements of internal signals and external cues - light being most effective. An extension of the OBF to allow multiple inputs could further improve the overall accuracy of the algorithm in entrained ambulatory settings.

Declarations

Author contribution statement

Chukwuemeka O. Ike: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

John T. Wen: Analyzed and interpreted the data; Wrote the paper.

Meeko M. K. Oishi, Lee K. Brown: Contributed reagents, materials, analysis tools and data; Wrote the paper.

A. Agung Julius: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Funding statement

Anak Agung Julius was supported by National Science Foundation [DMS-2037357], Army Research Office [W911NF-17-1-0562 & W911NF-22-10039], New York State Foundation for Science, Technology and Innovation [C130145] and Smart Lighting Engineering Research Centre [EEC-0812056].

Data availability statement

Data will be made available on request.

Declaration of interests statement

The authors declare the following conflict of interests: Chukwuemeka Ike reports financial support was provided by National Science Foundation. Agung Julius, John Wen, Meeko Oishi, Lee Brown reports financial support was provided by National Science Foundation. Agung Julius and John Wen report financial support was provided by US Army Research Office. Agung Julius, John Wen, Meeko Oishi, Lee Brown reports a relationship with National Science Foundation that includes: funding grants. Agung Julius and John Wen report a relationship with US Army Research Office that includes: funding grants.

Additional information

No additional information is available for this paper.

References

[1] D.F. Swaab, E. Fliers, T. Partiman, The suprachiasmatic nucleus of the human brain in relation to sex, age and senile dementia, Brain Res. 342 (1) (1985) 37–44.
[2] S.R. Pandi-Perumal, M. Smits, W. Spence, V. Srinivasan, D.P. Cardinali, A.D. Lowe, L. Kayumov, Dim light melatonin onset (dlmo): a tool for the analysis of circadian phase in human sleep and chronobiological disorders, Prog. Neuro-Psychopharmacol. Biol. Psychiatry 31 (1) (2007) 1–11.
[3] M.H. Vitaterna, J.S. Takahashi, F.W. Turek, Overview of circadian rhythms, Alcohol Res. Health 25 (2) (2001) 85.
[4] A. Knutsson, Health disorders of shift workers, Occup. Med. 53 (2) (2003) 103–108.
[5] J. Zhang, J.T. Wen, A. Julius, Optimal circadian rhythm control with light input for rapid entrainment and improved vigilance, in: 2012 IEEE 51st IEEE Conference on Decision and Control (CDC), IEEE, 2012, pp. 3007–3012.
[6] K. Serkh, D.B. Forger, Optimal schedules of light exposure for rapidly correcting circadian misalignment, PLoS Comput. Biol. 10 (4) (2014) e1003523.
[7] J. Yin, A.A. Julius, J.T. Wen, Optimization of light exposure and sleep schedule for circadian rhythm entrainment, PLoS ONE 16 (6) (2021) e0251478.
[8] W.A. Hofstra, A.W. de Weerd, How to assess circadian rhythm in humans: a review of literature, Epilepsy Behav. 13 (3) (2008) 438–444.
[9] A.J. Lewy, R.L. Sack, The dim light melatonin onset as a marker for orcadian phase position, Chronobiol. Int. 6 (1) (1989) 93–102.
[10] J. Yin, A. Julius, J.T. Wen, M.M. Oishi, L.K. Brown, Actigraphy-based parameter tuning process for adaptive notch filter and circadian phase shift estimation, Chronobiol. Int. 37 (11) (2020) 1552–1564.
[11] D.-J. Dijk, J.F. Duffy, Novel approaches for assessing circadian rhythmicity in humans: a review, J. Biol. Rhythms 35 (5) (2020) 421–438.
[12] A.M. Reiter, C. Sargent, G.D. Roach, Finding dlmo: estimating dim light melatonin onset from sleep markers derived from questionnaires, diaries and actigraphy, Chronobiol. Int. 37 (9–10) (2020) 1412–1424.

[13] M.A. Bonmati-Carrion, B. Middleton, V. Revell, D.J. Skene, M. Rol, J.A. Madrid, Circadian phase assessment by ambulatory monitoring in humans: correlation with dim light melatonin onset, Chronobiol. Int. 31 (1) (2014) 37–51.
[14] M.E. Jewett, D.B. Forger, R.E. Kronauer, Revised limit cycle oscillator model of human circadian pacemaker, J. Biol. Rhythms 14 (6) (1999) 493–500.
[15] K.M. Hannay, V. Booth, D.B. Forger, Macroscopic models for human circadian rhythms, J. Biol. Rhythms 34 (6) (2019) 658–671.
[16] L.S. Brown, M.A.St. Hilaire, A.W. McHill, A.J. Phillips, L.K. Barger, A. Sano, C.A. Czeisler, F.J. Doyle III, E.B. Klerman, A classification approach to estimating human circadian phase under circadian alignment from actigraphy and photometry data, J. Pineal Res. (2021) e12745.
[17] E.A. Gil, X.L. Aubert, E.I. Møst, D.G. Beersma, Human circadian phase estimation from signals collected in ambulatory conditions using an autoregressive model, J. Biol. Rhythms 28 (2) (2013) 152–163.
[18] P. Cheng, O. Walch, Y. Huang, C. Mayer, C. Sagong, A. Cuamatzi Castelan, H.J. Burgess, T. Roth, D.B. Forger, C.L. Drake, Predicting circadian misalignment with wearable technology: validation of wrist-worn actigraphy and photometry in night shift workers, Sleep 44 (2) (2021) zsaa180.
[19] C. Mott, G. Dumont, D.B. Boivin, D. Mollicone, Model-based human circadian phase estimation using a particle filter, IEEE Trans. Biomed. Eng. 58 (5) (2011) 1325–1336.
[20] T. Woelders, D.G. Beersma, M.C. Gordijn, R.A. Hut, E.J. Wams, Daily light exposure patterns reveal phase and period of the human circadian clock, J. Biol. Rhythms 32 (3) (2017) 274–286.
[21] M.D. Coyne, C.M. Kesick, T.J. Doherty, M.A. Kolka, L.A. Stephenson, Circadian rhythm changes in core temperature over the menstrual cycle: method for noninvasive monitoring, Am. J. Physiol., Regul. Integr. Comp. Physiol. 279 (4) (2000) R1316–R1320.
[22] M. Ito, M. Kohsaka, K. Honma, N. Fukuda, S. Honma, Y. Katsuno, I. Kawai, H. Honma, N. Morita, T. Miyamoto, Changes in biological rhythm and sleep structure during the menstrual cycle in healthy women, Seishin Shinkeigaku Zasshi, Psychiatr. Neurol. Jpn. 97 (3) (1995) 155–164.
[23] A. Kivelä, A. Kauppila, P. Ylöstalo, O. Vakkuri, J. Leppäluoto, Seasonal, menstrual and circadian secretions of melatonin, gonadotropins and prolactin in women, Acta Physiol. Scand. 132 (3) (1988) 321–327.
[24] E. Ruffio, D. Saury, D. Petit, M. Girault, Tutorial 2: zero-order optimization algorithms, in: Eurotherm School METTI, 2011.
[25] M. Mojiri, M. Karimi-Ghartemani, A. Bakhshai, Time-domain signal analysis using adaptive notch filter, IEEE Trans. Signal Process. 55 (1) (2007) 85–93.
[26] J. Zhang, J.T. Wen, A.A. Julius, Adaptive circadian rhythm estimator and its application to locomotor activity, IEEE Signal Proc. Med. Biol. Symp. (2012) 1–6.
[27] J. Zhang, J.T. Wen, A.A. Julius, Adaptive circadian argument estimator and its application to circadian argument control, Am. Control Conf. (2013) 2295–2300.
[28] A. Julius, J. Zhang, W. Qiao, J.T. Wen, Multi-input adaptive notch filter and observer for circadian phase estimation, Int. J. Adapt. Control Signal Process. 30 (8–10) (2016) 1375–1388.
[29] C.A. Czeisler, J.F. Duffy, T.L. Shanahan, E.N. Brown, J.F. Mitchell, D.W. Rimmer, J.M. Ronda, E.J. Silva, J.S. Allan, J.S. Emens, et al., Stability, precision, and near-24-hour period of the human circadian pacemaker, Science 284 (5423) (1999) 2177–2181.
[30] C.A. Czeisler, J.S. Allan, S.H. Strogatz, J.M. Ronda, R. Sánchez, C.D. Ríos, W.O. Freitag, G.S. Richardson, R.E. Kronauer, Bright light resets the human circadian pacemaker independent of the timing of the sleep-wake cycle, Science 233 (4764) (1986) 667–671.