PHYSICAL REVIEW E **77**, 011922 (2008)

# Temporal differentiation and the optimization of system output

Emmanuel Tannenbaum[*]
*Department of Chemistry, Ben-Gurion University of the Negev, Be'er-Sheva, Israel*

(Received 4 June 2007; revised manuscript received 13 November 2007; published 31 January 2008)

We develop two simplified dynamical models with which to explore the conditions under which temporal differentiation leads to increased system output. By temporal differentiation, we mean a division of labor whereby different subtasks associated with performing a given task are done at different times. The idea is that, by focusing on one particular set of subtasks at a time, it is possible to increase the efficiency with which each subtask is performed, thereby allowing for faster completion of the overall task. In the first model, we consider the filling and emptying of a tank in the presence of a time-varying resource profile. If a given resource is available, the tank may be filled at some rate $r_f$. As long as the tank contains a resource, it may be emptied at a rate $r_e$, corresponding to processing into some product, which is either the final product of a process or an intermediate that is transported for further processing. Given a resource-availability profile over some time interval $T$, we develop an algorithm for determining the fill-empty profile that produces the maximum quantity of processed resource at the end of the time interval. We rigorously prove that the basic algorithm is one where the tank is filled when a resource is available and emptied when a resource is not available. In the second model, we consider a process whereby some resource is converted into some final product in a series of three agent-mediated steps. Temporal differentiation is incorporated by allowing the agents to oscillate between performing the first two steps and performing the last step. We find that temporal differentiation is favored when the number of agents is at intermediate values and when there are process intermediates that have long lifetimes compared to other characteristic time scales in the system. Based on these results, we speculate that temporal differentiation may provide an evolutionary basis for the emergence of phenomena such as sleep, distinct REM and non-REM sleep states, and circadian rhythms in general. The essential argument is that in sufficiently complex biological systems, a maximal amount of information and tasks can be processed and completed if the system follows a temporally differentiated "work plan," whereby the system focuses on one or a few tasks at a time.

DOI: 10.1103/PhysRevE.77.011922 PACS number(s): 87.18.Yt, 87.55.N−, 87.55.−x, 87.85.Xd

## I. INTRODUCTION

Differentiation and the division of labor is a ubiquitous phenomenon characterizing the emergence of complex systems. Different enzymes, nucleic acids, and other biopolymers are involved in the proper function of living cells. In multicellular organisms, cells differentiate and specialize in the performance of one or a few tasks. At higher levels of complexity, multicellular organisms (e.g., humans) can themselves form highly differentiated structures (e.g., a modern networked economy), where each organism performs one or a few tasks [1–13].

As a result of the ubiquity of the division of labor in biology, considerable experimental and theoretical work has been devoted to understanding both its genetic basis and the selection pressures that give rise to such behaviors [14–21]. In a general sense, division of labor is favored when transport costs associated with delivering process intermediates to the appropriate agents are small. Therefore, division of labor is generally favored at high population densities, though this may not always be the case [10,22].

In this paper, we wish to discuss another form of differentiation that we term *temporal differentiation*. Temporal differentiation refers to a division of labor where a given task is broken up into several subtasks and the various subtasks are performed at different times. That is, with temporal differentiation, a given set of agents performs all the subtasks associated with a given task. However, these agents concentrate their efforts on one set of subtasks for a certain period of time and then concentrate their efforts on another set of subtasks for another period of time. This is in contrast to the "standard" picture of division of labor, whereby all subtasks associated with a given task are performed simultaneously by different sets of agents.

Although the concept of time-varying strategies has been considered before [23–25], temporal differentiation as a form of division of labor has thus far received little to no attention. This is not surprising, since temporal differentiation is a subtle form of division of labor. Nevertheless, temporal differentiation is also a ubiquitous phenomenon. At the level of task completion by humans, it is quite common that various tasks are often done in intermittent blocks. Examples include paying of bills, housekeeping chores, and the procurement of food. Such forms of temporal differentiation are likely prevalent in other organisms, since a temporally differentiated labor strategy is a natural approach in many contexts for increasing system efficiency (i.e., it makes sense to have a routine or "work plan" and not try to do everything at once).

Another possible manifestation of temporal differentiation is the phenomenon of sleep, and, more generally, of circadian rhythms. In regards to sleep, it may be readily observed that sleep is prevalent in organisms with highly complex central nervous systems. It is characterized by periods of high levels

---

[*] emanuelt@bgu.ac.il

of alertness, followed by periods during which the brain goes into an unconscious state.

One theory, due to Crick and Mitchison, for the existence of sleep, is that sleep is a time when the brain engages in various garbage collection activities [26]. More specifically, sleep is a time when the brain sorts through and consolidates information accumulated from the previous period of wakefulness [27–32]. Increasing evidence suggests that Crick and Mitchison's basic hypothesis may be correct [33]. Nevertheless, while this "garbage collection" hypothesis provides an explanation for what happens during sleep, it does not explain the selection pressures driving the emergence of this phenomenon.

In the presence of a day-night cycle, it may be optimal for a highly complex brain to engage in information collection activities during the day, when light information is most available, and then to engage in information consolidation (i.e., "garbage collection") activities at night, when light information is far less available. The idea is that, by concentrating on information collection when it is available and information consolidation when external information is less available, the brain can process a maximal amount of information in a given amount of time, which presumably confers a survival advantage to the organism.

However, such an explanation is incomplete, since some organisms are nocturnal. Further, what is interesting about the sleep state itself is that it is divided into distinct, alternating cycles of REM and non-REM sleep (where REM stands for "rapid eye movement") [26,27,34]. However, not all sleeping organisms exhibit REM and non-REM sleep. In particular, the organism *Tachyglossus aculeatus*, a representative of the earliest branch of mammalian evolution (the monotremes), combines both REM and non-REM sleep into one sleep state [35]. This suggests that REM and non-REM sleep are not fundamental to sleep itself, but rather emerged via the differentiation of a single, older sleep state.

Therefore, it is possible that temporal differentiation of various brain tasks leads to improved brain function in higher organisms, independent of any external day-night regulation cycle. The existence of a day-night cycle simply regulates the optimal start times for each task period. Furthermore, it is also possible that temporal differentiation is favored the more work needs to be completed within a given time period. In the context of REM and non-REM sleep, it is possible that mammals with simpler brains can perform the information consolidation tasks in one sleep state. However, for larger, more complex mammalian brains, the amount of information consolidation to be performed becomes sufficiently large that it becomes more efficient to divide the various information consolidation subtasks into two distinct sleep states.

Motivated by these various considerations, we present two highly simplified models describing the processing of some resource into a final product. The first model considers the filling and emptying of a tank with a time-varying resource flow. Filling the tank corresponds to the input of external resource or information into a system, while emptying the tank corresponds to processing the resource or information. Not surprisingly, the model considered here is able to capture certain basic features of the wake-sleep cycle, such as its synchronization with the day-night cycle (at least for diurnal organisms).

The second model considers the processing of some resource into a final product via a series of agent-mediated, or in the language of chemical kinetics, enzyme-catalyzed, steps. In contrast to the tank-filling model, where the resource flow is time varying to mimic a day-night cycle, in this second model the resource flow is constant. The point of this second model is to illustrate how temporal differentiation can lead to increased output even when the inputs to the system are not time varying.

While the connection between temporal differentiation, sleep, and circadian rhythms is speculative, what is interesting is that if sleep is indeed an example of a temporally differentiated labor strategy that optimizes brain function, then it suggests that techniques from operations research, game theory, and mathematical economics could be used to model various aspects of sleep [37–39]. Because these fields have been fairly well developed, a fundamental connection between sleep and temporal differentiation could pave the way for the development of sophisticated models that can make quantitative predictions about various aspects of brain function that are related to the phenomenon of sleep.

We point out that while mathematical models related to sleep have been previously developed [40–43], such models have primarily focused on the dynamics associated with the underlying pathways controlling the wake-sleep cycle. However, the need for such pathways has not been addressed: nor were the models framed in a manner illustrating a fundamental connection between sleep and problems in engineering and behavioral sciences.

This paper is organized as follows: In the following section (Sec. II), we develop the tank-filling model. We develop a canonical algorithm for maximizing the amount of resource that is processed in a given time interval and rigorously prove that this algorithm is indeed optimal. In Sec. III, we use our algorithm to construct optimal fill-empty profiles for several resource availability profiles and compare these fill-empty profiles to known aspects of the sleep-wake cycle. In Sec. IV, we develop our three-process model and derive the limiting form of the model when the total number of enzymes is small. We solve our model with and without temporal differentiation. We compare both the temporally differentiated and nondifferentiated strategies, and determine the regimes where one strategy is expected to be favored over the other. In Sec. V, we discuss our results in the context of sleep, circadian rhythms, and other examples of temporal differentiation. Finally, in Sec. VI we summarize our main conclusions and discuss plans for future research.

## II. TANK-FILLING MODEL

### A. Model description

Consider a tank that can be filled and emptied with some unspecified material (see Fig. 1). At any given time $t$, an external resource is available or it is not. We denote the resource availability profile by a function $\delta(t)$, where $\delta(t)=1$ if a resource is available at time $t$ and 0 if not. If $\delta(t)=1$, then the tank may be filled at a rate $r_f$. The tank may also

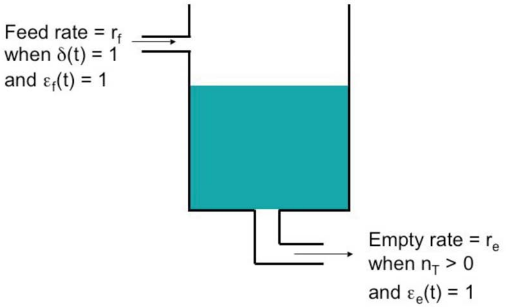

FIG. 1. (Color online) Illustration of a tank that may be filled at a rate $r_f$ when an external resource is present or emptied at a rate $r_e$ when there is material in the tank to empty.

be emptied at a rate $r_e$ as long as it is not empty. Over any finite interval $[t_1,t_2]$, we assume that $\delta(t)$ is discontinuous at a finite number of points. This implies that $\delta(t)$ may be taken to be a piecewise constant function.

At any given time $t$, we assume that the tank is carrying out one of the fill or empty tasks, but that the tasks cannot occur simultaneously. We let $\epsilon_f(t)$ denote the fill profile function, so that $\epsilon_f(t)=1$ if the tank is in the fill mode at time $t$ and $\epsilon_f(t)=0$ otherwise. We let $\epsilon_e(t)$ denote the empty profile function, so that $\epsilon_f(t)+\epsilon_e(t)=1$ at all times. As with $\delta(t)$, we assume $\epsilon_f(t)$ and $\epsilon_e(t)$ are piecewise constant, with a finite number of discontinuities over a finite interval.

We also let $n_T(t)$ denote the total amount of material in the tank at time $t$ and $n_P(t)$ denote the total amount of material that has been processed through the tank at time $t$. It should be apparent that

$$\frac{dn_T}{dt} = r_f\epsilon_f(t)\delta(t) - r_e\epsilon_e(t)(1-\delta_{n_T,0}),$$

$$\frac{dn_P}{dt} = r_e\epsilon_e(t)(1-\delta_{n_T,0}), \tag{1}$$

where $\delta_{n_T,0}=1$ if $n_T=0$ and 0 otherwise.

Although we assume that the fill and empty modes of the tank cannot occur simultaneously, a simultaneous fill-empty profile can be approximated with an arbitrary degree of accuracy by considering a profile where the tank oscillates between the fill and empty modes on sufficiently short time scales. Therefore, there is no loss of generality in assuming that the fill and empty modes occur separately. If there is a time associated with switching between the two tasks, then we do need to explicitly consider a mode where the fill and empty tasks operate simultaneously. However, for the purposes of this paper, we assume that the switching time between tasks is negligible.

Finally, while we are considering the filling and emptying of a tank, it should not be assumed that the actual fill-empty dynamics is necessarily governed by the rules of fluid mechanics. The external resource simply represents an input into a system, represented by a tank. The above system of differential equations define a highly simplified dynamics that broadly represents the dynamics associated with the input and processing of some resource.

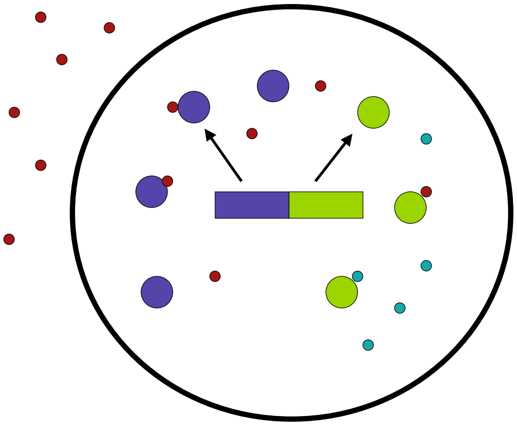

FIG. 2. (Color online) Illustration of the two-gene resource processing model discussed in this paper. The dark blue rectangle denotes gene $G_1$ (black rectangle), and the dark blue circles denote protein $P_1$ (black circles), while the green rectangle (white rectangle) and circles (white circles) denote gene $G_2$ and protein $P_2$, respectively. The red circles (small black circles) denote the available external resource, which enters the system via $P_1$. The proteins $P_2$ process the resource. Processed resource is denoted by the small blue-green circles (small white circles).

### B. Relation of the tank-filling model to biological networks

Consider a system that is capable of engaging in two tasks: (i) an input task, whereby external resource is input into the system, and (ii) a processing task, whereby the resource is processed for use by the system.

If we are dealing with a biological system, then we assume that the function of each task is mediated by some protein. For the input tasks, we denote the protein by $P_1$, while for the processing tasks, we denote the relevant protein by $P_2$. Now, $P_1$ and $P_2$ are encoded in a genome, with corresponding genes denoted by $G_1$ and $G_2$. If we assume a total transcription plus translation rate of $k_T$, then when only one gene is active, the active gene $G_i$ produces protein at a rate $k_T$, while when both genes are active, each protein is produced at a rate $(1/2)k_T$. We assume that at least one of the genes is active at any given time (see Fig. 2).

Furthermore, we assume that the proteins have a decay rate given by a first-order constant $k_d$. Therefore, if $\epsilon_1(t)$ is defined to be 0 if gene $G_1$ is off, and 1 if gene $G_1$ is on, and $\epsilon_2(t)$ is defined analogously for gene $G_2$, then, letting $n_{P_1}$ and $n_{P_2}$ denote the number of proteins in the system at a given time, we have

$$\frac{dn_{P_1}}{dt} = \frac{1}{2}(\epsilon_1(t)-\epsilon_2(t)+1)k_T - k_d n_{P_1},$$

$$\frac{dn_{P_2}}{dt} = \frac{1}{2}(\epsilon_2(t)-\epsilon_1(t)+1)k_T - k_d n_{P_2}. \tag{2}$$

Now, if we let $\delta(t)$ denote a resource availability profile, defined to be 1 when an external resource is available and 0 otherwise, then in the simplest assumption the rate at which resource enters the system is proportional to $n_{P_1}$ and the rate at which resource is processed is proportional to $n_{P_2}$. If $n_1$ denotes the amount of unprocessed resource at any given time $t$ and $n_2$ denotes the amount of processed resource, then we have

$$\frac{dn_1}{dt} = r_1 n_{P_1}\delta(t) - r_2 n_{P_2}(1-\delta_{n_1,0}),$$

$$\frac{dn_2}{dt} = r_2 n_{P_2}(1-\delta_{n_1,0}), \tag{3}$$

where $\delta_{n_1,0}=1$ if $n_1=0$ and 0 otherwise.

Note then that if $\delta(t)$ experiences oscillatory periods of resource availability, then keeping both genes on at all times may not lead to an optimal processing of resource. The reason for this is that when $\delta(t)=1$, the resource only enters the system at half the maximal rate. Depending on the resource availability profile, it may be optimal for the system to only take in an external resource when it is available and then process that resource during periods when the resource is significantly less available.

We note that for finite values of $k_T$ and $k_d$, switching from one task to another involves a transient during which the proteins involved in one task degrade and the proteins for the other task reach their steady-state levels. However, if $k_T, k_d \to \infty$ in such a way that $k_T/k_d$ is fixed, then there is no time associated with switching tasks. In this case, we may assume, for simplicity, that only one task can be active at any given time, since a profile where both tasks are on over a time interval can be approximated to arbitrary accuracy by a profile that rapidly oscillates between one task and the other.

### C. Optimal fill-empty profiles

Given a resource availability profile $\delta(t)$ over some time interval $[0,T]$, we wish to determine the fill-empty profile $\epsilon=(\epsilon_f,\epsilon_e)$ that maximizes $n_P(T)$, given the initial conditions $n_T(0)=n_P(0)=0$. As a notational convenience, we let $n_{\epsilon,T}(t)$, $n_{\epsilon,P}(t)$ denote the $n_T$ and $n_P$ values associated with the fill-empty profile $\epsilon$.

A natural fill-empty profile, denoted $\epsilon_0=(\epsilon_{f,0},\epsilon_{e,0})$, is defined by the following prescription: Fill whenever $\delta(t)=1$ and empty whenever $\delta(t)=0$ as long as $n_T(t)>0$. Continue with this fill-empty profile until $n_T(t)=r_e(T-t)$, at which point the tank should be emptied until time $T$. Let $t_{\epsilon_0}$ denote the critical time at which emptying until time $T$ begins. Then for notational convenience, we define $I_1=[0,t_{\epsilon_0}]$ and $I_2=[t_{\epsilon_0},T]$.

We can prove that $\epsilon_0$ yields a maximal value for $n_P(T)$. However, because the proof makes use of some definitions that will be used in the subsequent analysis, we will first provide these definitions before going on to the proof itself.

We begin by defining, for an arbitrary fill-empty profile $\epsilon$ over some set $\mathbf{S}$, the quantities $T_f(\mathbf{S};\epsilon)$, $T_e(\mathbf{S};\epsilon)$, and $T_w(\mathbf{S};\epsilon)$, as follows: We define $\mathbf{S}_f=\{t\in\mathbf{S}\,|\,\delta(t)=1$ and $\epsilon_f(t)=1\}$, $\mathbf{S}_e=\{t\in\mathbf{S}\,|\,n_{\epsilon,T}(t)>0$ and $\epsilon_e(t)=1\}$, and $\mathbf{S}_w=\mathbf{S}/(\mathbf{S}_f\cup\mathbf{S}_e)$. If $\mu(\Omega)$ denotes the measure of a set $\Omega$ (essentially the total length of the set), then $T_f(\mathbf{S};\epsilon)\equiv\mu(\mathbf{S}_f)$, $T_e(\mathbf{S};\epsilon)\equiv\mu(\mathbf{S}_e)$, and $T_w(\mathbf{S};\epsilon)\equiv\mu(\mathbf{S}_w)$. We should point out that because $\delta$, $\epsilon_f$, and $\epsilon_e$ are assumed to be piecewise constant, all sets considered in this paper are unions of disjoint intervals and hence are measurable.

Given a set $\mathbf{S}$, define $\mathbf{S}^0=\{t\in\mathbf{S}\,|\,\delta(t)=0\}$ and $\mathbf{S}^1=\{t\in\mathbf{S}\,|\,\delta(t)=1\}$. Then define $T_{e,0}(\mathbf{S};\epsilon)=\mu(\mathbf{S}_e^0)$ and $T_{e,1}(\mathbf{S};\epsilon)=\mu(\mathbf{S}_e^1)$. Note that $T_e(\mathbf{S};\epsilon)=T_{e,0}(\mathbf{S};\epsilon)+T_{e,1}(\mathbf{S};\epsilon)$. Also, note that since $\epsilon_{0,f}(t)=1$ whenever $\delta(t)=1$ for $t\in I_1$, it follows that $T_f(I_1;\epsilon)\leqslant T_f(I_1;\epsilon_0)-T_{e,1}(I_1;\epsilon)$.

We should note that, although the optimal fill-empty profile $\epsilon_0$ may not necessarily be unique, if $\epsilon$ denotes any other optimal fill-empty profile, then we must have $T_e([0,T];\epsilon)=T_e([0,T];\epsilon_0)$. It may be readily shown that $n_{\epsilon,T}(T)=n_{\epsilon_0,T}(T)=0$.

An $\epsilon$ for which $n_{\epsilon,T}(T)>0$ is not optimal, for letting $t_\epsilon$ denote when $n_{\epsilon,T}(t)=r_e(T-t)$ (by the intermediate value theorem, such a $t$ exists), we have $n_{\epsilon,P}(T)=n_{\epsilon,P}(t_\epsilon)+r_eT_e([t_\epsilon,T];\epsilon)\leqslant n_{\epsilon,P}(t_\epsilon)+r_e(T-t_\epsilon)$, with equality only occurring when $T_e([t_\epsilon,T];\epsilon)=r_e(T-t_\epsilon)$. However, $T_e([t_\epsilon,T];\epsilon)=r_e(T-t_\epsilon)$ implies that $n_{\epsilon,T}(T)=0$, and so our claim is proved.

But this implies that $r_fT_f([0,T];\epsilon)=r_eT_e([0,T];\epsilon)=r_eT_e([0,T];\epsilon_0)=r_fT_f([0,T];\epsilon_0)$, so that $T_f([0,T];\epsilon)=T_f([0,T];\epsilon_0)$. Finally, $T_w([0,T];\epsilon)=T-T_f([0,T];\epsilon)-T_e([0,T];\epsilon)=T-T_f([0,T];\epsilon_0)-T_e([0,T];\epsilon_0)=T_w([0,T];\epsilon_0)$.

Therefore, although the optimal fill-empty profile may not be unique, the $T_f$, $T_e$, and $T_w$ values are uniquely specified.

### D. Proof of the optimality of the fill-empty profile $\epsilon_0$

For any fill-empty profile $\epsilon$, we have

$$\begin{aligned} n_{\epsilon,T}(T) &= n_{\epsilon,T}(t_{\epsilon_0}) + r_fT_f(I_2;\epsilon) - r_eT_e(I_2;\epsilon) \\ &= r_fT_f(I_1;\epsilon) - r_eT_e(I_1;\epsilon) + r_fT_f(I_2;\epsilon) - r_eT_e(I_2;\epsilon). \end{aligned} \tag{4}$$

Since $T_f(I_2;\epsilon)+T_e(I_2;\epsilon)\leqslant T-t_{\epsilon_0}$, we have

$$\begin{aligned} 0 \leqslant n_{\epsilon,T}(T) \leqslant{}& r_f[T_f(I_1;\epsilon_0) - T_{e,1}(I_1;\epsilon)] - r_eT_e(I_1;\epsilon) \\ &+ r_f[T - t_{\epsilon_0} - T_e(I_2;\epsilon)] - r_eT_e(I_2;\epsilon), \end{aligned} \tag{5}$$

which may be rearranged to give

$$\begin{aligned} T_e(I_2;\epsilon) \leqslant{}& \frac{r_fT_f(I_1;\epsilon_0) - r_eT_e(I_1;\epsilon)}{r_e} \\ &- \frac{r_f}{r_f+r_e}[T_e(I_1;\epsilon_0) + T_{e,1}(I_1;\epsilon) - T_e(I_1;\epsilon)], \end{aligned} \tag{6}$$

where the derivation of this inequality makes use of the identity $r_fT_f(I_1;\epsilon_0)-r_eT_e(I_1;\epsilon_0)=r_e(T-t_{\epsilon_0})$.

We then obtain that

$$n_{\epsilon,P}(T) = r_e[T_e(I_1;\epsilon) + T_e(I_2;\epsilon)] \leqslant r_f T_f(I_1;\epsilon_0) - \frac{r_f r_e}{r_f + r_e}[T_e(I_1;\epsilon_0) + T_{e,1}(I_1;\epsilon) - T_e(I_1;\epsilon)]. \quad (7)$$

Since $n_{\epsilon_0,P}(T)=r_fT_f(I_1;\epsilon_0)$, then if we can show that $T_e(I_1;\epsilon_0)+T_{e,1}(I_1;\epsilon)-T_e(I_1;\epsilon)\geqslant 0$, we will have established that $n_{\epsilon,P}(T)\leqslant n_{\epsilon_0,P}(T)$, thereby proving the maximality of $\epsilon_0$.

So suppose $T_e(I_1;\epsilon_0)+T_{e,1}(I_1;\epsilon)-T_e(I_1;\epsilon)<0$. Then, defining $I(t)=[0,t]$, we may note that the function $\tau(t) \equiv T_e(I(t);\epsilon_0)+T_{e,1}(I(t);\epsilon)-T_e(I(t);\epsilon)$ is continuous and satisfies $\tau(0)=0$, $\tau(t_{\epsilon_0})<0$. Let us then define $t^*=\inf\{t \in [0,t_{\epsilon_0}] \mid \tau(t)<0\}$. By continuity of $\tau$ and from the definition of inf, we have that $t^*<t_{\epsilon_0}$, $\tau(t^*)=0$ and that for any $t>t^*$ there exists a $t' \in (t^*,t)$ such that $\tau(t')<0$.

Now, by assumption $\delta(t)$ is piecewise constant; hence, if $\delta$ is discontinuous at $t^*$, then there exists an interval $(t^*,t^*+h)\subset I_1$ over which $\delta$ is constant. If $\delta$ is continuous at $t^*$, then there also exists an interval $(t^*,t^*+h)\subset I_1$ over which $\delta$ is constant.

Suppose $\delta(t)=1$ on $(t^*,t^*+h)$. Then, for any $h'<h$, we have that $T_e(I(t^*+h');\epsilon_0)=T_e(I(t^*);\epsilon_0)$, since the prescription for $\epsilon_0$ is to fill when $\delta(t)=1$ on $I_1$. We also have that $T_{e,1}(I(t^*+h');\epsilon)-T_e(I(t^*+h');\epsilon) = T_{e,1}(I(t^*);\epsilon)-T_e(I(t^*);\epsilon) +T_{e,1}([t^*,t^*+h'];\epsilon)-T_e([t^*,t^*+h'];\epsilon)$. Since $\delta(t)=1$ on $(t^*,t^*+h')$, it follows that $T_{e,1}([t^*,t^*+h'];\epsilon)=T_e([t^*,t^*+h'];\epsilon)$ and hence that $T_{e,1}(I(t^*+h');\epsilon)-T_e(I(t^*+h');\epsilon)=T_{e,1}(I(t^*);\epsilon)-T_e(I(t^*);\epsilon)$. Therefore, $\tau(t^*+h')=\tau(t^*)$, so that $\tau(t)=0$ on $[t^*,t^*+h]$, contradicting the definition of $t^*$.

So suppose $\delta(t)=0$ on $(t^*,t^*+h)$. Then it should be clear that $T_{e,1}(I(t);\epsilon)$ is constant over $(t^*,t^*+h)$. If $n_{\epsilon_0,T}(t^*)>0$, then according to our prescription there exists an $h' \in (0,h)$ such that $\epsilon_e(t)=1$ with $n_{\epsilon_0,T}(t)>0$ over $(t^*,t^*+h')$. Therefore, given $h'' \in (0,h')$, we have $T_e(I(t^*+h'');\epsilon_0) =T_e(I(t^*);\epsilon_0)+h''$, while $T_e(I(t^*+h'');\epsilon)=T_e(I(t^*);\epsilon) +T_e([t^*,t^*+h''];\epsilon)$. The result is that $\tau(t^*+h'')=\tau(t^*)+h'' -T_e([t^*,t^*+h''];\epsilon)\geqslant\tau(t^*)=0$. Therefore, $\tau(t)\geqslant 0$ on $[t^*,t^*+h']$, again contradicting the definition of $t^*$.

So suppose that $\delta(t)=0$ on $(t^*,t^*+h)$ with $n_{\epsilon_0,T}(t^*)=0$. Then,

$$\begin{aligned} n_{\epsilon,T}(t^*) &= r_fT_f(I(t^*);\epsilon) - r_eT_e(I(t^*);\epsilon) \leqslant r_f[T_f(I(t^*);\epsilon_0) \\ &\quad - T_{e,1}(I(t^*);\epsilon)] - r_e[T_e(I(t^*);\epsilon_0) + T_{e,1}(I(t^*);\epsilon)] \\ &= n_{\epsilon_0,T}(t^*) - (r_f + r_e)T_{e,1}(I(t^*);\epsilon) \\ &= -(r_f + r_e)T_{e,1}(I(t^*);\epsilon), \end{aligned} \quad (8)$$

which is only possible if $n_{\epsilon,T}(t^*)=0$ with $T_{e,1}(I(t^*);\epsilon)=0$. But since $n_{\epsilon,T}(t^*)=n_{\epsilon_0,T}(t^*)=0$, then since $\delta(t)=0$ on $(t^*,t^*+h)$, it follows that $n_{\epsilon,T}(t)=n_{\epsilon_0,T}(t)=0$ on $(t^*,t^*+h)$, and hence $T_{e,1}([t^*,t^*+h'];\epsilon)=T_e([t^*,t^*+h'];\epsilon)=T_e([t^*,t^*+h'];\epsilon_0)=0$ on for all $h' \in [0,h]$, so that $\tau(t)=\tau(t^*)=0$ on $[t,t+h]$, which is again a contradiction.

Since we have exhausted all possibilities, we have established that $\tau(t_{\epsilon_0})<0$ leads to a contradiction. Therefore, $\tau(t_{\epsilon_0})\geqslant 0$ and the proof is complete.

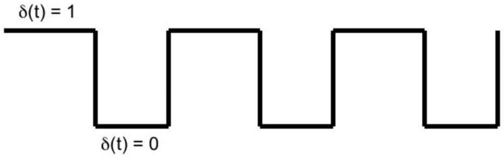

FIG. 3. Illustration of a periodic resource-availability profile.

## III. EXAMPLES OF OPTIMAL FILL-EMPTY PROFILES

For simplicity, we consider optimal fill-empty profiles generated by a $\delta(t)$ that is periodic over the time interval $[0,T]$. Specifically, we consider a basic profile, denoted $\sigma_{(T_1,T_2)}$, defined by

$$\sigma_{(T_1,T_2)}(t) = \begin{cases} 1 & \text{if } t \in [0,T_1], \\ 0 & \text{if } t \in (T_1,T_1+T_2]. \end{cases}$$

Over the time interval $[0,T]$, we then define $\delta(t)$ by setting $\delta(t)=\sigma_{(T_1,T_2)}(t)$ on $[0,T_1+T_2]$ and then imposing the periodicity relation $\delta(t)=\delta(t+T_1+T_2)$. We also assume that $T=N(T_1+T_2)$ for some positive integer $N$. An example of such a profile is illustrated in Fig. 3.

### A. Profile 1: Excess emptying time

If $r_eT_2\geqslant r_fT_1$, then it should be apparent that an optimal fill-empty profile, defined by our $\epsilon_0$ prescription, is to fill whenever $\delta(t)=1$ and to empty whenever $\delta(t)=0$ and $n_{\epsilon,T}(t)>0$, until time $T$. For this profile, then, we have that the optimal values of $T_f$, $T_e$, and $T_w$ are given by

$$T_f([0,T];\epsilon_0) = NT_1,$$

$$T_e([0,T];\epsilon_0) = \frac{r_f}{r_e}NT_1,$$

$$T_w([0,T];\epsilon_0) = N\frac{r_eT_2 - r_fT_1}{r_e}, \quad (9)$$

and so any other optimal fill-empty profile must fill exactly when $\delta(t)=1$. Note that $T_w$ only occurs in intervals where $\delta(t)=0$ (since $T_f([0,T];\epsilon_0)=NT_1$). Therefore, during these periods, since there is a lack of available resource (analogous to periods of night when making the analogy to sleep), the tank is either filling when there is nothing to fill it with (analogous to wakefulness periods during lack of available sensory information) or is emptying when the tank is already empty (analogous to excess sleeping). When $\delta(t)=0$, any combination of filling, or emptying when $n_T=0$, over these time intervals will not affect the final value of $n_P(T)$. Thus, there is a considerable degree of freedom in choosing an optimal fill-empty profile, which is consistent with the observed breakdown in sleeping patterns in environments with limited exposure to the sun [44].

### B. Profile 2: Lack of emptying time

Now consider the case when $r_eT_2<r_fT_1$. Then at the end of every resource availability cycle $[n(T_1+T_2),(n+1)(T_1$

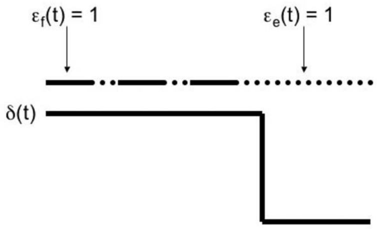

FIG. 4. Illustration of the "microsleep" solution profile.

$+T_2)]$, we have that $n_{\epsilon_0,T}$ increases by $r_f T_1 - r_e T_2$, as long as we are before the final emptying phase. So, if $n$ is chosen so that $n(r_f T_1 - r_e T_2) \leqslant r_e[T - n(T_1+T_2)]$ but $(n+1)(r_f T_1 - r_e T_2) > r_e[T-(n+1)(T_1+T_2)]$, then we have $t_{\epsilon_0} \in [n(T_1+T_2),(n+1)(T_1+T_2)]$, which further implies that $t_{\epsilon_0} \in [n(T_1+T_2), n(T_1+T_2)+T_1]$. Therefore, $n_{\epsilon,T}(t_{\epsilon_0}) = n(r_f T_1 - r_e T_2) + r_f[t_{\epsilon_0} - n(T_1+T_2)] = r_e(T - t_{\epsilon_0})$.

Solving for $t_{\epsilon_0}$, we obtain

$$t_{\epsilon_0} = nT_2 + \frac{r_e}{r_e + r_f} T, \tag{10}$$

from which it follows that

$$T_f([0,T];\epsilon_0) = \frac{r_e}{r_e + r_f} T,$$

$$T_e([0,T];\epsilon_0) = \frac{r_f}{r_e + r_f} T,$$

$$T_w([0,T];\epsilon_0) = 0. \tag{11}$$

An interesting fill-empty profile that arises from this condition is defined as follows: Over the interval $[0,T]$, $\epsilon_e(t) = 1$ whenever $\delta(t)=0$. Whenever $\delta(t)=1$, $\epsilon_f(t)$ and $\epsilon_e(t)$ alternate in being 1, with corresponding time lengths $t_f$ and $t_e$. To determine these time lengths, we first assume that there are $M$ such cycles over each $\delta(t)=1$ interval of length $T_1$, so that $t_f + t_e = T_1/M$. The net accumulation in the tank over each such interval should be $r_e T_2 = M(r_f t_f - r_e t_e)$, which may be solved to give

$$t_f = \frac{1}{M}\frac{r_e}{r_f + r_e}(T_1 + T_2),$$

$$t_e = \frac{1}{M}\frac{r_f T_1 - r_e T_2}{r_f + r_e}. \tag{12}$$

Note that $r_f t_f > r_e t_e$ and that $T_e([0,T];\epsilon) = r_f/(r_e+r_f)T$, so that this profile is indeed an optimal one.

This solution profile is illustrated in Fig. 4. Essentially, when the amount of time during which $\delta(t)=0$ is not sufficiently long to process all of the resource that can fill the tank, then one optimal solution profile is, during periods when $\delta(t)=1$, to fill and empty the tank in alternating time intervals of lengths $t_f$ and $t_e$, and then to empty the tank whenever $\delta(t)=0$. We argue that the $t_e$ empty periods are analogous to the phenomenon of "microsleep" that occurs during sleep deprivation.

In general, we claim that any optimal solution profile for this form of $\delta(t)$ will have $\epsilon_e(t)=1$ whenever $\delta(t)=0$. Otherwise, we obtain $r_f/(r_f+r_e)T = T_e([0,T];\epsilon) = T_{e,0}([0,T];\epsilon) + T_{e,1}([0,T];\epsilon) < NT_2 + T_{e,1}([0,T];\epsilon) \Rightarrow T_{e,1}([0,T];\epsilon) > N(r_f T_1 - r_e T_2)/(r_f + r_e)$.

But this implies that $T_f([0,T];\epsilon) \leqslant NT_1 - T_{e,1}([0,T];\epsilon) < NT_1 - N(r_f T_1 - r_e T_2)/(r_f+r_e) = r_e/(r_e+r_f)T \Rightarrow \Leftarrow$. With this contradiction, our claim is proved.

### C. Profile 3: Union of two resource availability profiles

As a final example for this subsection, we consider a resource availability profile given by

$$\delta(t) = \begin{cases} \sigma_{(T_1,T_2)}(t) & \text{if } t \in [0, T_1 + T_2], \\ \sigma_{(T_1,T_2')}(t - T_1 - T_2) & \text{if } t \in [T_1 + T_2, \\ & T = 2T_1 + T_2 + T_2'], \end{cases}$$

where $r_e T_2' < r_f T_1 < r_e T_2$. Then it is possible to show that

$$T_f([0,T];\epsilon_0) = T_1 + \frac{r_e}{r_e + r_f}(T_1 + T_2'),$$

$$T_e([0,T];\epsilon_0) = \frac{r_f}{r_e}T_1 + \frac{r_f}{r_e + r_f}(T_1 + T_2'),$$

$$T_w([0,T];\epsilon_0) = \frac{r_e T_2 - r_f T_1}{r_e}. \tag{13}$$

We now show that this uniquely determines $T_f([0,T_1+T_2];\epsilon_0)$, $T_f([T_1+T_2,T];\epsilon_0)$, $T_e([0,T_1+T_2];\epsilon_0)$, $T_e([T_1+T_2,T];\epsilon_0)$, $T_w([0,T_1+T_2];\epsilon_0)$, and $T_w([T_1+T_2,T];\epsilon_0)$.

Note that $T_f([0,T];\epsilon_0) \geqslant T_f([0,T_1+T_2];\epsilon_0) + \frac{r_e}{r_e+r_f}(T_1+T_2') \Rightarrow T_f([T_1+T_2,T];\epsilon_0) \geqslant \frac{r_e}{r_e+r_f}(T_1+T_2')$. Therefore, $T_e([T_1+T_2,T];\epsilon_0) \leqslant \frac{r_f}{r_e+r_f}(T_1+T_2')$. However, $n_{\epsilon_0,T}(T) = 0 \geqslant r_f T_f([T_1+T_2,T];\epsilon_0) - r_e T_e([T_1+T_2,T];\epsilon_0) \geqslant \frac{r_f r_e}{r_e+r_f}(T_1+T_2') - r_e T_e([T_1+T_2,T];\epsilon_0) \Rightarrow T_e([T_1+T_2,T];\epsilon_0) \geqslant \frac{r_f}{r_e+r_f}(T_1+T_2')$, and so $T_e([T_1+T_2,T];\epsilon_0) = \frac{r_f}{r_e+r_f}(T_1+T_2')$. But this implies that $T_f([T_1+T_2,T];\epsilon_0) \leqslant \frac{r_e}{r_e+r_f}(T_1+T_2')$, so $T_f([T_1+T_2,T];\epsilon_0) = \frac{r_e}{r_e+r_f}(T_1+T_2')$. This further implies that $T_w([T_1+T_2,T];\epsilon_0) = 0$, so that $T_w([0,T_1+T_2];\epsilon_0) = \frac{r_e T_2 - r_f T_1}{r_e}$. Finally, we obtain $T_f([0,T_1+T_2];\epsilon_0) = T_1$ and $T_e([0,T_1+T_2];\epsilon_0) = \frac{r_f}{r_e}T_1$.

Summarizing the results, we have

$$T_f([0,T_1+T_2];\epsilon_0) = T_1,$$

$$T_f([T_1+T_2,T];\epsilon_0) = \frac{r_e}{r_e + r_f}(T_1 + T_2') < T_1,$$

$$T_e([0,T_1+T_2];\epsilon_0) = \frac{r_f}{r_e}T_1,$$

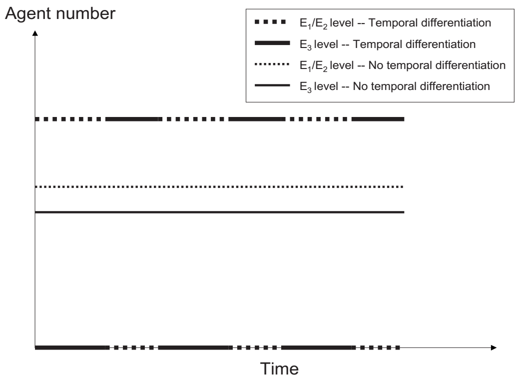

FIG. 5. Illustration of the time dependence of various enzyme levels in both the temporally differentiated and undifferentiated labor strategies.

$$T_e([T_1+T_2,T];\epsilon_0)=\frac{r_f}{r_f+r_e}(T_1+T_2')>T_2',$$

$$T_w([0,T_1+T_2];\epsilon_0)=\frac{r_eT_2-r_fT_1}{r_e},$$

$$T_w([T_1+T_2,T];\epsilon_0)=0. \tag{14}$$

Therefore, maximal processing of the external resource still requires emptying the tank when $\delta(t)=1$ on the interval $[T_1+T_2,T]$. In making the analogy with sleep, this implies that excess sleep in a certain time interval will not prevent sleep deprivation in a later time interval when the period of resource availability exceeds the period when the resource may be processed. Intuitively, this makes sense, since once the brain has performed all of the information processing and upkeep tasks associated with sleep, further sleep will not prevent the accumulation of sleep-related brain tasks during a later cycle.

## IV. AGENT-BASED MODEL

In this section, we introduce our agent-based, three-process model, whereby an external resource is converted into a final product in a series of three steps. We set up dynamical equations governing the production of the final product for both the temporally differentiated and nondifferentiated cases (see Fig. 5).

### A. Definition of the model

Our model consists of some compartment of fixed volume $V$, into which flows a resource, denoted $R$, at some fixed rate $f_R$. This resource is processed into a final product, denoted $P$, via a series of three agent-mediated, or in the language of chemical kinetics, enzyme-catalyzed, steps.

In the first step, an enzyme, denoted $E_1$, binds to $R$ and then converts $R$ into an intermediate $I_1$. In the second step, an enzyme, denoted $E_2$, binds to $I_1$ and then converts $I_1$ into an intermediate $I_2$. In the third and final step, an enzyme, denoted $E_3$, binds to $I_2$ and then converts $I_2$ into the final product $P$.

The resource $R$ and intermediates $I_1$ and $I_2$ have finite lifetimes in the compartment, defined by first-order decay constants $k_{D,1}$, $k_{D,2}$, and $k_{D,3}$, respectively. These decay terms can be due to various factors, such as diffusion out of the compartment, or simply the physical decay of the components themselves. In the context of networked systems and data processing, these decay constants can also correspond to a finite lifetime during which an information packet is relevant (say stock information that is used by an investor to decide whether or not to invest in a given stock).

For greater generality, we also assume that the resource $R$ and intermediates $I_1$ and $I_2$ can decay when bound to their respective enzymes. The decay constants are given by $k'_{D,1}$, $k'_{D,2}$, and $k'_{D,3}$, respectively. We assume that $0\leq k'_{D,i}\leq k_{D,i}$, for $i\in\{1,2,3\}$.

In the language of chemical kinetics, the set of reactions in the compartment is given by

$$R\rightarrow \text{decay products (first-order rate constant } k_{D,1}),$$

$$E_1+R\rightarrow E_1R \text{ (second-order rate constant } k_{11}),$$

$$E_1R\rightarrow E_1+\text{decay products (first-order rate constant } k'_{D,1}),$$

$$E_1R\rightarrow E_1+I_1 \text{ (first-order rate constant } k_{12}),$$

$$I_1\rightarrow \text{decay products (first-order rate constant } k_{D,2}),$$

$$E_2+I_1\rightarrow E_2I_1 \text{ (second-order rate constant } k_{21}),$$

$$E_2I_1\rightarrow E_2+\text{decay products (first-order rate constant } k'_{D,2}),$$

$$E_2I_1\rightarrow E_2+I_2 \text{ (first-order rate constant } k_{22}),$$

$$I_2\rightarrow \text{decay products (first-order rate constant } k_{D,3}),$$

$$E_3+I_2\rightarrow E_3I_2 \text{ (second-order rate constant } k_{31}),$$

$$E_3I_2\rightarrow E_3+\text{decay products (first-order rate constant } k'_{D,3}),$$

$$E_3I_2\rightarrow E_3+P \text{ (first-order rate constant } k_{32}). \tag{15}$$

We now define the following quantities: We define $r$, $i_1$, and $i_2$ to be the number of particles of $R$, $I_1$, and $I_2$, respectively. We define $e_{b,i}$, $e_{f,i}$, and $e_i$ to be the number of bound particles of enzyme $E_i$, the number of free particles of enzyme $E_i$, and the total number of particles of enzyme $E_i$, where $i\in 1,2,3$. Note that $e_i=e_{b,i}+e_{f,i}$.

With these definitions, the system of ordinary differential equations governing the dynamics inside the compartment is given by

$$\frac{dr}{dt}=f_R-\left(\frac{k_{11}}{V}\right)(e_1-e_{b,1})r-k_{D,1}r,$$

$$\frac{de_{b,1}}{dt}=\left(\frac{k_{11}}{V}\right)(e_1-e_{b,1})r-(k_{12}+k'_{D,1})e_{b,1},$$

$$\frac{di_1}{dt} = k_{12}e_{b,1} - \left(\frac{k_{21}}{V}\right)(e_2 - e_{b,2})i_1 - k_{D,2}i_1,$$

$$\frac{de_{b,2}}{dt} = \left(\frac{k_{21}}{V}\right)(e_2 - e_{b,2})i_1 - (k_{22} + k'_{D,2})e_{b,2},$$

$$\frac{di_2}{dt} = k_{22}e_{b,2} - \left(\frac{k_{31}}{V}\right)(e_3 - e_{b,3})i_2 - k_{D,3}i_2,$$

$$\frac{de_{b,3}}{dt} = \left(\frac{k_{31}}{V}\right)(e_3 - e_{b,3})i_2 - (k_{32} + k'_{D,3})e_{b,3}. \qquad (16)$$

### B. Temporal differentiation in the three-enzyme compartment model

Temporal differentiation can occur in our model if we do not assume that $e_1$, $e_2$, and $e_3$ are fixed, but rather can oscillate in time. For the purposes of this paper, we will assume that $e_1$, $e_2$, and $e_3$ oscillate in such a way that $e_1 = e_1^+$, $e_2 = e_2^+$, and $e_3 = e_3^-$ over some time period of length $T_1$, followed by a time period of length $T_2$ where $e_1 = e_1^- \leq e_1^+$, $e_2 = e_2^- \leq e_2^+$, and $e_3 = e_3^+ \geq e_3^-$. We assume that the total number of enzymes remains fixed, however, so that $e_1^+ + e_2^+ + e_3^- = e_1^- + e_2^- + e_3^+$.

Essentially, if we switch from an enzyme-based viewpoint to an agent-based viewpoint, our model assumes that agents can switch from one set of tasks to another. In this model, the agents alternate between focusing on the first two processes and the third process. In chemical kinetics notation, we have

$$E_{1/2} \leftrightarrow E_3. \qquad (17)$$

### C. Limiting forms of the model

We will now study how our model behaves when $e_1 + e_2 + e_3$ and $k_{D,3}$ may each be regarded as small in some sense. The criterion for smallness will be defined later, once we have established the behavior of the model in these regimes.

To begin, we note from the previous subsection that the various enzyme numbers fluctuate in time. More precisely, there exist $T_1, T_2 > 0$ such that for every integer $s$, the total enzyme numbers for each enzyme are at $e_1^+$, $e_2^+$, and $e_3^-$, respectively, during the time interval $[s(T_1+T_2), s(T_1+T_2)+T_1]$, while during the time interval $[s(T_1+T_2)+T_1, (s+1)(T_1+T_2)]$, the enzyme numbers are at $e_1^-$, $e_2^-$, and $e_3^+$.

If we define, for $i \in \{1,2,3\}$,

$$\bar{e}_i = \frac{e_i^+ + e_i^-}{2} \qquad (18)$$

and

$$\lambda_i = \frac{e_i^+ - e_i^-}{e_i^+ + e_i^-}, \qquad (19)$$

then it may be readily shown that

$$e_i^{\pm} = \bar{e}_i(1 \pm \lambda_i). \qquad (20)$$

For fixed values of $\lambda_1$ and $\lambda_2$, we wish to develop a form for the first four equations assuming that $\bar{e}_1$ and $\bar{e}_2$ are small. The overall strategy is as follows: Because $I_2$ is the intermediate that feeds into the third reaction, our goal is to determine the rate of production of $I_2$ when $\bar{e}_1$ and $\bar{e}_2$ are small. By this we mean that we seek to determine, with respect to $\bar{e}_1$ and $\bar{e}_2$, the lowest-order term contributing to the production rate of $I_2$.

If we let $r_0$ denote $r$ when $\bar{e}_1 = 0$, we obtain

$$\frac{dr_0}{dt} = f_R - k_{D,1}r_0. \qquad (21)$$

If we define $e_{b,1,1} = (\partial e_{b,1}/\partial \bar{e}_1)_{\bar{e}_1=0}$, we obtain

$$\frac{de_{b,1,1}}{dt} = \left(\frac{k_{11}}{V}\right)(1 \pm \lambda_1)r_0 - \left[\left(\frac{k_{11}}{V}\right)r_0 + k_{12} + k'_{D,1}\right]e_{b,1,1}. \qquad (22)$$

If we define $i_{1,1} = (\partial i_1/\partial \bar{e}_1)_{\bar{e}_1=\bar{e}_2=0}$, we obtain

$$\frac{di_{1,1}}{dt} = k_{12}e_{b,1,1} - k_{D,2}i_{1,1}. \qquad (23)$$

Note that $e_{b,2} = 0$ when either $\bar{e}_1 = 0$ or $\bar{e}_2 = 0$. Therefore, the lowest-order derivative of $e_{b,2}$ that is possibly nonvanishing at $(\bar{e}_1, \bar{e}_2) = (0,0)$ is $\partial^2 e_{b,2}/(\partial \bar{e}_1 \partial \bar{e}_2)$. Defining $e_{b,2,1} = [\partial^2 e_{b,2}/(\partial \bar{e}_1 \partial \bar{e}_2)]_{(\bar{e}_1,\bar{e}_2)=(0,0)}$, we obtain

$$\frac{de_{b,2,1}}{dt} = \left(\frac{k_{21}}{V}\right)(1 \pm \lambda_2)i_{1,1} - (k_{22} + k'_{D,2})e_{b,2,1}. \qquad (24)$$

For the final two equations, we assume that $k_{D,3} = k'_{D,3} = 0$, giving

$$\frac{di_2}{dt} = k_{22}e_{b,2} - \left(\frac{k_{31}}{V}\right)(e_3 - e_{b,3})i_2,$$

$$\frac{de_{b,3}}{dt} = \left(\frac{k_{31}}{V}\right)(e_3 - e_{b,3})i_2 - k_{32}e_{b,3}. \qquad (25)$$

### D. Long-term behavior of the model

In the absence of temporal differentiation ($\lambda_1 = \lambda_2 = \lambda_3 = 0$), the values of $e_1$, $e_2$, and $e_3$ remain constant, and so we expect the dynamics to converge to a steady state. With temporal differentiation, the values of $e_1$, $e_2$, and $e_3$ oscillate between two sets of values over a well-defined time period $T_1 + T_2$. In this situation, we cannot expect the dynamics to settle into a steady-state solution, but rather into a periodic solution.

In this subsection, we will compare the long-term behavior of the model in the absence of temporal differentiation with a temporally differentiated labor strategy where $\lambda_1 = \lambda_2 = \lambda_3 = 1$. Here, the enzymes devote themselves entirely to the first two tasks over a time interval of length $T_1$ and then devote themselves entirely to the third task over a time interval of length $T_2$.

### 1. No temporal differentiation ($\lambda_1=\lambda_2=\lambda_3=0$)

When $\lambda_1=\lambda_2=\lambda_3=0$, the long-term solution is simply a steady-state solution. Setting the left-hand sides of the four linearized equations to 0 gives

$$r_0=\frac{f_R}{k_{D,1}},$$

$$e_{b,1,1}=\frac{\left(\frac{k_{11}}{V}\right)\left(\frac{f_R}{k_{D,1}}\right)}{k_{12}+k'_{D,1}+\left(\frac{k_{11}}{V}\right)\left(\frac{f_R}{k_{D,1}}\right)},$$

$$i_{1,1}=\frac{k_{12}}{k_{D,2}}\frac{\left(\frac{k_{11}}{V}\right)\left(\frac{f_R}{k_{D,1}}\right)}{k_{12}+k'_{D,1}+\left(\frac{k_{11}}{V}\right)\left(\frac{f_R}{k_{D,1}}\right)},$$

$$e_{b,2,1}=\frac{(k_{21}/V)}{k_{22}+k'_{D,2}}\frac{k_{12}}{k_{D,2}}\frac{\left(\frac{k_{11}}{V}\right)\left(\frac{f_R}{k_{D,1}}\right)}{k_{12}+k'_{D,1}+\left(\frac{k_{11}}{V}\right)\left(\frac{f_R}{k_{D,1}}\right)}. \tag{26}$$

If we define $A$ via

$$A=k_{22}\frac{(k_{21}/V)}{k_{22}+k'_{D,2}}\frac{k_{12}}{k_{D,2}}\frac{\left(\frac{k_{11}}{V}\right)\left(\frac{f_R}{k_{D,1}}\right)}{k_{12}+k'_{D,1}+\left(\frac{k_{11}}{V}\right)\left(\frac{f_R}{k_{D,1}}\right)}, \tag{27}$$

then we obtain that the rate of production of $I_2$ is $A\bar{e}_1\bar{e}_2$.

Because the production rate of $P$ is given by $k_{32}e_{b,3}$ and because $e_{b,3}\leqslant e_3$, the production rate of $P$ is bounded from above by $k_{32}e_3$. Therefore, if we can find a steady-state solution where the production rate of $P$ is given by $k_{32}e_3$, then we will have found a solution that maximizes system output.

In order for the production rate of $P$ to equal $e_3$, we must have that $e_{b,3}=e_3$; that is, all of the $E_3$ molecules must be bound to an $I_2$ molecule. If we solve for the steady-state value of $e_{b,3}$, we obtain that

$$e_{b,3,ss}=\frac{\left(\frac{k_{31}}{V}\right)i_{2,ss}}{k_{32}+\left(\frac{k_{31}}{V}\right)i_{2,ss}}e_3, \tag{28}$$

where the subscript *ss* signifies "steady state."

Note then that $e_{b,3,ss}$ is 0 when $i_{2,ss}=0$ and increases to $e_3$ as $i_{2,ss}\to\infty$. Therefore, the maximal production rate of $P$ is obtained in the limiting case of $i_{2,ss}\to\infty$. While this condition is impossible to reach in practice, we can always choose a steady state where $i_{2,ss}$ can be made as large as possible, so that the production rate of $P$ can be made as close to the maximal value as desired. The reason for this is that $k_{D,3}=k'_{D,3}=0$. As a result, the only way that $I_2$ can be lost is via its conversion to $P$. If we set $i_2\to\infty$ and choose parameters so that the rate of production of $I_2$ is equal to the rate of consumption, then we will have a steady state where the production rate of $P$ is maximal.

When $i_2=\infty$, the rate of production of $P$ is given by $k_{32}\bar{e}_3$. Therefore, at steady state,

$$A\bar{e}_1\bar{e}_2=k_{32}\bar{e}_3. \tag{29}$$

Now, defining $e_{tot}=\bar{e}_1+\bar{e}_2+\bar{e}_3$, $\alpha=(\bar{e}_1+\bar{e}_2)/e_{tot}$, and $\beta=\bar{e}_1/(\bar{e}_1+\bar{e}_2)$, we have that the steady-state production rate of $P$ is

$$A\alpha^2\beta(1-\beta)e_{tot}^2=k_{32}(1-\alpha)e_{tot}. \tag{30}$$

The steady-state production rate of $P$ is maximized when $\beta=1/2$, and so we wish to solve

$$\alpha^2+\frac{1}{\gamma e_{tot}}\alpha-\frac{1}{\gamma e_{tot}}=0, \tag{31}$$

where $\gamma\equiv A/(4k_{32})$.

This gives

$$1-\alpha=\frac{1+2\gamma e_{tot}-\sqrt{1+4\gamma e_{tot}}}{2\gamma e_{tot}}. \tag{32}$$

Since this value of $\alpha$ is the one obtained for the temporally undifferentiated system, we denote the $\alpha$ derived here as $\alpha_{undiff}$.

### 2. Temporal differentiation with $\lambda_1=\lambda_2=\lambda_3=1$

When $\lambda_1=\lambda_2=\lambda_3=1$, the enzyme levels for enyzmes $E_1$ and $E_2$ are at their maximal levels during the time interval $[s(T_1+T_2),s(T_1+T_2)+T_1]$ and are not present during the time interval $[s(T_1+T_2)+T_1,(s+1)(T_1+T_2)]$. During the labor cycles where the enzymes are focused on the third reaction, $e_1=e_{b,1}=e_2=e_{b,2}=0$, so that when the enzymes switch to the first two reactions in the following labor cycle, $e_{b,2}$ starts out at 0 and rises to a steady-state value.

Therefore, in order to maximize the average production rate of $I_2$, we want to choose a large value of $T_1$, so that the time it takes for $e_{b,2}$ to rise to its steady-state value is small compared to $T_1$. Mathematically, this means that we want to take $T_1\to\infty$, so that the average production rate of $I_2$ is simply given by the steady-state production rate.

When $e_1=e_1^+=2\bar{e}_1$ and $e_2=e_2^+=2\bar{e}_2$, the steady-state production rate of $P$ is $4A\bar{e}_1\bar{e}_2$, and so, when $T_1$ is large, the total amount of $I_2$ that is produced during the time period $T_1$ is given by

$$\Delta i_2=4A\bar{e}_1\bar{e}_2T_1. \tag{33}$$

To ensure periodicity of the solution, $T_2$ must be such that the amount of $I_2$ consumed in the third step is equal to the amount of $I_2$ accumulated. Assuming that the amount of $I_2$ is large at all times (as with the undifferentiated case, this assumption maximizes the overall production rate of $P$), the amount of $I_2$ consumed is given by $2k_{32}\bar{e}_3T_2$. We then have

$$A(2\bar{e}_1)(2\bar{e}_2)T_1=k_{32}(2\bar{e}_3)T_2. \tag{34}$$

Now, since all the enzymes are focused on either the first two tasks or the third task in the temporally differentiated model, it follows that $e_{tot}=2\bar{e}_1+2\bar{e}_2=2\bar{e}_3$. Defining $\beta=2\bar{e}_1/e_{tot}$ gives

$$A\beta(1-\beta)e_{tot}^2T_1=k_{32}e_{tot}T_2. \tag{35}$$

Note that the average production rate of $P$ is simply given by $k_{32}e_{tot}T_2/(T_1+T_2)$. We can determine the optimal split between the two work cycles that maximizes the average production rate of $P$.

Defining $\alpha=T_1/(T_1+T_2)$, we get that the average production rate of $P$ is simply $k_{32}(1-\alpha)e_{tot}$. Dividing both sides of the previous equation by $T_1+T_2$ gives

$$A\beta(1-\beta)\alpha e_{tot}^2=k_{32}(1-\alpha)e_{tot}. \tag{36}$$

As with the steady-state solution, we can maximize the output of $P$ when $\beta=1/2$, and so, defining $\gamma$ as before, we obtain

$$1-\alpha=\frac{\gamma e_{tot}}{1+\gamma e_{tot}}. \tag{37}$$

To distinguish this $\alpha$ from the $\alpha$ defined in the temporally undifferentiated case, we redenote the $\alpha$ defined in this subsection by $\alpha_{diff}$.

Although a given realization of the system must have finite values of $i_2$, $T_1$, and $T_2$, the production rate of $P$ is maximized as $i_2$, $T_1$, and $T_2$ take on arbitrarily large values, where the $T_1$ to $T_2$ ratio is held fixed. This is the basis for considering steady states and periodic solutions where $i_2$, $T_1$, and $T_2$ are taken to be infinite. We can consider such limiting forms of the model because $k_{D,3}=k'_{D,3}=0$, so that we can allow the cycle times and the value of $i_2$ to take on arbitrarily large values without worrying about loss of intermediate that can reduce the overall process efficiency.

### E. Comparison of system output with and without temporal differentiation

It may be shown that $1-\alpha_{diff}>1-\alpha_{undiff}$ for $\gamma e_{tot}>0$, so that temporal differentiation leads to a rate of production of $P$ that is faster than the undifferentiated case (see Fig. 6). It may also be shown that, as $\gamma e_{tot}$ increases from 0 to $\infty$, the ratio $(1-\alpha_{diff})/(1-\alpha_{undiff})$ increases from 1 to a maximum of $4/3$ at $\gamma e_{tot}=2$ and then decreases to 1. This implies that the temporally differentiated labor strategy can achieve a 33% greater production rate of $P$ than the undifferentiated strategy.

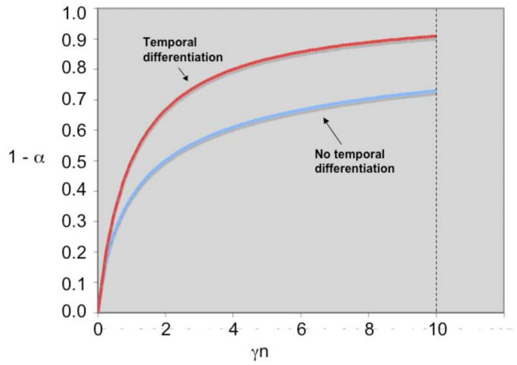

FIG. 6. (Color online) Comparison of the production rate of product $P$, as measured by the factor $1-\alpha$, for the temporally differentiated and undifferentiated pathways.

### F. Small-$\bar{e}_1,\bar{e}_2$ criterion

The analytical solution of the first four differential equations governing our model explicitly made use of the assumption that $\bar{e}_1$ and $\bar{e}_2$ are small. We therefore need to investigate what the criteria for smallness are.

When $\bar{e}_1$ and $\bar{e}_2$ are small, the rate of the first reaction is given by

$$k_{12}e_{b,1}=k_{12}\frac{\left(\frac{k_{11}}{V}\right)\left(\frac{f_R}{k_{D,1}}\right)}{k_{12}+k'_{D,1}+\left(\frac{k_{11}}{V}\right)\left(\frac{f_R}{k_{D,1}}\right)}\bar{e}_1. \tag{38}$$

As $\bar{e}_1$ increases, there will eventually be enough enzyme $E_1$ present to process all of the incoming resource. At this point, the reaction rate becomes $f_R$. Therefore, the criterion that $\bar{e}_1$ be small is

$$k_{12}e_{b,1}\ll f_R\Rightarrow\bar{e}_1\ll\frac{f_R}{k_{12}}+\frac{k_{D,1}}{(k_{11}/V)}\left(1+\frac{k'_{D,1}}{k_{12}}\right). \tag{39}$$

Now, in the small-$\bar{e}_1,\bar{e}_2$ regime, the rate of the second reaction is given by

$$k_{22}e_{b,2}=k_{22}\frac{(k_{21}/V)}{k_{22}+k'_{D,2}}\frac{k_{12}e_{b,1}}{k_{D,2}}\bar{e}_2. \tag{40}$$

To determine the small-$\bar{e}_1,\bar{e}_2$ criterion for the second reaction, note that at steady state, the rate of production of $I_2$ for small $\bar{e}_2$ is given by the small-$\bar{e}_1$ formula for the rate of production of $I_1$, except with the replacements $f_R\rightarrow k_{12}e_{b,1}$, $k_{12}\rightarrow k_{22}$, $k_{D,1}\rightarrow k_{D,2}$, $k'_{D,1}\rightarrow k'_{D,2}$, and $k_{11}\rightarrow k_{21}$. This gives

$$k_{22}e_{b,2}=k_{22}\frac{\left(\frac{k_{21}}{V}\right)\left(\frac{k_{12}e_{b,1}}{k_{D,2}}\right)}{k_{22}+k'_{D,2}+\left(\frac{k_{21}}{V}\right)\left(\frac{k_{12}e_{b,1}}{k_{D,2}}\right)}\bar{e}_2. \tag{41}$$

As with the criterion for small $\bar{e}_1$, this formula only holds when

$$\bar{e}_2\ll\frac{k_{12}e_{b,1}}{k_{22}}+\frac{k_{D,2}}{(k_{21}/V)}\left(1+\frac{k'_{D,2}}{k_{22}}\right). \tag{42}$$

However, once this inequality holds, the only way for the rate of the second reaction to reduce to the formula in Eq. (27) in the limit of small $\bar{e}_2$ is if

$$k_{22} + k'_{D,2} \gg \left(\frac{k_{21}}{V}\right)\left(\frac{k_{12}e_{b,1}}{k_{D,2}}\right), \tag{43}$$

which implies that

$$\bar{e}_1 \ll (k_{22} + k'_{D,2})\frac{k_{D,2}}{(k_{21}/V)}\left[\frac{1}{k_{12}} + \frac{k_{D,1}}{(k_{11}/V)}\left(1 + \frac{k'_{D,1}}{k_{12}}\right)\frac{1}{f_R}\right]. \tag{44}$$

It also implies that the above inequality for $\bar{e}_2$ reduces to

$$\bar{e}_2 \ll \frac{k_{D,2}}{(k_{21}/V)}\left(1 + \frac{k'_{D,2}}{k_{22}}\right). \tag{45}$$

Setting $\bar{e}_1 = \bar{e}_2$ in order to maximize the reaction rate and combining all the inequalities together, we obtain

$$\bar{e}_1, \bar{e}_2 \ll \frac{f_R}{k_{12}} + \frac{k_{D,1}}{(k_{11}/V)}\left(1 + \frac{k'_{D,1}}{k_{12}}\right),$$

$$\frac{k_{D,2}}{(k_{21}/V)}\left(1 + \frac{k'_{D,2}}{k_{22}}\right),$$

$$(k_{22} + k'_{D,2})\frac{k_{D,2}}{(k_{21}/V)}\left[\frac{1}{k_{12}} + \frac{k_{D,1}}{(k_{11}/V)}\left(1 + \frac{k'_{D,1}}{k_{12}}\right)\frac{1}{f_R}\right]. \tag{46}$$

If we take $f_R \to \infty$, then the first inequality for small $\bar{e}_1, \bar{e}_2$ is automatically satisfied. If we also assume that $k_{12} = k_{22}$, then the second and third inequalities are equivalent.

Now, in the limit $f_R \to \infty$, we have

$$\gamma e_{tot} = \frac{k_{12}(k_{21}/V)}{4k_{32}k_{D,2}\left(1 + \dfrac{k'_{D,2}}{k_{22}}\right)}e_{tot}, \tag{47}$$

so that, if we want $\gamma e_{tot}$ to equal some specified value, denoted $\eta$, we obtain

$$e_{tot} = 4\eta\frac{k_{32}}{k_{12}}\frac{k_{D,2}}{(k_{21}/V)}\left(1 + \frac{k'_{D,2}}{k_{22}}\right). \tag{48}$$

Combining this with the requirement that $\bar{e}_1, \bar{e}_2 \ll [k_{D,2}/(k_{21}/V)](1 + k'_{D,2}/k_{22})$, we obtain

$$\frac{k_{32}}{k_{12}} \ll \frac{1}{2\eta}. \tag{49}$$

Therefore, by choosing values for $k_{12} = k_{22}$ and $k_{32}$ so that $k_{32}/k_{12}$ is sufficiently small, we can guarantee that we will be in a parameter regime where the small-$\bar{e}_1, \bar{e}_2$ criterion applies, while at the same time achieving a prespecified value for $\gamma e_{tot}$. As the value of $\gamma e_{tot}$ increases, the ratio $k_{32}/k_{12}$ must be made correspondingly smaller so that the small-$e_{tot}$ criterion is valid at the given value of $\eta = \gamma e_{tot}$.

### G. Numerical example

We chose to test the plausibility of our theory with a simple numerical example. We considered a system with the following parameter values: $k_{12} = k_{22} = k_{D,2} = k'_{D,2} = 1000$, $k_{21}/V = 1$, $k_{32} = 6.25$, and $e_{tot} = 100$. With the assumption that $f_R \to \infty$, we obtain $\gamma e_{tot} = 2$. These parameters were chosen so that our system would be in a regime where the small-$\bar{e}_1, \bar{e}_2$ approximation holds.

It should be noted that we have not specified the values of $f_R$, $k_{11}/V$, $k_{D,1}$, $k'_{D,1}$, $k_{31}/V$, $k_{D,3}$, or $k'_{D,3}$. The reason for this is that we are assuming that $f_R \to \infty$ and that $i_2 \to \infty$ with $k_{D,3} = k'_{D,3} = 0$. Therefore, the steady-state value of $r \to \infty$, so that the rate of the first reaction is simply $k_{12}e_1$, since $e_{b,1} = e_1$ under these conditions. Similarly, the rate of the third reaction is simply $k_{32}e_3$.

For the temporally undifferentiated case, we numerically solve for the steady-state production rate of $I_2$ and compare it to the steady-state production rate of $P$. The steady state for the whole process is found by adjusting $\alpha$ until the two production rates are equal. We then find the value of $\beta$ which maximizes the overall steady-state production rate of $P$.

For the temporally differentiated case, we consider very long cycle times, so that we may assume steady-state dynamics in each phase of the temporal labor cycle. We then find the value of $\alpha$ such that the average production rate of $I_2$ over a full cycle is equal to the average production rate of $P$ over a full cycle. This corresponds to the steady state for the whole process. We then find the value of $\beta$ which maximizes the overall steady-state production rate of $P$.

While this method might seem somewhat cumbersome and limited to only solving certain special cases of the dynamics, we used it because we found that direct numerical integration of the ordinary differential equations was difficult for parameter values corresponding to the small-$\bar{e}_1, \bar{e}_2$ regime. In this regime, we obtained a stiff system of differential equations that required small time steps, on the one hand, and long integration times on the other. Since the numerical scheme we used provides the steady-state or periodic solution to arbitrarily desired accuracy (under the assumptions described above), we decided to numerically solve our system in this much simpler manner.

For the undifferentiated case, we found a maximum production rate of 309, under the conditions $\alpha = 0.506$ and $\beta = 0.50$. For comparison, theory predicts that $\alpha_{undiff} = 0.500$ and $\beta = 0.50$.

For the differentiated case, we found a maximum production rate of 407, under the conditions $\alpha = 0.351$ and $\beta = 0.53$. For comparison, theory predicts that $\alpha_{diff} = 0.333$ and $\beta = 0.50$.

Therefore, for the parameters chosen, we found that a temporally differentiated labor strategy led to a 32% increase in the total production rate of $P$, in good agreement with the theoretical prediction of 33% for the parameter values chosen.

## V. DISCUSSION

### A. When can a temporally differentiated process outperform a nondifferentiated process?

In the tank-filling model, we found that a maximal amount of resource is processed if the tank fills when resource is available and empties when resource is not available. Although this result is obvious, it leads to fill-empty profiles that parallel various aspects of the sleep-wake cycle in diurnal organisms.

For the agent-based model, which assumes a continuous flow of resource, we have that if $k_{D,3} > 0$, then the temporally differentiated pathway will only perform optimally with a finite cycle time. For if the time period during which the agents focus on the first two subtasks is infinite, then the accumulated $I_2$ will decay away, resulting in a low production rate of $P$. As $k_{D,3}$ increases, the optimal cycle time will decrease.

This analysis suggests that when $k_{D,3}$ becomes sufficiently large, the advantage for a temporally differentiated labor strategy will disappear entirely. The reason for this is that when the agents switch tasks, there is a delay time during which the system reaches its new steady state, where the overall process rate is maximized. When $k_{D,3}$ is small, this delay time is a small fraction of the overall cycle time, so the effect on the overall process rate is negligible. However, when $k_{D,3}$ is large, this delay time is a significant fraction of the overall cycle time, and so the reduction in overall process rate can be sufficient to eliminate the advantage for the temporally differentiated strategy.

Therefore, as a general rule, temporal differentiation can only outperform a nontemporally differentiated labor strategy if there are process intermediates that have large characteristic decay times. This affords the agents the opportunity to focus on one set of subtasks at a time, without significant loss of the corresponding process intermediates.

### B. Possible implications for sleep and circadian rhythms

The implications of the results presented here are that sleep emerges because the brain can process more tasks if it adopts a temporally differentiated labor strategy. In this case, the individual agents may be taken to correspond to individual neurons, each of which is involved in performing various tasks associated with proper brain function (within each neuron, the individual agents may be taken to be the various biochemicals involved with proper neuronal function). If the neurons collectively engage in certain brain subtasks together, they can possibly perform each subtask more efficiently than if they engage in various subtasks simultaneously. Presumably, the more tasks a brain can accomplish within a given amount of time, the greater the survival advantage for the organism, providing an evolutionary selection pressure for temporally differentiated labor strategies.

We have seen that an optimal ratio between the times devoted to two different sets of process subtasks emerges even with a constant inflow of external resource. In the context of sleep, this suggests a natural sleep cycle that can exist independently of any external day-night regulation. This also suggests an evolutionary basis for sleep that could apply to nocturnal organisms [36]. This being said, the presence of a day-night cycle could nevertheless regulate the exact location of the various subtask time intervals. By analogy with the tank-filling model, it presumably makes sense for most organisms to remain alert during the day, when external information is most available, and to process that information at night, when external information is less available. The ability to avoid predators and to hunt stealthily are probably the major selection pressures driving the emergence of nocturnal organisms.

We also argue that the agent-based model presented in this paper suggests an evolutionary basis for the emergence of distinct REM and non-REM sleep states from an earlier undifferentiated sleep state. As brain complexity increases and the amount of information that must be processed during the sleep state increases, it becomes more efficient for the brain to oscillate between various information processing and consolidation subtasks associated with the sleep state itself [35].

Finally, the models presented in this paper suggest that temporal differentiation may not only provide an evolutionary basis for the emergence of sleep, but for circadian rhythms in general. That is, the various tasks associated with proper organismal function may be performed most efficiently if different tasks are performed at different times and in a well-defined manner. By having a fixed daily routine for performing the various “housekeeping chores” necessary for proper organismal function, the organism may be able to optimally carry out the various tasks necessary for organismal survival.

## VI. CONCLUSIONS AND FUTURE RESEARCH

This paper presented two simplified models showing that a temporally differentiated labor strategy can lead to increased system output. The first model considered the emptying and filling of a tank in the presence of a time-varying resource-availability profile. The second model considered a three-step process for the conversion of some external resource into a final product. In both models we showed that oscillating between the various process subtasks could lead to a greater system output than if all subtasks were performed simultaneously.

The two models presented in this paper suggest that a possible evolutionary basis for phenomena such as sleep and circadian rhythms is that it is optimal for an organism to oscillate between performing various tasks necessary for organismal survival. A temporally differentiated labor strategy for performing these various tasks requires less energy and time for the same amount of system output than an undifferentiated strategy. As is the case for task completion by human agents, it may make sense for an organism to adopt a set routine, or “work plan,” for carrying out various tasks in an optimal manner.

The second model considered in this paper shows that external factors may not be necessary for a temporally differentiated work plan to be an optimal labor strategy. This suggests that temporal differentiation may provide an evolutionary basis for sleep that applies to nocturnal organisms and may also provide an evolutionary basis for the emergence of distinct REM and non-REM sleep states.

Although external factors may not be necessary for a temporally differentiated work plan to be optimal, the first model considered in this paper nevertheless shows that it is optimal for a system to adopt a work plan that is synchronized with the external resource-availability profile. In the context of sleep and circadian rhythms, this provides a natural explanation for the observation that the day-night cycle serves as a powerful regulator of the sleep-wake cycle and of other circadian cycles.

For future research, we would like to develop more sophisticated models that consider the interaction of a large collection of system components (e.g., neurons) and to study the optimal temporally differentiated labor strategies that emerge. An interesting question is whether, given certain kinds of network topologies, universal scaling laws emerge that allow the prediction of various aspects of the sleep-wake cycle in one organism, based on known aspects of the sleep-wake cycle in another organism. If a model based on the concept of temporal differentiation could be used to make quantitative predictions of various aspects of the sleep-wake cycle in various organisms, this would provide compelling evidence for temporal differentiation as the evolutionary basis for the emergence of sleep.

We should note that, although mathematical modeling of sleep-wake cycles and circadian rhythms has been done before [45–48], such models have analyzed these rhythms from the perspective of coupled-oscillator dynamics and have therefore not addressed the evolutionary basis for the emergence of such oscillator systems in the first place. This paper, by contrast, developed simple mathematical models that suggests that temporal differentiation may be an optimal labor strategy in certain cases, so that organisms that employ temporally differentiated labor strategies for these cases may have a survival advantage.

Finally, we also recognize that temporal differentiation is a general concept that does not only apply to sleep and circadian rhythms, but may also be useful for understanding oscillatory dynamics in other complex systems. Temporal differentiation may also have practical applications as a tool for maximizing productivity in various multistep processes. If this is indeed the case, then it is possible that sleep is a subtle manifestation of a fairly general optimization strategy, so that tools from game theory, mathematical economics, and operations research could be used to analyze the phenomenon of sleep.

## ACKNOWLEDGMENTS

This research was supported by the Israel Science Foundation and by a Start-Up Grant from the United States-Israel Binational Science Foundation.

[1] J. F. Stuefer, H. de Kroon, and H. J. During, Funct. Ecol. **10**, 328 (1996).
[2] Y. Ben-Shahar, A. Robichon, M. B. Sokolowski, and G. E. Robinson, Science **296**, 741 (2002).
[3] J. A. Shapiro, Annu. Rev. Microbiol. **52**, 81 (1998).
[4] G. E. Robinson, Annu. Rev. Entomol. **37**, 637 (1992).
[5] Z. Huang and G. E. Robinson, Proc. Natl. Acad. Sci. U.S.A. **89**, 11726 (1992).
[6] G. S. Becker and K. M. Murphy, Quart. J. Econom. **107**, 1137 (1992).
[7] H. Demsetz, Am. Econ. Rev. **62**, 777 (1972).
[8] A. Strauss, Sociol. Q. **26**, 1 (1985).
[9] J. F. Francois, Oxford Econ. Papers **42**, 715 (1990).
[10] P. Bolton, Quart. J. Econom. **109**, 809 (1994).
[11] C. I. Jones, J. Polit. Econ. **103**, 759 (1995).
[12] A. J. Scott, Econo. Geogr. **62**, 215 (1986).
[13] R. H. Hall, N. J. Johnson, and J. E. Haas, Am. Sociol. Rev. **32**, 903 (1967).
[14] G. S. Becker and K. M. Murphy, Quart. J. Econom. **107**, 1137 (1992).
[15] Y. Ben-Shahar, A. Robichon, M. B. Sokolowski, and G. E. Robinson, Science **296**, 741 (2002).
[16] P. Bolton, Quart. J. Econom. **109**, 809 (1994).
[17] S. S. Chow, C. O. Wilke, C. Ofria, R. E. Lenski, and C. Adami, Science **305**, 84 (2004).
[18] H. Demsetz, Am. Econ. Rev. **62**, 777 (1992).
[19] J. F. Francois, Oxford Econ. Papers **42**, 715 (1990).
[20] C. Furusawa and K. Kaneko, J. Theor. Biol. **209**, 395 (2001).
[21] R. H. Hall, N. J. Johnson, and J. E. Haas, Am. Sociol. Rev. **32**, 903 (1967).
[22] E. Tannenbaum, J. Theor. Biol. **247**, 413 (2007).
[23] E. Akiyama and K. Kaneko, Artif. Life **2**, 293 (1995).
[24] S. Lammer, R. Donner, and D. Helbing (unpublished).
[25] R. Filliger and M. O. Hongler, Europhys. Lett. **70**, 285 (2005).
[26] F. Crick and G. Mitchison, Nature (London) **304**, 111 (1983).
[27] T. J. Sejnowski and A. Destexhe, Brain Res. **886**, 208 (2000).
[28] M. P. Walker and R. Stickgold, Annu. Rev. Psychol. **57**, 139 (2006).
[29] J. M. Krueger, F. Obal, Jr., and J. Fang, Sleep Med. Rev. **3**, 119 (1999).
[30] R. Stickgold and M. P. Walker, Trends Neurosci. **28**, 408 (2005).
[31] M. P. Walker, Behav. Brain Sci. **28**, 51 (2005).
[32] W. Wetzel, T. Wagner, and D. Balschun, Eur. J. Neurosci. **18**, 2611 (2003).
[33] C. B. Saper, T. C. Chou, and T. E. Scammell, Trends Neurosci. **24**, 726 (2001).
[34] R. B. Machado, D. Suchecki, and S. Tufik, Behav. Brain Res. **160**, 356 (2005).
[35] J. M. Siegel, P. R. Manger, R. Nienhuis, H. M. Fahringer, and J. D. Pettigrew, J. Neurosci. **16**, 3500 (1996).
[36] I. Tobler and T. Deboer, Sleep **24**, 147 (2001).
[37] F. S. Hillier and G. J. Lieberman, *Introduction to Operations Research*, 8th ed. (McGraw-Hill, Boston, 2005).
[38] A. C. Chiang and K. Wainwright, *Fundamental Methods of Mathematical Economics*, 4th ed. (McGraw-Hill, Boston, 2005).
[39] M. J. Osborne and A. Rubinstein, *A Course in Game Theory* (MIT Press, Cambridge, MA, 1994).
[40] A. A. Borbely and P. Achermann, J. Sleep Res. **1**, 63 (1992).
[41] A. A. Borbely and P. Achermann, J. Biol. Rhythms **14**, 557 (1999).
[42] M. E. Jewett and R. E. Kronauer, J. Biol. Rhythms **14**, 588 (1999).

[43] S. H. Strogatz, *The Mathematical Structure of the Human Sleep-Wake Cycle, Lecture Notes in Biomathematics* (Springer-Verlag, New York, 1986).
[44] D. Leger, C. Guilleminault, R. Defrance, A. Domont, and M. Paillard, Clin. Sci. **97**, 193 (1999).
[45] A. T. Winfree, *The Geometry of Biological Time*, 2nd ed. (Springer-Verlag, Berlin, 2001).
[46] A. T. Winfree, *The Timing of Biological Clocks* (Scientific American Library, New York, 1987).
[47] A. T. Winfree, Nature (London) **297**, 23 (1982).
[48] A. T. Winfree, J. Comp. Physiol. [A] **85**, 105 (1973).