Why Don't We Move Slower? The Value of Time in the Neural Control of Action

Bastien Berret, Frédéric Jean

Behavioral/Cognitive

Title

Why don't we move slower? The value of time in the neural control of action

Running Title: The value of time in the neural control of action

Authors

Bastien Berret[1] and Frédéric Jean[2]

1. Univ Paris-Sud, CIAMS EA4532, MHAPS team
2. ENSTA Paris-Tech, Unité de Mathématiques Appliquées

Correspondence should be addressed to Bastien Berret, Université Paris-Sud, Bâtiment 335, Laboratoire CIAMS, F-91405 Orsay, France (bastien.berret@u-psud.fr).

9 Figures and 0 Table
35 Pages (without figures and tables)
Number of words for Abstract: 250
Number of words for Introduction: 603
Number of words for Discussion: 1757

Acknowledgments: This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the « Investissement d'Avenir » program, through the "iCODE Institute project" funded by the IDEX Paris-Saclay, ANR-11-IDEX-0003-02. We thank Thomas Deroche, Carole Castanier, and Jean Jeuvrey for their help during the acquisition of the experimental data. We are indebted to François Bonnetblanc, Francesco Nori and Jérémie Gaveau for their useful comments on earlier versions of the manuscript.

Abbreviations: CoT = cost of time, MT = movement time, LQ = linear quadratic, OC = optimal control, PMP = Pontryagin Maximum Principle, BG = basal ganglia

Abstract

To want something now rather than later is a common attitude that reflects the brain's tendency to value the passage of time. Because the time taken to accomplish an action inevitably delays task achievement and reward acquisition, this idea was ported to neural movement control within the "cost of time" theory. This theory provides a normative framework to account for the underpinnings of movement time formation within the brain and the origin of a self-selected pace in human and animal motion. Then, how does the brain exactly value time in the control of action? To tackle this issue, we employed an inverse optimal control approach and developed a general methodology allowing to squarely sample infinitesimal values of the time cost from experimental motion data. The cost of time underlying saccades was found to have a concave growth, thereby confirming previous results on hyperbolic reward discounting yet without making any prior assumption about this hypothetical nature. For self-paced reaching, however, movement time was primarily valued according to a striking sigmoidal shape; its rate of change consistently presented a steep rise before a maximum was reached and a slower decay was observed. Theoretical properties of uniqueness and robustness of the inferred time cost were established for the class of problems under investigation, thus reinforcing the significance of the present findings. These results may offer a unique opportunity to uncover how the brain values the passage of time in healthy and pathological motor control and shed new light on the processes underlying action invigoration.

Significance Statement

Movement time is a fundamental characteristic of neural motor control but the principles underlying its formation remain little known. This work addresses that question within the inverse optimal control framework where the challenge is to uncover what optimality criterion underlies a system's behavior. Here we rely on the "cost of time" theory that finds its roots into the brain's tendency to discount the actual value of future reward. It asserts that the time elapsed until action completion entails a cost, thereby making slow moves non-optimal. By means of a thorough theoretical analysis, the present article shows how to sample the infinitesimal values of the time cost without prior assumption about its hypothetical nature and emphasizes its sigmoidal shape for reaching.

Introduction

Movement time (MT) is an inherent characteristic of motor control but the profound principles underlying its formation, be they neural or computational, remain in practice little understood. Prediction of MT is however a critical issue in all fields concerned with biological movement such as humanoid/rehabilitation robotics, neuroprosthetics, brain-machine interfaces or computer animation. In neuroscience, investigating the underpinnings of one's own motion pace is of crucial importance as many motor disorders like Parkinson's disease lead to bradykinesia (Berardelli et al., 2001). This symptomatic movement slowness, not reducible to musculoskeletal or sensorimotor deficits, seems to be due to a lack of implicit "motor motivation" (Mazzoni et al., 2007; Baraduc et al., 2013). Such a reduced action vigor mainly originates from a dysfunction of the basal ganglia (BG) and the dopamine system (Turner and Desmurget, 2010; Desmurget and Turner, 2010; Brücke et al., 2012), thereby showing that MT is not just an emergent property of the task but a planned quantity. While it is well-understood why movements cannot be too fast, as exemplified by the speed/accuracy trade-off (Fitts, 1954; Harris and Wolpert, 1998; Tanaka et al., 2006), few studies actually addressed the complementary question of why movements are not slower.

Recently, the theory of the cost of time (CoT) has emerged as a promising avenue to tackle the issue of action invigoration (Shadmehr, 2010; Shadmehr et al., 2010; Shadmehr and Mussa-Ivaldi, 2012). This theory was originally motivated by the natural tendency of the brain to decrease the actual value of future reward (Myerson and Green, 1995). Porting this psycho-economical decision making to motor control (Wolpert and Landy, 2012), Shadmehr and colleagues proposed that the purpose of any action would be to put the neural system in a more rewarding state. Therefore, the slower the movement, the smaller the acquired reward. In other words, slow movements may be undesirable because the very passage of time entails a cost *per se*, thus placing the problem within the normative framework of optimal control (OC) theory (Todorov, 2004; Scott, 2004). The generic concept of a CoT revealed itself sufficient to account for MT in a variety of motor tasks (Hoff, 1994; Liu and Todorov, 2007; Shadmehr, 2010; Shadmehr et al., 2010; Rigoux and Guigon, 2012). The actual shape of the time cost remains however rather elusive in neural motor control because it was chosen *a priori* in most existing investigations. Researchers have assumed linear (Hoff, 1994; Harris and Wolpert, 2006; Liu and Todorov, 2007), concave (hyperbolic, Shadmehr,

2010; Haith et al., 2012 or exponential, Huh et al., 2010; Rigoux and Guigon, 2012) and even convex (quadratic, Shadmehr et al., 2010) shapes for the CoT (Fig. 1a). This choice could be motivated by psychological/behavioral observations, imposed by the mathematical constraints of the problem's formulation or just guided by the researcher's intuition.

In the present study, we instead employed an inverse OC approach where the goal is to automatically uncover what optimality criterion underlies a (presumably optimal) system's behavior. To this aim, we developed a complete and robust methodology for characterizing how the brain values time from basic experimental motion data. The hyperbolic nature of the temporal discounting of reward (Shadmehr et al., 2010; Haith et al., 2012) was reaffirmed for saccades but without making any prior assumption about its hypothetical nature. When applied to reaching, a striking sigmoidal shape of the CoT was discovered over the time interval of actual reach durations, thereby ruling out purely concave or convex time costs for limb movement control. These methodological and empirical breakthroughs may offer a unique opportunity to assess how the brain values time in neural movement control across tasks, individuals or species.

Materials and Methods

Theoretical analysis

Working hypotheses This work relies upon the general assumption that biological movement is optimal with respect to some cost function (Engelbrecht, 2001; Todorov, 2004). It generally supposes that the trajectories triggered by the central nervous system can be accounted for by a certain infinitesimal cost $h(\mathbf{x}, \mathbf{u}, t)$ depending on the system state, the motor command and the time, respectively. Originally, models of OC for biological movement were developed in fixed time where MT was typically taken from experimental measurements. This restriction can be released in free-time OC whereby MT emerges from the optimality of behavior (Pontryagin et al., 1964; Kirk, 1970). In this context and in agreement with the CoT theory, we further assume that h can be separated into a term that values time only, $g(t)$ (the infinitesimal CoT) and a term that depends on the state/control variables, $l(\mathbf{x}, \mathbf{u})$ (see Fig. 1b). A mathematical treatment of that problem, given below, shows that it is actually possible to compute $g(t)$ by resolving an OC problem in fixed time t with known initial/final states, given a system dynamics and trajectory

cost. When the dynamics is linear with a single control variable and the trajectory cost is quadratic (i.e. a single-input LQ problem), there moreover exist mathematical properties of uniqueness and robustness for the CoT, guaranteeing that the CoT can be identified unequivocally for that class of problem. Our main hypothesis is thus about the additive separability of temporal and trajectory costs, which is consistent with the computational motor control literature and coherent with all the previous works devised in fixed time (see below for a justification of this assertion). An alternative formulation, often seen in reinforcement learning, may have assumed multiplicative separability (e.g. exponential discounting in an infinite-horizon setting, Sutton and Barto, 1998). While we do not have theoretical results in the latter case, Rigoux and Guigon (2012) showed that under certain restrictive assumptions (e.g. reward acquired on a single time step and state) a similar free-time additive setting can be recovered. Note however that assuming a multiplicative temporal cost would be a major change of paradigm with respect to the motor control literature as the presence of such a term could modify all the predictions made by classical fixed-time optimal control models.

General methodology for identifying the cost of time Consider a system dynamics $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u})$, with state $\mathbf{x} \in \mathbb{R}^n$ and control $\mathbf{u} \in \mathbb{R}^m$, and fix a target $\mathbf{x}^f$ (f stands for *final* throughout the text). A dot above a variable stands for its time derivative. Given an input $\mathbf{u}(\cdot)$ defined on an interval $[0, t_\mathbf{u}]$, we denote by $\mathbf{x}_\mathbf{u}(\cdot)$ the trajectory of $\dot{\mathbf{x}}(t) = \mathbf{f}\big(\mathbf{x}(t), \mathbf{u}(t)\big)$ arriving at $\mathbf{x}^f$, i.e. satisfying $\mathbf{x}_\mathbf{u}(t_\mathbf{u}) = \mathbf{x}^f$. We assume that the cost function writes as the combination of the CoT plus a trajectory cost, as follows:

$$C(\mathbf{u}, t_\mathbf{u}) = \int_0^{t_\mathbf{u}} \big(g(t) + l\big(\mathbf{x}_\mathbf{u}(t), \mathbf{u}(t)\big)\big)\ dt, \tag{1}$$

where the functions g and l are non-negative. The function l is classical to motor control and may capture various aspects of the trajectory such as effort, energy, smoothness or accuracy. It has been the subject of extensive investigations (e.g. Berret et al., 2011). It will be referred to as the trajectory cost throughout the study. What we know about l is that it must account for the system trajectories in fixed time (i.e. when MT is known). To get persuaded of this, it suffices to realize that a model predicting MT accurately would be meaningless if it does not predict the correct system trajectories in fixed time as well. In what follows, we shall assume that l is either

known from the literature or can be identified via some inverse OC procedure performed in fixed-time (e.g. Mombaur et al., 2009; Berret et al., 2011). The function g is the infinitesimal (i.e. instantaneous) CoT we are looking for, whose antiderivative is the actual CoT and will be denoted by $G(t) - G(0) = \int_0^t g(s)ds$ (in the sequel, we will assume that $G(0) = 0$).

We consider the following *free-time* OC problems:

Given an initial condition $\mathbf{x}^0$*, minimize the cost* $C(\mathbf{u}, t_\mathbf{u})$ *among all inputs* $\mathbf{u}(\cdot)$ *and all times* $t_\mathbf{u}$ *such that* $\mathbf{x}_\mathbf{u}(0) = \mathbf{x}^0$ *and* $\mathbf{x}_\mathbf{u}(t_\mathbf{u}) = \mathbf{x}^f$ *(by definition of* $\mathbf{x}_\mathbf{u}$*).*

The existence of minimal solutions $\mathbf{u}(\cdot)$ with a finite time $t_\mathbf{u}$ may be guaranteed under some technical conditions on the dynamics and on the cost (typically, compactness of the set of controls and convexity of $\mathbf{f}$ and l with respect to $\mathbf{u}$, or $\mathbf{f}$ linear and l quadratic; e.g. Lee and Markus, 1967). We will assume that such conditions hold here.

Let $\mathbf{u}(\cdot)$ be a minimal solution of the problem. Then there exists a curve $\mathbf{p}(t)$, $t \in [0, t_\mathbf{u}]$, in $\mathbb{R}^n$ called the *adjoint or co-state vector*, and a real number $\lambda = 0$ or 1, such that $\big(\mathbf{x}(\cdot), \mathbf{p}(\cdot), \mathbf{u}(\cdot), \lambda\big)$ satisfies the conditions of the Pontryagin Maximum Principle (PMP for short, e.g. Pontryagin et al. 1964; Lee and Markus 1967; Todorov 2006). In particular, defining the Hamiltonian of the problem as

$$\mathcal{H}(t, \mathbf{x}, \mathbf{p}, \mathbf{u}, \lambda) = \mathbf{p}^\top \mathbf{f}(\mathbf{x}, \mathbf{u}) + \lambda\big(l(\mathbf{x}, \mathbf{u}) + g(t)\big), \tag{2}$$

the adjoint vector satisfies the Hamiltonian equation $\dot{\mathbf{p}} = -\frac{\partial \mathcal{H}}{\partial \mathbf{x}}(t, \mathbf{x}, \mathbf{p}, \mathbf{u}, \lambda)$; moreover, the free-time setting gives the following additional equation (called transversality condition):

$$\mathcal{H}\big(t_\mathbf{u}, \mathbf{x}(t_\mathbf{u}), \mathbf{p}(t_\mathbf{u}), \mathbf{u}(t_\mathbf{u}), \lambda\big) = 0. \tag{3}$$

Note that the case $\lambda = 0$ occurs only for singular controls of the dynamics $\mathbf{f}(\mathbf{x}, \mathbf{u})$ (these controls are such that the linearization of the system around the associated trajectory $\mathbf{x}_\mathbf{u}$ is not controllable). We will assume that no such control exist for $\mathbf{f}(\mathbf{x}, \mathbf{u})$ (this is always true for controllable linear systems for instance), and so we can choose $\lambda = 1$ in $\mathcal{H}$. As a consequence we obtain

$$g(t_\mathbf{u}) = -\mathbf{p}(t_\mathbf{u})^\top \mathbf{f}\big(\mathbf{x}(t_\mathbf{u}), \mathbf{u}(t_\mathbf{u})\big) - l\big(\mathbf{x}(t_\mathbf{u}), \mathbf{u}(t_\mathbf{u})\big). \tag{4}$$

On the other hand, it is obvious that $\mathbf{u}(\cdot)$ is also a minimal solution of the following OC problem

in *fixed time* $t_{\mathbf{u}}$:

Minimize the cost

$$C_{t_{\mathbf{u}}}(\mathbf{v}) = \int_0^{t_{\mathbf{u}}} l\big(\mathbf{x}_{\mathbf{v}}(t), \mathbf{v}(t)\big)dt, \tag{5}$$

among all inputs $\mathbf{v}(\cdot)$ *such that* $\mathbf{x}_{\mathbf{v}}(0) = \mathbf{x}^0$ *(and* $\mathbf{x}_{\mathbf{v}}(t_{\mathbf{u}}) = \mathbf{x}^f$*).* The Hamiltonian associated with this problem is

$$\mathcal{H}_0(\mathbf{x}, \mathbf{p}, \mathbf{u}, \lambda) = \mathbf{p}^\top \mathbf{f}(\mathbf{x}, \mathbf{u}) + \lambda l(\mathbf{x}, \mathbf{u}). \tag{6}$$

Note that $\mathcal{H}_0$ defines the same Hamiltonian equation than $\mathcal{H}$ since $\frac{\partial \mathcal{H}_0}{\partial \mathbf{x}} = \frac{\partial \mathcal{H}}{\partial \mathbf{x}}$, so $\mathbf{p}(\cdot)$ is also an adjoint vector for this OC problem in fixed time, and it is the only one since the adjoint vector associated with $\lambda = 1$ is unique.

Thus, for a given time $t_{\mathbf{u}}$ and a given cost l, we can compute the value $g(t_{\mathbf{u}})$ in the following way. First, solve the fixed-time OC problem in time $t_{\mathbf{u}}$. The result is a triple $(\mathbf{x}(\cdot), \mathbf{p}(\cdot), \mathbf{u}(\cdot))$ from which the Hamiltonian can be calculated. Second, set

$$g(t_{\mathbf{u}}) = -\mathcal{H}_0\big(\mathbf{x}(t_{\mathbf{u}}), \mathbf{p}(t_{\mathbf{u}}), \mathbf{u}(t_{\mathbf{u}}), 1\big) \tag{7}$$

to get the value of the infinitesimal CoT at time $t_{\mathbf{u}}$. A flowchart illustrating the process is given in Figure 2.

A more elegant way to obtain the same conclusion is to proceed as follows. Let $V(t, \mathbf{x}^0)$ be the value function of the OC problem in fixed-time t, that is

$$V(t, \mathbf{x}^0) = \inf \int_0^t l\big(\mathbf{x}_{\mathbf{u}}(s), \mathbf{u}(s)\big)ds, \tag{8}$$

where the infimum is taken among all inputs $\mathbf{u}(\cdot)$ such that $\mathbf{x}_{\mathbf{u}}(0) = \mathbf{x}^0$ (and $\mathbf{x}_{\mathbf{u}}(t) = \mathbf{x}^f$). It is the optimal cost of a motion in time t between $\mathbf{x}^0$ and $\mathbf{x}^f$. Then the time $t_{\mathbf{u}}$ of an optimal solution $\mathbf{u}(\cdot)$ of the free-time problem satisfies

$$t_{\mathbf{u}} \in \operatorname{argmin}_{t \geq 0} \left(\int_0^t g(s)ds + V(t, \mathbf{x}^0) \right), \tag{9}$$

and so necessarily $g(t_{\mathbf{u}}) + \frac{\partial V}{\partial t}(t_{\mathbf{u}}, \mathbf{x}^0) = 0$ (we assume here that V is differentiable with respect to t). It is well-known from the Hamilton-Jacobi-Bellman (HJB) theory that $\frac{\partial V}{\partial t}(t_{\mathbf{u}}, \mathbf{x}^0) = \mathcal{H}_0^\star\big(\mathbf{x}(t_{\mathbf{u}}), \mathbf{p}(t_{\mathbf{u}}), 1\big)$,

where $\mathcal{H}_0^\star(\mathbf{x},\mathbf{p},1) = \max_{\mathbf{v}} \mathcal{H}_0(\mathbf{x},\mathbf{p},\mathbf{v},1)$. Since for the optimal control $\mathbf{u}$ we have $\mathcal{H}_0^\star(\mathbf{x}(t_\mathbf{u}),\mathbf{p}(t_\mathbf{u}),1) = \mathcal{H}_0\big(\mathbf{x}(t_\mathbf{u}),\mathbf{p}(t_\mathbf{u}),\mathbf{u}(t_\mathbf{u}),1\big)$, we recover in this way Eq. (7). Note that we did not use the standard way to define the value function: for a MT equal to t, this is usually $\tilde{V}(w,\mathbf{x}(w)) = \inf \int_w^t l\big(\mathbf{x}_\mathbf{u}(s),\mathbf{u}(s)\big)ds$. Here we set $V(t-w,\mathbf{x}(w)) = \tilde{V}(w,\mathbf{x}(w))$, hence $\frac{\partial V}{\partial t} = -\frac{\partial \tilde{V}}{\partial t}$.

Remarkably, the latter analysis shows that the derivation extends to stochastic settings (Stengel, 1986; Todorov, 2006). In that case, the infinitesimal CoT is simply the partial time derivative of the expected value function of the stochastic OC problem in fixed-time $t_\mathbf{u}$. More precisely, for the stochastic dynamics $d\mathbf{x} = \mathbf{f}(\mathbf{x},\mathbf{u})dt + \mathbf{g}(\mathbf{x},\mathbf{u})d\xi$ where $d\xi$ is a standard Wiener process, $g(t_\mathbf{u})$ results from the stochastic HJB equation:

$$\begin{aligned} g(t_\mathbf{u}) &= \tfrac{\partial V}{\partial t}(t_\mathbf{u},\mathbf{x}^0) \\ &= \min_\mathbf{u} \Big(l(\mathbf{x},\mathbf{u}) + \tfrac{\partial V}{\partial \mathbf{x}}(t_\mathbf{u},\mathbf{x}^0)^\top \mathbf{f}(\mathbf{x},\mathbf{u}) \\ &\quad + \frac{1}{2}\mathrm{trace}\big(\tfrac{\partial^2 V}{\partial \mathbf{x}^2}(t_\mathbf{u},\mathbf{x}^0)\mathbf{g}(\mathbf{x},\mathbf{u})\mathbf{g}(\mathbf{x},\mathbf{u})^\top\big)\Big) \end{aligned} \tag{10}$$

In the linear quadratic Gaussian (LQG) case, the infinitesimal CoT can be easily computed because the value function has a parametric form whose parameters can be evaluated via the resolution of decoupled ordinary differential equations (e.g. Kappen, 2011).

In summary, the above theoretical considerations show that $g(t)$ can be calculated by solving a deterministic or stochastic OC problem in fixed time t for some initial/final states of the system ($\mathbf{x}^0$ and $\mathbf{x}^f$). Movement duration and initial/final states are basic information that can be retrieved from experimental data. Thus, given a trajectory cost l and a model of the system dynamics, it is straightforward to sample the values of the mapping $t \mapsto g(t)$ on some time interval $[t_{\min}, t_{\max}]$. The boundary values correspond to the range where real observations can be taken in practice and come out experimentally. This is the empirical range of movement durations. As a consequence, the integral CoT (i.e. G) can only be identified up to a constant since the values of $g(t)$ for $t < t_{\min}$ are unknown/unobservable. This constant shift however does not affect neither the solution of the free-time OC problem nor the predicted amplitude-duration relationship over the range $[t_{\min}, t_{\max}]$ (that will be exactly recovered on this interval). To infer the CoT outside of the range of actual measurements, we may devise at least two ways. First, the identified CoT may just be extrapolated for $t < t_{\min}$ and $t > t_{\max}$. Alternatively, values of $g(t)$ may be computed for unobserved times by

relying on predictions of the empirical amplitude-duration relationship (e.g. obtained via regression analysis). Values inferred using the first and second methods will be respectively displayed as dotted and solid lines in all graphs depicting the CoT. Values inferred over the range of actual reach durations will be emphasized by filled circles added on the top of solid lines, which locate where the CoT was the most reliably identified.

Detailed solution for single-input LQ problems For deterministic single-input linear-quadratic problems, a more detailed account of the methodology can be given and closed-form solutions of the CoT can be obtained. This case is relevant as it includes basic eye and arm movements. The state of such systems can be described by $\mathbf{x} = (\theta, \dots, \theta^{(n-1)}) \in \mathbb{R}^n$ and then the dynamics has the form

$$\theta^{(n)} + c_{n-1}\theta^{(n-1)} + \cdots + c_0\theta = u, \tag{11}$$

which is a single-input linear system $\dot{\mathbf{x}} = A\mathbf{x} + Bu$, $u \in \mathbb{R}$. Typically $n = 2$ or 3 for dynamical models of the arm or the eye (see below). Let us assume that a quadratic cost $l(\mathbf{x}, u) = (u - \mathbf{k}^T\mathbf{x})^2$ has been determined (e.g. an "effort" cost). We choose the final point $\mathbf{x}^f$ as an equilibrium state of the system. Up to a translation in θ, we can always assume $\mathbf{x}^f = 0$. We then choose a family of initial conditions $\mathbf{x}^0(a) = (a, 0, \dots, 0)$, which are equilibrium states parameterized by the movement extent $a > 0$ (i.e. the amplitude of the motion). For every amplitude $a > 0$ we denote by $t^*(a)$ the duration of the motion between $\mathbf{x}^0(a)$ and $\mathbf{x}^f$ (which can be estimated experimentally), and by $u^a(\cdot)$ a control minimizing the integral cost $C_{t^*(a)}(u) = \int_0^{t^*(a)} \left(u - \mathbf{k}^T\mathbf{x}_u(t)\right)^2 dt$ in fixed time $t^*(a)$ between $\mathbf{x}_u(0) = \mathbf{x}^0(a)$ and $\mathbf{x}_u(t^*(a)) = \mathbf{x}^f = 0$. Applying the above methodology, we have to compute $\mathcal{H}_0(\mathbf{x}(t^*(a)), \mathbf{p}(t^*(a)), u(t^*(a)), 1)$ where

$$\mathcal{H}_0(\mathbf{x}, \mathbf{p}, u, 1) = \mathbf{p}^T(A\mathbf{x} + Bu) + (u - \mathbf{k}^T\mathbf{x})^2. \tag{12}$$

From the PMP, $u^a(t^*(a))$ minimizes $\mathcal{H}_0\big(\mathbf{x}(t^*(a)), \mathbf{p}(t^*(a)), u, 1\big)$ with respect to u, which means $\frac{\partial \mathcal{H}_0}{\partial u} = 0$, and replacing the solution in Eq. 12, we obtain

$$\mathcal{H}_0(\mathbf{x}(t^*(a)), \mathbf{p}(t^*(a)), u^a(t^*(a)), 1) = -u^a(t^*(a))^2. \tag{13}$$

Moreover, the value $u^a(t^*(a))$ can be seen to depend linearly on $\mathbf{x}_u(0)$ in the LQ case, and so it depends linearly on a since $\mathbf{x}_u(0) = \mathbf{x}^0(a) = a\mathbf{x}^0(1)$. In other words, $u^a(t^*(a)) = a\varphi(t^*(a))$, where the function $\varphi(\cdot)$ is defined as follows: *for every* $\tau > 0$, $\varphi(\tau)$ *is the value* $u^1(\tau)$ *of the control minimizing the integral cost* $C_\tau(u) = \int_0^\tau (u - \mathbf{k}^T\mathbf{x}_u(t))^2 dt$ *in fixed time* τ *between* $\mathbf{x}_u(0) = \mathbf{x}^0(1)$ *and* $\mathbf{x}_u(\tau) = \mathbf{0}$. Note that the function $\varphi(\cdot)$ is a *universal value of time* that depends only on the system dynamics (A, B) and the trajectory cost ($\mathbf{k}$) and not on the specific behavior of an individual. This universal value of time can be computed explicitly thanks to the equations given in Ferrante et al. (2005). Applying formula given in Eq. (7), we finally obtain $g(t^*(a)) = \varphi(t^*(a))^2 a^2$.

Empirical observations show that the time $t^*(a)$ is typically an increasing function of the amplitude, so that its inverse $a^*(t)$ exists. We can then determine the function $g(\cdot)$ by $g(t) = \varphi(t)^2 a^*(t)^2$. In particular, if it appears from experiments that the function t^* is approximately affine of the form $t^*(a) = \alpha a + \beta$, then the infinitesimal CoT can be written $g(t) = \varphi(t)^2(\frac{1}{\alpha}t - \frac{\beta}{\alpha})^2$. Hence, it suffices to compute the universal value of time $\varphi(t)$, which can be done explicitly, to recover the actual infinitesimal CoT from the experimental duration/amplitude mapping.

Interestingly, this analysis also reveals the effect of globally rescaling g by multiplying it by some positive parameter κ. This would induce a new amplitude $\sqrt{\kappa}a^*(t)$. Therefore, the rescaled CoT would lead to the affine amplitude-duration relationship $t^*(a) = \alpha\frac{a}{\sqrt{\kappa}} + \beta$. It can be concluded that simply rescaling the CoT does not allow to change both the slope and the intercept of the amplitude/duration relationship. Hence, if both the intercept and slope are found to change experimentally, this variation cannot be attributed to a global rescaling of the time cost. The latter point is important if attempting to account for speed variations and inter-individual differences (see Results).

Uniqueness and robustness properties for single-input LQ problems From the above considerations, it is clear that the values of the infinitesimal CoT may critically depend on the trajectory cost. This raises the question about the robustness and uniqueness of g in general as the cost l may be known only approximately or may even be quite different from the true cost without it being noticeable in the motion data under investigation. Although important, this question was not addressed in previous studies. The problem can be exemplified for planar point-to-point reaches where the torque change (Uno et al., 1989) and jerk (Flash and Hogan, 1985) trajectory costs are

hardly distinguishable despite their different nature (dynamic versus kinematic costs). For the CoT theory, different trajectory costs may conceivably lead to divergent time costs. Thus, before identifying the CoT, the question of what trajectory cost underlies the system trajectories with known MT must be posed. It may be argued that the trajectory cost must be identified through the resolution of the following inverse OC problem in fixed time:

Given a set of recorded trajectories in time τ, find l such that every recorded trajectory $\mathbf{y}(t)$ is a minimum of the integral cost

$$C_\tau(\mathbf{x}_u) = \int_0^\tau l(\mathbf{x}_u(t), u(t))dt, \tag{14}$$

among all trajectories $\mathbf{x}_u$ such that $\mathbf{x}_u(0) = \mathbf{y}(0)$ and $\mathbf{x}_u(\tau) = \mathbf{y}(\tau)$.

However, this inverse problem is ill-posed in general, which means that its solutions $l(\mathbf{x}, u)$ are neither unique nor stable with respect to perturbations of the recorded data (which are affected by sensorimotor and measurement noise and subject to inter-individual variability). Within the single-input LQ framework, however, powerful results of uniqueness and robustness can be obtained, as explained hereafter.

We thus assume that the class of admissible trajectory costs is the class $\mathcal{L}$ of quadratic costs $\mathcal{L} = \{l(\mathbf{x}, u) = u^2 + \mathbf{x}^T Q\mathbf{x} + 2\mathbf{x}^T Su\}$, with standard hypothesis on the matrices Q and S (i.e. $Q = Q^T \geq 0$ and S is such that l is a positive semidefinite quadratic form). Given a cost $l \in \mathcal{L}$, the classical fixed-time LQ theory asserts that the solution of the corresponding OC problem in time τ satisfies $u(t) = K_\tau(t)\mathbf{x}(t)$, where the time-dependent matrices $K_\tau(t)$ are determined through some Riccati equations. Alternatively, Ferrante and collaborators Ferrante et al. (2005) showed that all solutions in fixed time are determined by only two $(1 \times n)$ matrices K_-, K_+, i.e., the set $\{K_\tau(\cdot), \tau > 0\}$ is uniquely characterized by a pair (K_-, K_+). Following the latter approach, let us introduce the set $\mathcal{K}$ of all such pairs, that is, (K_-, K_+) belongs to $\mathcal{K}$ if there exists a LQ problem associated with a cost $l \in \mathcal{L}$ whose solutions are characterized by (K_-, K_+). The following Theorem is proved in Nori and Frezza (2004).

Theorem 1. *For any pair $(K_-, K_+) \in \mathcal{K}$ there exists a unique vector $\mathbf{k} \in \mathbb{R}^n$ such that (K_-, K_+) describes the optimal solutions of the fixed-time OC problems associated with*

$$\int_0^\tau (u - \mathbf{k}^T\mathbf{x})^2 dt. \tag{15}$$

Moreover the map $(K_-, K_+) \mapsto \mathbf{k}$ *is continuous (actually* $\mathbf{k} = -K_+^T$*).*

As a consequence, for single-input LQ problems, there exists a complete methodology for determining unequivocally and robustly the cost of the time g. The method is summarized as follows: **(*i*)** determine a pair $(K_-, K_+) \in \mathcal{K}$ such that the associated optimal trajectories match accurately the recorded ones in fixed time (this can be written as a least-square problem); **(*ii*)** set $\mathbf{k} = -K_+^T$ and $l(\mathbf{x}, u) = (u - \mathbf{k}^T\mathbf{x})^2$; **(*iii*)** compute the function $\frac{\partial V}{\partial t}(t, \mathbf{x})$ by using the Hamiltonian $\mathcal{H}_0$ associated with $l(\mathbf{x}, u)$; **(*iv*)** for every time t, identify from experimental data an initial condition $\mathbf{x}^*(t)$ corresponding to a movement in time t (if it exists for that t); **(*v*)** set $g(t) = -\frac{\partial V}{\partial t}(t, \mathbf{x}^*(t))$. The latter analysis proves that the CoT can be unequivocally and robustly identified from experimental data for single-input LQ problems. Although restricted, this class of problems covers most of the applications considered in the present study, which reinforces the significance of our findings. For more general problems, either the literature already provides some widely-accepted trajectory cost or numerical techniques of inverse optimal control can be employed to identify the cost function best replicating the system trajectories in fixed time (e.g. Berret et al., 2011).

Computational procedures

Models for saccades The eye plant model was taken from Shadmehr et al. (2010). Briefly, the eye dynamics was single-input linear system as follows:

$$c_1 \dddot{\theta} + c_2 \ddot{\theta} + c_3 \dot{\theta} + c_4 \theta = u \tag{16}$$

where $c_1 = 1.165 \times 10^{-5}$, $c_2 = 3.9 \times 10^{-3}$, $c_3 = 0.241$ and $c_4 = 1$ for a human eye. The system state was $\mathbf{x}^\top = (\theta, \dot{\theta}, \ddot{\theta})^\top$ and an effort cost $l(\mathbf{x}, u) = u^2$ was minimized. Note that this problem falls within the single-input LQ framework described above.

We also performed simulations with a signal-dependent noise following Shadmehr et al. (2010) and identified the CoT in this stochastic context. Signal-dependent noise is a random variable with a normal distribution of mean zero and variance equal to $k_{\mathrm{SDN}} u^2$. This problem then falls within the LQG framework, for which closed-form solutions are available to compute the value function and in particular its partial time derivative (see Kappen, 2011 for the equations). The default noise level (signal-dependent noise) was set to $k_{\mathrm{SDN}} = 0.0075$ and a terminal cost accounting for accuracy

of the form $\mathbf{x}^\top D\mathbf{x}$ with $D = \mathrm{diag}(10^3, 1, 10^{-2})$ was added to the integral cost.

Models for reaching *Single degree-of-freedom (dof) limb.* For a 1-dof arm moving in the horizontal plane, the basic model used throughout the study was already described in numerous other studies (e.g. Hogan 1984; Gentili et al. 2007; Tanaka et al. 2006; Gaveau et al. 2014) and is as follows:

$$\begin{cases} I\ddot{\theta} &= \tau - b\dot{\theta} \\ \dot{\tau} &= u \end{cases} \tag{17}$$

where is θ the shoulder joint angle, τ is the muscle torque, b is the friction coefficient ($b = 0.87$ here), I is the moment of inertia of the arm with respect to the shoulder joint (value estimated based upon Winter's table for each participant; Winter, 1990) and u is the single control variable.

For the trajectory cost we typically considered canonical quadratic costs of the form $l(\mathbf{x}, u) = (u - \mathbf{k}^\top\mathbf{x})^2 dt$, where $\mathbf{x}^\top = (\theta, \dot{\theta}, \ddot{\theta})$ denotes the system state. The two most famous examples are the minimum torque change corresponding to $l(\mathbf{x}, u) = u^2$ (i.e. $\mathbf{k} = \mathbf{0}$; Uno et al., 1989) and the minimum jerk corresponding to $l(\mathbf{x}, u) = \dddot{\theta}^2$ (i.e. $\mathbf{k}^\top = (0, 0, b)$; Flash and Hogan, 1985). However, other costs, possibly composite, may account for such planar movements. Therefore, additional perturbations of the effort cost were tested as follows: a torque-based cost ($l(\mathbf{x}, u) = 0.05u^2 + \tau^2$) (Nelson, 1983), an energy-based cost (absolute work, $l(\mathbf{x}, u) = 0.05u^2 + |\tau\dot{\theta}|$) (Berret et al., 2008), and an acceleration-based cost ($l(\mathbf{x}, u) = 0.05u^2 + \ddot{\theta}^2$) (Ben-Itzhak and Karniel, 2008). All these costs were checked to predict symmetrical and smooth velocity profiles and thus to be plausible trajectory costs for such a type of movement under consideration. For visualization purposes, we considered when convenient $l(\mathbf{x}, u) = 0.05u^2$ and $l(\mathbf{x}, u) = 0.05\dddot{\theta}^2$ for the torque change and jerk costs, respectively.

Sensorimotor noise is known to play a role on motor control during reaching, especially signal-dependent noise (Todorov and Jordan, 2002; van Beers et al., 2004). Therefore, we also tested signal-dependent noise among values in the following range $k_{\mathrm{SDN}} = [0; 0.05; 0.075; 0.1; 0.125]$. In this stochastic context, a terminal cost accounting for accuracy of the form $\mathbf{x}^\top D\mathbf{x}$ with $D = 5\mathrm{diag}(10^5, 10^3, 10^2)$ was added to the integral cost.

1-dof with muscle dynamics. For the sake of completeness, we also tested a more advanced model of the arm, namely a rigid body actuated by two antagonist muscles modeled as second-

order low-pass filters (Van der Helm and Rozendaal, 2000) with certain time constants. In that case, the single-joint arm dynamics could be written:

$$\begin{cases} I\ddot{\theta} & = \tau_1 - \tau_2 - b\dot{\theta} \\ \dot{\tau}_i & = (\epsilon_i - \tau_i)/\sigma,\ i = 1, 2 \\ \dot{\epsilon}_i & = (u_i - \epsilon_i)/\sigma,\ i = 1, 2 \end{cases} \tag{18}$$

where the subscripts 1 and 2 denote the agonist and antagonist muscles respectively, ϵ is the muscle excitation, u_i is the control variable, which can be thought of as motor neuron input, and σ is the time constant of the filter. Typically the value is chosen about 0.04 s in the literature for fast movements (e.g. Guigon et al., 2007; Liu and Todorov, 2007) but we tested values ranging from 0.03 s to 0.08 s to evaluate the effects of varying this parameter on the shape of the CoT. In the reported simulations, we minimized the following effort cost: $l(\mathbf{x}, u) = 0.005(u_1^2 + u_2^2)$.

2-dof arm. For a nonlinear multijoint planar arm model, a full description and the parameters used for the present simulations can be found in previously published articles (e.g. Berret et al., 2008, 2011). Briefly,

$$\tau = \mathcal{M}(\theta)\ddot{\theta} + \mathcal{C}(\theta, \dot{\theta})\dot{\theta} + \mathcal{F}\dot{\theta}, \tag{19}$$

where $\theta = (\theta_1, \theta_2)^\top$ and $\tau = (\tau_1, \tau_2)^\top$ denote the joint angle and torque vectors, respectively. The quantities $\mathcal{M}$, $\mathcal{C}$, and $\mathcal{F}$ are the inertia, the Coriolis/centripetal and the viscosity matrices, respectively.

Smooth torque signals were obtained by controlling the torque derivative in order to emulate the low-pass filter characteristics of muscles, which is a classical assumption done in other studies (e.g. Uno et al., 1989; Nakano et al., 1999) and is as follows:

$$\dot{\tau} = \mathbf{u}, \tag{20}$$

where $\mathbf{u}$ is the control variable.

Anthropometric parameters were taken from Winter's tables (Winter, 1990) and values were adjusted to each participant.

Numerical optimal control In order to solve a free-time optimal control problem, a naive strategy might consist in testing different movement times until the best duration is found, by solving fixed-time problems at each iteration. This approach would be of course computationally intensive and rather inefficient. However it can be seen from the PMP that the free time setting just introduces an additional equation (see Eq. 3). Indeed, the Hamiltonian has to be zero along the optimal state/costate trajectory, which can be exploited to find the solution without resorting to any trial-error procedure and without additional computational complexity. For resolving general optimal control problems (fixed- or free-time), we relied here a classical method that consists of transforming the OCP into a nonlinear programming (NLP) problem with constraints. More specifically, we used an orthogonal collocation method that is efficiently implemented in Matlab (The MathWorks, Inc., Natick, Massachusetts, United States), called $\mathcal{GPOPS}$ (Benson et al., 2006; Garg et al., 2009; Rao et al., 2010). The NLP problem was solved by means of the well-established numerical software SNOPT (Gill et al., 2005). We checked *a posteriori* that the Hamiltonian was either zero (in free-time) or constant (in fixed-time) along the resulting optimal state/costate trajectories. For LQ or LQG problems, the simulations were also performed by using the closed-form solutions of the problem, and we checked that both numerical methods yielded the same trajectories and time costs.

Experimental procedures

Main experimental task *Participants.* Twenty right-handed healthy adults, without neuromuscular diseases and with normal or corrected-to-normal vision, participated in this study (10 males, age: 30 ± 5 years, mass: 67 ± 12 kg). Written informed consent was obtained from each participant in the study as required by the Helsinki declaration and the EA 4042 local Ethics Committee. All the participants were naive to the purpose of the experiment.

Motor task. Participants performed visually-guided single-joint movements (rotation around the shoulder joint) in the horizontal plane. Participants stood in front of a large vertical screen where spotlight targets (3 cm diameter) were displayed. We tested 10 amplitudes ranging from 5° to 95° and 10 repetitions were recorded for each amplitude (5 in the rightward and 5 in the leftward direction). Participants started with the arm orthogonal to the screen. A sequence of 20 targets was then displayed (a new target appeared every 3.5 seconds). The starting arm position was varied across the experiment and coincided with the previous target location. For each participant, 5

blocks of 20 movements were recorded, for a total of 100 movements per participant. The sequence of trials was fully randomized within each block, however ensuring that one leftward and one rightward movement of each amplitude was presented within each sequence of 20 trials.

Instructions. Participants were instructed to move at a spontaneous, comfortable speed. They were asked to point toward the displayed spotlight target without trying to touch it because it would otherwise induce trunk movements. They were required to keep the arm fully extended, to minimize trunk rotations while performing the task and to perform one-shot movements without correction. They also had to support the weight of their arm by themselves and try to move in a transverse plane as much as possible. A familiarization phase consisting of 20 trials was performed prior to the recording of the 100 trials. Participants were allowed to rest and relax their arm every 20 trials for several minutes (a "pause" word appeared on the screen for that purpose).

Data Collection and Processing *Materials.* Arm and head motion of the dominant arm were recorded by means of a motion capture system (Optitrack device). Ten cameras were used to capture the movement of five retro reflective markers (14 mm in diameter), placed at well-defined anatomical locations on the moving arm and head (acromial process, humeral lateral condyle, apex of the index finger, left and right sides of the frontal bone).

Data analysis. All the analyses were performed with custom software written in Matlab (Mathworks, Natick, MA) from the recorded 3D positions of the markers (sampling frequency of 250 Hz). Recorded signals were low-pass filtered using a digital fifth-order Butterworth filter at a cut-off frequency of 10 Hz (Matlab filtfilt function). Velocity profiles were computed via numerical differentiation.

The movement onset time was defined as the instant at which the linear tangential velocity of the fingertip exceeded 5% of its peak and the end of movement as the point at which the same velocity dropped below the 5% threshold. Movement duration was first inferred from those values as it is a traditional method. However, given the purpose of the paper, a second method was used where the velocity profiles were fitted to minimum jerk velocity profiles (Flash and Hogan, 1985) in order to infer the movement duration (in the spirit of Botzer and Karniel, 2009). To do so, a global optimization using Matlab's *patternsearch* function was carried out and the parameters of the best-fitting minimum jerk problem were identified. On average, both methods gave similar

results but the second method revealed itself to be more robust to variations of motion amplitude and noise, and was therefore much more reliable in general. This second method was thus retained. Finally, the angular position of the arm was also evaluated from the shoulder and finger markers and using some trigonometry. The error between the theoretical amplitude and the one inferred from our recorded data was 2.0 ± 0.9 degrees on average across participants. This indicated that the participants fulfilled the task correctly by respecting the required amplitudes and that the stimuli displayed on the screen were correctly calibrated.

Additional experiment Four right-handed healthy adults were tested during an additional experiment addressing the case of small amplitudes/times (age: 32 ± 2 years, mass: 66 ± 4 kg). The same protocol than the main experiment was used except that a smaller target width (0.3 cm instead of 3 cm) and a smaller fingertip marker (9 mm instead of 14 mm) were used since specific focus was put on small motion extents. A first block of 100 movements was recorded for amplitudes ranging from 5° to 95° as for the main task. A second block of 100 movements was recorded for amplitudes ranging from 1° to 5°. The order of the two blocks was counterbalanced across subjects. The smallest tested movements were about 12 mm here, which was still much above the accuracy of the motion capture system (<0.5 mm from calibration). A total of 200 reaches per participant was thus available in this data set, ranging from very small to very large motion extents.

Results

Illustration for saccades

We first considered the case of saccades to test the effectiveness and relevance of our methodology. In a series of paper (Shadmehr et al., 2010; Shadmehr, 2010), Shadmehr and colleagues already investigated in great details the extent to which the CoT theory could account for the duration of horizontal saccades and their variations with respect to the reward assigned by the brain to the task. They concluded that a hyperbolic CoT replicated better the amplitude-duration relationship than linear or quadratic alternatives. However, "the timescales [were] too short to allow [them] to dissociate between hyperbolic and exponential temporal discount functions" (Shadmehr, 2010). Considering intertrial intervals, they nevertheless concluded later that an exponential cost of time

was inconsistent with the data (Haith et al., 2012). Here we relied upon the same experimental data (reported in Collewijn et al., 1988) and we used the same linear model of the oculomotor plant and trajectory cost (i.e. "effort" cost, quadratic in the control variable) to identify the CoT underlying human saccades (as described also in Harris and Wolpert, 1998; Tanaka et al., 2006; Kardamakis and Moschovakis, 2009; Shadmehr et al., 2010).

In contrast with Shadmehr et al. (2010), the present investigation did not require guessing a parametric form of the way the brain values the passage of time. Instead, we directly computed the infinitesimal CoT at different times based upon either the affine amplitude-duration relationship or the true data points exhibiting a growth larger than linear for ample saccades reported in Collewijn et al. (1988). We identified two similar yet slightly different time costs (black and gray traces in Fig. 3a,b). First, our results confirmed previous findings by showing that the CoT underlying saccades had undeniably a concave shape, thereby ruling out other time costs such as linear and convex ones. When fitting the infinitesimal CoT on the range of actual durations (i.e., [36.5-295 ms]), we found that the root mean square error (RMSE) for the exponential fit was only slightly better than the hyperbolic one (0.010 vs 0.015). Fitting g on a larger interval would be possible but the result would then depend heavily on the extrapolation method. Second, the reward assigned to the task could be estimated from the asymptotic value of G, without the need for a parametric guess of the CoT. However, when a parametric model of CoT was available (not based on blind intuition though but on the shape of g), the reward could be estimated via the parameter α (see Fig. 3c.), without resorting to any trial/error adjustments. In that case, the reward discount rate could be evaluated as well via the parameter β (Fig. 3c). Third, when modeling signal-dependent noise and thus switching to the stochastic context like in Shadmehr et al. (2010) or Harris and Wolpert (1998), the identified values of the infinitesimal CoT were found to be larger than for the deterministic case (Fig. 3d). This increase of the CoT was mainly a consequence of the larger expected trajectory cost due to the effects of sensorimotor noise. As the level of noise increased, the CoT approached a hyperbolic rather than an exponential discounting of reward (see RMSE in Fig. 3d). Since our methodology provides an exact solution to the identification problem, the original amplitude-duration relationship was perfectly recovered from free-time OC simulations as expected (using the inferred CoT, see Fig. 3e). The associated optimal saccade velocity profiles are reported in Figure 3f. The important result was that in all cases the instantaneous CoT was

a decreasing function of time such that G (the integrated CoT) always had a clear concave shape. Importantly, our theoretical results for single-input LQ problems ensure that this inferred CoT is not just an artifact but does possess such well-identified characteristics. Our findings thus rule out non-concave time costs for the control of saccades and confirm the hyperbolic shape of the CoT (Shadmehr et al., 2010; Haith et al., 2012).

Application to reaching

We then investigated in depth the CoT underlying the control of reaching. As far as we know, inferring the CoT for reaching has not been attempted yet. To start, we considered the case of a single-joint arm moving in the horizontal plane whose linear musculoskeletal dynamics was described in several studies (Hogan, 1984; Tanaka et al., 2006; Berret et al., 2008; Gaveau et al., 2014) and detailed in Materials and Methods. We relied on the effort required to generate a reach as trajectory cost, i.e. a quadratic cost in the control variable. Incidentally, it coincided here with the torque change (Uno et al., 1989), a prominent model known to replicate well the bell shape of velocity profiles during single-joint rotations (Gottlieb et al., 1989). Note that this problem also falls within the single-input LQ class for which we have powerful mathematical results.

In order to uncover the CoT underlying such self-paced arm pointing movements, experimental data were first gathered and parameters such as movement extent and duration were computed. For all the participants, the relationship between amplitude and duration was approximately linear, in agreement with the existing literature (Brown et al., 1990; Gentili et al., 2007). Figure 4 illustrates this relationship for 4 participants separately, which was quantified via linear regressions based on all the 100 trials recorded and depicted as single gray dots ($R^2 > 0.75$ for all the 20 participants). Even though the task was fairly simple, inter-individual differences were noticeable. In particular, both the slope and the intercept of the regression line varied across participants.

For each single movement, information about its amplitude and duration was eventually used to estimate the value of g for that time. Figure 5a displays the values of $g(t)$ when sampled on such a single-trial basis (gray dots). Each dot in this figure corresponds to a dot in Figure 4. Because of sensorimotor noise and inter-trial fluctuations (clearly visible in Fig. 4), the identified CoT was itself quite noisy when inferred from single trials. To get a more reliable and noiseless estimation of g, we then exploited the affine law linking amplitude and MT to infer the CoT on the time interval

where actual measurements were taken in practice (black traces in Fig. 5a). The time interval containing the empirical reach durations is emphasized with filled circles added on the CoT traces. Its boundary values, corresponding respectively to the faster and slower motions observed during the experimental session, varied between participants as they were instructed to move at their own pace. Outside of this interval, the shape of the infinitesimal CoT was less reliable as it was either computed based upon regression equations (yielding predicted MTs for untested amplitudes) or just extrapolated by assuming that the initial and asymptotic values of g were both zero (respectively plotted as solid and dotted lines in Fig. 5a). In Figure 5b, the corresponding integral CoTs, G, are depicted. The CoT exhibited a striking logistic shape for all the 20 participants, irrespective of their spontaneous pace and anthropometric characteristics. Importantly, this conclusion could be drawn from the most reliable part of the curve, that is, over the range of MTs that were recorded in practice (and without resorting to any sort of extrapolation). This non-trivial shape would have been hardly predictable using a parametric approach as the CoT is neither concave or convex but sigmoidal (at least on the time interval of interest, i.e. from about 500 ms to 1500 ms). As a by-product, these results show that for reaching the CoT is not purely hyperbolic or exponential.

Verification through direct optimal control simulations in free time Before going further, we performed free-time optimal control simulations by using the identified g to verify numerically the effectiveness of our methodology. As expected, the fitted relationship between amplitude and duration was exactly the same in simulation and experiment (see Fig. 6a). This was expected given that the method seeks the CoT that will exactly account for the original amplitude-duration relationship. The velocity profiles were also bell-shaped and simply scaled with speed in accordance with well-known experimental observations (see Fig. 6b and Gottlieb et al., 1989 for example). It is worth noting that solving free-time optimal control problems is not computationally more intensive than solving their fixed-time counterparts. In particular, resolving a free-time problem does not necessarily require a trial/error procedure as it is often assumed or done in the literature (e.g. Shadmehr et al., 2010; Rigoux and Guigon, 2012 but see Hoff, 1994; van Beers, 2008).

Consistency with respect to the fitting of the amplitude-duration relationship In the preceding paragraphs, we assumed an affine relationship between amplitude and duration with the

consequence of having non-zero MT at zero amplitude. Actually, other functions such as concave ones may fit the data equally well, with or without bias at the origin. To investigate the time cost consistency with respect to data fitting, we collected additional data involving reaching movements of amplitudes less than 5° since non-linearity could mainly arise for small times/amplitudes (see additional experiment in Materials and Methods). While the shape of the amplitude-duration relationship was very close to linear for reaches greater than 5°, these data allowed us to refine our analysis near the origin. Figure 7a shows the amplitude-duration relationship for amplitudes ranging from 1° to 5° (100 reaches) and 5° to 95° (100 reaches) for horizontal reaches with a fully-extended arm. The data revealed that the relationship was more concave than linear in the vicinity of small amplitudes, even though an affine fit still performed relatively well overall ($R^2 > 0.85$). Although it was not possible to decide between concave fits with or without intercept from the R^2 values (light vs. dark gray traces), the CoT appeared to be sigmoidal over the range of empirical MTs in any case (i.e., between approximately 300 ms and 1500 ms, see Fig. 7b). For all the participants, a similar bell shape for the function g starting from times about 300 ms was found. The only noticeable difference between fits with or without intercept occurred for small times (corresponding to very short motion extents because the fitted curve was very steep near the origin) where g was found to increase as MT tended to zero in the latter case (dark gray traces). This finding is however hypothetical since it is not grounded on real data as we did not measure reaches with duration smaller than 300 ms in this experiment. Nevertheless, the integral CoT, G, exhibited a clear sigmoidal shape over the range of actual reach durations. The latter was fully characterized up to a constant shift corresponding to the (unknown) values of $g(t)$ for $t < 300$ ms. These values however do not affect the ability of the identified G to predict the correct amplitude-duration relationship for times of practical interest for reaching.

Consistency with respect to trajectory costs, motor noise and muscle dynamics Before concluding about the sigmoidal shape of the CoT on the time interval of empirical reach durations, we further tested its consistency with respect to changes of trajectory cost, motor noise and muscle dynamics (Fig. 8). Thus far, we relied on a simple quadratic effort cost (i.e. the torque change here). However, since a variety of trajectory costs could replicate the bell-shaped velocity profiles of such single-joint arm rotations, we assessed the robustness of the CoT with respect to other plausible

trajectory costs. We considered the minimum jerk (Flash and Hogan, 1985), torque (Nelson, 1983), absolute work (Berret et al., 2008) or acceleration (Ben-Itzhak and Karniel, 2008) models. The extracted time costs are reported in Figure 8a. This numerical analysis confirmed the consistency of the shape of g with respect to l. Note that a candidate trajectory cost must at least replicate the arm trajectories when MT is known and, therefore, this condition was ensured before extracting the CoT (i.e. the smooth bell shape of velocity profiles). We then tested the effect of motor noise as it is known play on role on arm movement control (Todorov and Jordan, 2002; van Beers et al., 2004). Results are reported in Figure 8b. Signal-dependent noise did not change the overall sigmoidal shape of the CoT but just tended to increase the CoT magnitude.

In the preceding simulations, a simple model of the arm mechanics and muscle dynamics was used. We tested the extent to which the sigmoidal shape of the CoT was dependent on the model of the musculoskeletal system. To this aim, we considered that the arm was driven by two antagonist muscles modeled as 2nd-order low-pass filters with specific time constants. When still minimizing the torque change (but controlling the input to muscles), we found that the CoT remained very precisely the same whatever the time constants of the muscles (we tested from 0.03 s to 0.15 s). However, when the cost function was modified such that a quadratic cost (i.e. a "neural effort") was minimized instead of the torque change, the CoT depended on the muscle characteristics (Fig. 8c). Importantly, although the precise values of the CoT obviously obviously varied as a function of the muscle time constants, its sigmoidal shape was nevertheless preserved. Its rate of change g was robustly characterized by a steep increase until a peak was reached and followed by a slower decay on the time interval where the CoT could be reliably sampled. These tests further show that the neural effort cost, yet another plausible candidate trajectory cost known to predict smooth bell-shaped velocity profiles (e.g. Guigon et al., 2007; Liu and Todorov, 2007), also yields a sigmoidal CoT (especially for $\sigma \approx 0.04$ s where velocity profiles are close to minimum jerk ones, Richardson and Flash, 2002).

Consistency with respect to multijoint dynamics, different initial/final states and speed instructions Experimentally, different initial/final arm configurations may lead to the same MT. From the identification standpoint, this may yield several possible values for $g(t)$ where the CoT theory would only assume one. For the above planar single-joint arm movements, this was however

not a problem. At the theoretical level, the invariance of g with respect to initial/final states follows trivially from the fact that the optimal solutions only depend on the amplitude a, and not specifically on the states $\mathbf{x}^0$ and $\mathbf{x}^f$. This fact is consistent with experimental observations where MT was found to mostly depend on a and neither specifically on the relative starting hand position in the workspace nor on the leftward/rightward direction of the motion for such horizontal single-joint rotations (Gentili et al., 2007). In general, however, reaching involves a non-linear plant and several degrees of freedom. The situation is thus more complex for a multi-articulated arm but our methodology allows for sampling $g(t)$ in the exact same way and effortlessly as soon as the associated fixed-time OC problem can be (numerically) resolved.

For a two-joint arm moving in the horizontal plane, the role of the state/control cost l is reinforced because the hand path depends on it. Attempting to predict MT would be completely irrelevant if the hand path cannot be accurately replicated when MT is known at least. Here we still considered the minimum torque change model for the trajectory cost because it is known to account well for both the quasi-straight path and bell-shaped velocity profile of the hand during planar point-to-point arm movements (Uno et al., 1989). We used the data of Young et al. (2009) and the reported equations for the amplitude-duration relationship to recover the CoT underlying such multijoint arm reaches. The task is illustrated in Figure 9a and the amplitude-duration mappings at natural and quick speeds are given in Figure 9b. The CoT g was then sampled pointwise from these linear fits using our methodology (Figure 9c,d) and it was extrapolated out of the range of real measurements. Since Young and colleagues did not report any leftward/rightward difference regarding MTs, we used the same affine relationship for leftward and rightward movements performed at natural speed. The infinitesimal CoT differed slightly depended on movement direction. Additionally, varying the starting posture of the arm also led to different instantaneous time costs for leftward hand displacements (Fig. 9c, left). These CoT variations are due to the arm's inertial anisotropy (Gordon et al., 1994) and the fact that the torque change is a dynamic cost. For the sake of completeness, we also identified the CoT for the hand jerk cost (Flash and Hogan, 1985) and this revealed that the CoT remained exactly the same irrespective of movement direction and initial posture (data not shown).

The above variations of CoT were relatively small compared to what was observed during a change of speed instruction (Fig. 9c, right). When asking the participants to move quickly, one can

observe that both the slope and the intercept of the affine relationship change drastically (Fig. 9b). This resulted in very different values for the function g. Not only the peak was multiplied by ~30x but MT was also penalized slightly earlier (onset time of g). This agrees with our theoretical analysis where we showed that globally rescaling the CoT cannot lead to a change in both slope and intercept of the amplitude/duration relationship.

Hitherto, we have identified the CoT from experimental data. We also assessed the extent to which the CoT identified in a given task could be used to predict MTs in another task. In Figure 9e, a center-out reaching task was thus considered as it has been the subject of extensive investigations (Gordon et al., 1994; van Beers et al., 2004). Using the CoT identified at natural speed for the central starting position and using the relationship between duration and extent for medio-lateral movements, it was possible to predict the directional dependence of MT during a center-out task (Fig. 9f). The same could be observed at quick speed using the corresponding CoT. We eventually identified the values of g from the data of van Beers et al. (2004) at quick speed for the 16 movement directions and plotted the corresponding values in Figure 9d (black dots), where the CoT had been identified from the data of Young et al. (2009). The location of the dots around the CoT uncovered at quick speed confirmed the relative directional invariance of the infinitesimal CoT.

Above all, these findings reveal the robust sigmoidal shape of the CoT with respect to modifications of data fitting, trajectory cost, system dynamics, initial/final states and task instructions, even though its precise values necessarily depend on all the details of modeling. These present results rule out purely concave or convex time costs for reaching whereas its concavity (e.g. hyperbolic) was confirmed for saccades.

Discussion

We addressed the issue of how the brain selects the duration of movement. We relied upon the "cost of time" theory that finds its roots into the brain's tendency to discount the actual value of future reward (Shadmehr, 2010; Shadmehr et al., 2010; Shadmehr and Mussa-Ivaldi, 2012). It asserts that the time elapsed until action completion entails a cost, thereby making slow moves non-optimal. By means of a thorough mathematical analysis, we developed a complete, robust and fairly general methodology to automatically identify the CoT from basic experimental data by

sampling its infinitesimal values without making prior assumption about its hypothetical nature. An application to reaching revealed that an overall sigmoidal CoT underlies the planning of arm movement. In the following, we discuss the possible origin, rationale and implication of such a CoT.

The CoT is a generic but appealing concept that has been recently popularized by Shadmehr's group. In particular, it is straightforward to interpret the CoT as a temporal discounting of reward by obvious mathematical equivalence (Shadmehr, 2010), thereby creating a link between the reward system and motor control (Wolpert and Landy, 2012). In most studies investigating this link, however, humans or animals were motivated via incentives and gratification was often made explicit to them (e.g. food, monetary, token or sensory information; Battaglia and Schrater, 2007; Niv et al., 2007; Dean et al., 2007; Hudson et al., 2008; Guitart-Masip et al., 2012). For the neural control of movement, the actual reward assigned by the brain to any given action is rather an intrinsic and hidden quantity. For example, consider reaching for an ordinary object. It is difficult to predict what reward the brain actually assigns to such a task. Furthermore, if the object is a glass of water, the theory predicts that movement vigor should increase when the individual is thirsty, but exactly knowing the extent to which the intrinsic value of reward is modified seems tricky. Interestingly, the present methodology offers the possibility to robustly estimate and differentially compare such intrinsic rewards. Indeed, the reward attributed to the task can be inferred from the asymptotic value of G. Then it would be possible to record movements under different conditions and analyze how reward varies, all other things being equal.

However, time could be costly in motor control for reasons other than just a temporal discounting of reward. In particular, slow moves might also be undesirable because they monopolize a significant amount of neural resources due to the inevitable processing of the sensorimotor flow of information by the central nervous system, and neural processing of information is metabolically expensive (Laughlin, 2001; Lennie, 2003; Howarth et al., 2012). The attentional cost associated with slow-paced movement control could also augment and interfere with the proficiency of achieving other tasks, as suggested by dual-task experimental paradigms (Krampe et al., 2010). Hence the signalling and processing of a quantity of information that likely grows with time during sensorimotor control might induce such a high cognitive load and large neural metabolic energy demand. Interestingly, a growing body of evidence also suggests that parkinsonian motor symptoms may be linked to neural metabolic deficits (Amano et al., 2014). The logistic shape identified for reaching may thus

reflect this dual nature of the CoT. Indeed, its sole interpretation as a temporal discounting of reward might be insufficient to account for the inflexion of G and this secondary interpretation as a metabolic cost related to the engagement into a sensorimotor task may be speculated. This is what a recent study tends to confirm for planar arm reaching (Huang and Ahmed, 2012). A surprising constant metabolic cost related to the simple fact of being engaged in a reaching task was discovered. Therefore, the CoT we inferred here may also reflect this part of the metabolic cost that does not depend on the state of the musculoskeletal apparatus but on the underlying neural mechanisms that are involved in on-line movement control and that presumably depend on motion duration. An interesting topic for future research would consist in trying to disentangle these putative reward and metabolic parts of the CoT. This may be possible by tuning the task parameters adequately and by analyzing how the CoT varies accordingly (e.g. Klein-Flügge et al., 2015 where evidence for an inverse sigmoidal effort-related discounting of reward was reported).

It could still be argued that the CoT is just a mathematical epiphenomenon required to compensate for some missing features of the trajectory cost l. Besides a temporal discounting of reward, motion duration could indirectly reflect onto other costs such as accuracy via constant noise accumulation (van Beers et al., 2004; van Beers, 2008) or metabolic energy expenditure (Alexander, 1997; Huang and Ahmed, 2012). For eye movement, it has been suggested that slow saccades are detrimental to movement accuracy as a consequence of constant noise which accumulates over time (van Beers, 2008). This would lead to a trajectory cost l exhibiting a "U" shape as a function of MT since endpoint variability would not only increase at large speeds as a consequence of signal-dependent noise (Harris and Wolpert, 1998) but also at small speeds. This would yield optimal MTs without the need to resort to an explicit CoT. While plausible for saccades and feedforward control, this view is unsustainable for visually-guided reaching since sensory feedback is processed on-line to correct for motor errors. For reaching, this is rather the metabolic energy expenditure that may critically increase during slow-paced movement. Gravity could then be a predominant factor for MT formation. If so, MT should drastically increase in weightless conditions but it is known that cosmonauts increase only slightly MT after adaptation to weightlessness (Papaxanthis et al., 1998). Moreover, criteria counting the effort required to counteract gravitational torques usually fail to predict realistic arm trajectories even in fixed time (Berret et al., 2011) and models minimizing the metabolic cost of muscle contraction had to resort to parameter tuning to be able to replicate

empirical trajectories for a wide range of speeds, making it difficult to estimate metabolic energy expenditure across speeds (Alexander, 1997). Interestingly, the metabolic cost per arm movement reported in Huang and Ahmed (2012) exhibited a "U" shape with respect to motion speed, as for the case of walking (Ralston, 1958; Alexander, 1991), but with the difference that the subject's arm was fully supported by a robotic device. This observation further supports that the cost measured at rest or very low speed was mostly due to a metabolic energy expenditure associated with the fact of being actively involved into a reaching task (which would therefore depend on time only) rather than with gravity itself or with the complex energetics of the muscle contraction (e.g. Kushmerick, 1983). Unfortunately, it seems rather difficult from current knowledge to dissociate the part of metabolic energy expenditure that does depend on arm trajectories (and thus contributes to shaping them) from the one that does not depend on them but just on the passage of time.

Whatever the exact nature of this time cost is (temporal discounting of reward, neural/attentional metabolic cost or even something else, e.g. working memory decay), the present methodology allowed us to reliably infer its global shape for times corresponding to empirical reach and saccade durations. For reaching, the time interval approximately ranged from 300 ms to 1500 ms whereas for saccades it ranged from about 30 ms to 300 ms. Since the two CoTs exhibited different shapes, the question of their link can be posed. On the one hand, it seems conceivable to argue that the time costs underlying saccades and reaches are just independent because the two systems and their functions are different. On the other hand, it also seems possible to assume a single time cost. Indeed, the two CoTs may be juxtaposed as they only slightly overlap on their respective time intervals. Support for this idea was found when assuming zero duration at zero amplitude: in that case, the CoT for reaching had a shape reminiscent of the one for saccades in the vicinity of small times (typically $<$200 ms). However, such a similarity was only hypothetical because reach data were unavailable on such a range and whether or not the amplitude-duration relationship passed through the origin was undecidable on the present data. Actually, a regression intercept is often assumed in the literature (e.g. Fitts' law and Young et al., 2009) but its relatively large value for reaching contrasts with the case of saccades. This discrepancy could be due to the role of feedback (which may require MT extensions to improve terminal reach accuracy) or to the incompressible duration of skeletal muscles activation/deactivation (which differ from extraocular muscles) making it impossible to attain zero duration as amplitude tends to zero. We nevertheless showed that this

point was not critical regarding our main conclusion about the sigmoidal shape of the time cost over the range of real movement times. Further work seems however required to assess the extent to which temporal costs inferred from different motor tasks and systems (e.g. eye, wrist, arm or whole-body) have compatible shapes on overlapping time intervals and whether or not they can be related.

Although the neural substrate of such a CoT is unknown, one may speculate that it is linked with the dopaminergic system since movement vigor is significantly altered in Parkinson's disease (Mazzoni et al., 2007). Indeed, dopamine plays a central role in the BG, a region related to the control of movement gain (extent and speed) and cost functions (Shadmehr and Krakauer, 2008; Schmidt et al., 2008; Desmurget and Turner, 2008; Turner and Desmurget, 2010; Amano et al., 2014). More specifically, a number of studies showed that transient inactivation or ablation of the globus pallidus internus, the principal output nucleus for the BG, reduced movement velocity and acceleration without altering hand path, reach accuracy or movement sequencing (Anderson and Horak, 1985; Mink and Thach, 1991; Inase et al., 1996; Desmurget and Turner, 2010). In addition, the striatum could play a critical role on MT since it has been shown to be related to the temporal discounting of reward (Kobayashi and Schultz, 2008). Therefore, valuation of MT could be partly encoded in that area as it is a fundamental center to gauge effort-benefit situations (Croxson et al., 2009) even in the absence of extrinsic reward (Schouppe et al., 2014). It would be the purpose of future studies to investigate the neural correlates of time cost variations arising from task modifications. The present robust and automated methodology may offer such an opportunity by providing the genuine shape of CoT in neural motor control.

References

Alexander RM (1991) Energy-saving mechanisms in walking and running. *J Exp Biol* 160:55–69.

Alexander RM (1997) A minimum energy cost hypothesis for human arm trajectories. *Biol Cybern* 76:97–105.

Amano S, Kegelmeyer D, Hong SL (2014) Rethinking energy in parkinsonian motor symptoms: a potential role for neural metabolic deficits. *Front Syst Neurosci* 8:242.

Anderson ME, Horak FB (1985) Influence of the globus pallidus on arm movements in monkeys. iii. timing of movement-related information. *J Neurophysiol* 54:433–448.

Baraduc P, Thobois S, Gan J, Broussolle E, Desmurget M (2013) A common optimization principle for motor execution in healthy subjects and parkinsonian patients. *J Neurosci* 33:665–677.

Battaglia PW, Schrater PR (2007) Humans trade off viewing time and movement duration to improve visuomotor accuracy in a fast reaching task. *J Neurosci* 27:6984–6994.

Ben-Itzhak S, Karniel A (2008) Minimum acceleration criterion with constraints implies bang-bang control as an underlying principle for optimal trajectories of arm reaching movements. *Neural Comput* 20(3):779–812.

Benson DA, Huntington GT, Thorvaldsen TP, Rao AV (2006) Direct trajectory optimization and costate estimation via an orthogonal collocation method. *Journal of Guidance, Control, and Dynamics* 29:1435–1440.

Berardelli A, Rothwell JC, Thompson PD, Hallett M (2001) Pathophysiology of bradykinesia in parkinson's disease. *Brain* 124:2131–2146.

Berret B, Chiovetto E, Nori F, Pozzo T (2011) Evidence for composite cost functions in arm movement planning: an inverse optimal control approach. *PLoS Comput Biol* 7:e1002183.

Berret B, Darlot C, Jean F, Pozzo T, Papaxanthis C, Gauthier JP (2008) The inactivation principle: mathematical solutions minimizing the absolute work and biological implications for the planning of arm movements. *PLoS Comput Biol* 4:e1000194.

Botzer L, Karniel A (2009) A simple and accurate onset detection method for a measured bell-shaped speed profile. *Front Neurosci* 3:61.

Brücke C, Huebl J, Schönecker T, Neumann WJ, Yarrow K, Kupsch A, Blahak C, Lütjens G, Brown P, Krauss JK, Schneider GH, Kühn AA (2012) Scaling of movement is related to pallidal gamma oscillations in patients with dystonia. *J Neurosci* 32:1008–1019.

Brown SH, Hefter H, Mertens M, Freund HJ (1990) Disturbances in human arm movement trajectory due to mild cerebellar dysfunction. *J Neurol Neurosurg Psychiatry* 53:306–313.

Collewijn H, Erkelens CJ, Steinman RM (1988) Binocular co-ordination of human horizontal saccadic eye movements. *J Physiol* 404:157–182.

Croxson PL, Walton ME, O'Reilly JX, Behrens TEJ, Rushworth MFS (2009) Effort-based cost-benefit valuation and the human brain. *J Neurosci* 29:4531–4541.

Dean M, Wu SW, Maloney LT (2007) Trading off speed and accuracy in rapid, goal-directed movements. *J Vis* 7:10.1–1012.

Desmurget M, Turner RS (2008) Testing basal ganglia motor functions through reversible inactivations in the posterior internal globus pallidus. *J Neurophysiol* 99:1057–1076.

Desmurget M, Turner RS (2010) Motor sequences and the basal ganglia: kinematics, not habits. *J Neurosci* 30:7685–7690.

Engelbrecht S (2001) Minimum principles in motor control. *J Math Psychol* 45:497–542.

Ferrante A, Marro G, Ntogramatzidis L (2005) A parametrization of the solutions of the finite-horizon lq problem with general cost and boundary conditions. *Automatica* 41:1359–1366.

Fitts PM (1954) The information capacity of the human motor system in controlling the amplitude of movement. *J Exp Psychol* 47:381–391.

Flash T, Hogan N (1985) The coordination of arm movements: an experimentally confirmed mathematical model. *J Neurosci* 5:1688–1703.

Garg D, Patterson MA, Hager WW, Rao AV, Benson DA, Huntington GT (2009) A unified framework for the numerical solution of optimal control problems using pseudospectral methods (to appear). *Automatica* .

Gaveau J, Berret B, Demougeot L, Fadiga L, Pozzo T, Papaxanthis C (2014) Energy-related optimal control accounts for gravitational load: comparing shoulder, elbow, and wrist rotations. *J Neurophysiol* 111:4–16.

Gentili R, Cahouet V, Papaxanthis C (2007) Motor planning of arm movements is direction-dependent in the gravity field. *Neuroscience* 145:20–32.

Gill PE, Murray W, Saunders MA (2005) Snopt: An sqp algorithm for large-scale constrained optimization. *SIAM Review* 47:99–131.

Gordon J, Ghilardi MF, Cooper SE, Ghez C (1994) Accuracy of planar reaching movements. ii. systematic extent errors resulting from inertial anisotropy. *Exp Brain Res* 99:112–130.

Gottlieb GL, Corcos DM, Agarwal GC (1989) Strategies for the control of voluntary movements with one mechanical degree of freedom. *Behavioral and Brain Sciences* 12:189–250.

Guigon E, Baraduc P, Desmurget M (2007) Computational motor control: redundancy and invariance. *J Neurophysiol* 97:331–347.

Guitart-Masip M, Chowdhury R, Sharot T, Dayan P, Duzel E, Dolan RJ (2012) Action controls dopaminergic enhancement of reward representations. *Proc Natl Acad Sci U S A* 109:7511–7516.

Haith AM, Reppert TR, Shadmehr R (2012) Evidence for hyperbolic temporal discounting of reward in control of movements. *J Neurosci* 32:11727–11736.

Harris CM, Wolpert DM (1998) Signal-dependent noise determines motor planning. *Nature* 394:780–784.

Harris CM, Wolpert DM (2006) The main sequence of saccades optimizes speed-accuracy trade-off. *Biol Cybern* 95:21–29.

Hoff B (1994) A model of duration in normal and perturbed reaching movement. *Biological Cybernetics* pp. 481–488.

Hogan N (1984) An organizing principle for a class of voluntary movements. *J Neurosci* 4:2745–2754.

Howarth C, Gleeson P, Attwell D (2012) Updated energy budgets for neural computation in the neocortex and cerebellum. *J Cereb Blood Flow Metab* 32:1222–1232.

Huang HJ, Ahmed AA (2012) Is there a reaching speed that minimizes metabolic cost? In *Proceedings of the Translational and Computational Motor Control 2012*, New Orleans, LA.

Hudson TE, Maloney LT, Landy MS (2008) Optimal compensation for temporal uncertainty in movement planning. *PLoS Comput Biol* 4:e1000130.

Huh D, Todorov E, Sejnowski TJ (2010) Infinite horizon optimal control framework for goal directed movements In *Society for Neuroscience Annual Meeting, Online: Program No. 492.411.*

Inase M, Buford JA, Anderson ME (1996) Changes in the control of arm position, movement, and thalamic discharge during local inactivation in the globus pallidus of the monkey. *J Neurophysiol* 75:1087–1104.

Kappen HJ (2011) Optimal control theory and the linear bellman equation In Barber D, Cemgil AT, Chiappa S, editors, *Bayesian Time Series Models*, pp. 363–387. Cambridge University Press Cambridge Books Online.

Kardamakis AA, Moschovakis AK (2009) Optimal control of gaze shifts. *J Neurosci* 29:7723–7730.

Kirk DE (1970) *Optimal control theory: An Introduction.* Prentice-Hall, New Jersey.

Klein-Flügge MC, Kennerley SW, Saraiva AC, Penny WD, Bestmann S (2015) Behavioral modeling of human choices reveals dissociable effects of physical effort and temporal delay on reward devaluation. *PLoS Comput Biol* 11:e1004116.

Kobayashi S, Schultz W (2008) Influence of reward delays on responses of dopamine neurons. *J Neurosci* 28:7837–7846.

Krampe RT, Doumas M, Lavrysen A, Rapp M (2010) The costs of taking it slowly: fast and slow movement timing in older age. *Psychol Aging* 25:980–990.

Kushmerick MJ (1983) Energetics of muscle contraction. *Comprehensive Physiology* .

Laughlin SB (2001) Energy as a constraint on the coding and processing of sensory information. *Curr Opin Neurobiol* 11:475–480.

Lee EB, Markus L (1967) *Foundations of Optimal Control Theory.* John Wiley.

Lennie P (2003) The cost of cortical computation. *Curr Biol* 13:493–497.

Liu D, Todorov E (2007) Evidence for the flexible sensorimotor strategies predicted by optimal feedback control. *J Neurosci* 27:9354–9368.

Mazzoni P, Hristova A, Krakauer JW (2007) Why don't we move faster? parkinson's disease, movement vigor, and implicit motivation. *J Neurosci* 27:7105–7116.

Mink JW, Thach WT (1991) Basal ganglia motor control. iii. pallidal ablation: normal reaction time, muscle cocontraction, and slow movement. *J Neurophysiol* 65:330–351.

Mombaur K, Truong A, Laumond JP (2009) From human to humanoid locomotion - an inverse optimal control approach. *Autonomous Robots* .

Myerson J, Green L (1995) Discounting of delayed rewards: Models of individual choice. *J Exp Anal Behav* 64:263–276.

Nakano E, Imamizu H, Osu R, Uno Y, Gomi H, Yoshioka T, Kawato M (1999) Quantitative examinations of internal representations for arm trajectory planning: minimum commanded torque change model. *J Neurophysiol* 81:2140–2155.

Nelson WL (1983) Physical principles for economies of skilled movements. *Biol Cybern* 46:135–147.

Niv Y, Daw ND, Joel D, Dayan P (2007) Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 191:507–520.

Nori F, Frezza R (2004) Linear optimal control problems and quadratic cost functions estimation In *12th Mediterranean Conference on Control and Automation, MED'04. Kusadasi, Aydin, Turkey.*

Papaxanthis C, Pozzo T, Popov KE, McIntyre J (1998) Hand trajectories of vertical arm movements in one-g and zero-g environments. evidence for a central representation of gravitational force. *Exp Brain Res* 120:496–502.

Pontryagin LS, Boltyanskii VG, Gamkrelidze RV, Mishchenko EF (1964) *The Mathematical Theory of Optimal Processes.* Pergamon Press.

Ralston HJ (1958) Energy-speed relation and optimal speed during level walking. *Int Z Angew Physiol* 17:277–283.

Rao AV, Benson DA, Darby CL, Patterson MA, Francolin C, Sanders I, Huntington GT (2010) Algorithm 902: Gpops, a matlab software for solving multiple-phase optimal control problems using the gauss pseudospectral method. *ACM Transactions on Mathematical Software* 37:1–39.

Richardson MJE, Flash T (2002) Comparing smooth arm movements with the two-thirds power law and the related segmented-control hypothesis. *J Neurosci* 22:8201–8211.

Rigoux L, Guigon E (2012) A model of reward- and effort-based optimal decision making and motor control. *PLoS Comput Biol* 8:e1002716.

Schmidt L, d'Arc BF, Lafargue G, Galanaud D, Czernecki V, Grabli D, Schüpbach M, Hartmann A, Lévy R, Dubois B, Pessiglione M (2008) Disconnecting force from money: effects of basal ganglia damage on incentive motivation. *Brain* 131:1303–1310.

Schouppe N, Demanet J, Boehler CN, Ridderinkhof KR, Notebaert W (2014) The role of the striatum in effort-based decision-making in the absence of reward. *J Neurosci* 34:2148–2154.

Scott SH (2004) Optimal feedback control and the neural basis of volitional motor control. *Nat Rev Neurosci* 5:532–546.

Shadmehr R (2010) Control of movements and temporal discounting of reward. *Curr Opin Neurobiol* 20:726–730.

Shadmehr R, Krakauer JW (2008) A computational neuroanatomy for motor control. *Exp Brain Res* 185:359–381.

Shadmehr R, Mussa-Ivaldi S (2012) *Biological Learning and Control* MIT Press.

Shadmehr R, Orban de Xivry JJ, Xu-Wilson M, Shih TY (2010) Temporal discounting of reward and the cost of time in motor control. *J Neurosci* 30:10507–10516.

Stengel R (1986) *Optimal Control and Estimation* Dover books on advanced mathematics. Dover Publications.

Sutton RS, Barto AG (1998) *Reinforcement Learning: An Introduction* MIT Press.

Tanaka H, Krakauer JW, Qian N (2006) An optimization principle for determining movement duration. *J Neurophysiol* 95:3875–3886.

Todorov E (2004) Optimality principles in sensorimotor control. *Nat Neurosci* 7:907–915.

Todorov E (2006) *Optimal control theory*, chapter 12, pp. 269–298 Bayesian Brain: Probabilistic Approaches to Neural Coding, Doya K (ed).

Todorov E, Jordan MI (2002) Optimal feedback control as a theory of motor coordination. *Nat Neurosci* 5:1226–1235.

Turner RS, Desmurget M (2010) Basal ganglia contributions to motor control: a vigorous tutor. *Curr Opin Neurobiol* 20:704–716.

Uno Y, Kawato M, Suzuki R (1989) Formation and control of optimal trajectory in human multijoint arm movement. minimum torque-change model. *Biol Cybern* 61:89–101.

van Beers RJ (2008) Saccadic eye movements minimize the consequences of motor noise. *PLoS One* 3:e2070.

van Beers RJ, Haggard P, Wolpert DM (2004) The role of execution noise in movement variability. *J Neurophysiol* 91:1050–1063.

Van der Helm FCT, Rozendaal LA (2000) Musculoskeletal systems with intrinsic and proprioceptive feedback In Winters JM, (Eds.) PC, editors, *Biomechanics and neural control of posture and movement*, pp. 164–174. New York: Springer.

Winter D (1990) *Biomechanics and Motor Control of Human Movement.* New York: John Wiley & Sons.

Wolpert DM, Landy MS (2012) Motor control is decision-making. *Curr Opin Neurobiol* 22:996–1003.

Young SJ, Pratt J, Chau T (2009) Target-directed movements at a comfortable pace: movement duration and fitts's law. *J Mot Behav* 41:339–346.

Tables, Legends, Figures

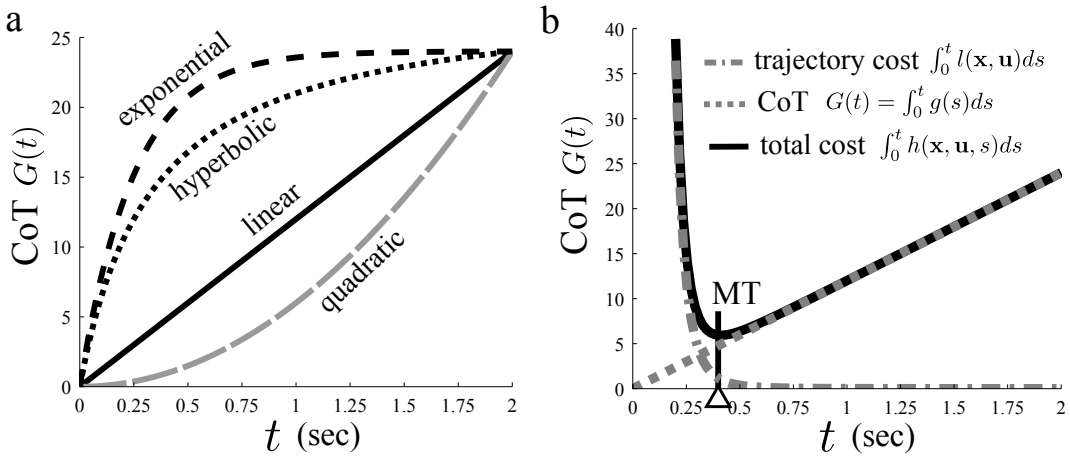

Figure 1: Illustration of the CoT theory. **a**. Previously proposed shapes of the CoT. Linear: $G(t) = \alpha t$ (Hoff, 1994; Liu and Todorov, 2007); Hyperbolic: $G(t) = \alpha(1 - \frac{1}{1+\beta t})$ (Shadmehr, 2010; Shadmehr et al., 2010; Shadmehr and Mussa-Ivaldi, 2012); Exponential: $G(t) = \alpha(1 - \exp(-\beta t))$ (Rigoux and Guigon, 2012); Quadratic: $G(t) = \alpha t^2$ (Shadmehr et al., 2010). **b**. The CoT $G(t) = \int_0^t g(s)ds$ (assumed to be linear here for simplicity) sums with the trajectory cost $\int_0^t l(\mathbf{x}(s), \mathbf{u}(s))ds$ yielding a total cost $\int_0^t h(\mathbf{x}(s), \mathbf{u}(s), s)ds$ that exhibits a "U" shape, the minimum being the optimal MT. It is noteworthy that, without the CoT, infinitely slow movements would be optimal for similar trajectory costs (which includes most of the costs proposed in the literature).

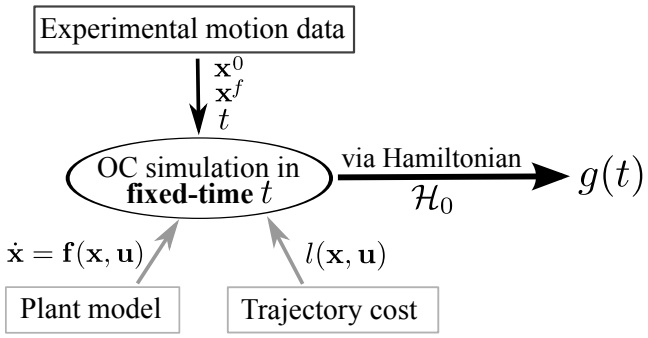

Figure 2: Flowchart of the methodology. First of all, a model of the plant dynamics $\mathbf{f}$ and a trajectory cost l replicating the experimental trajectories with known MT must be available (gray boxes). From these experimental trajectories, the initial and final states of the system can be estimated ($\mathbf{x}^0$ and $\mathbf{x}^f$), as well as the movement duration t (black box). Finally, the corresponding optimal control (OC) problem in fixed-time t must be resolved (central ellipse) in order to evaluate the (constant) Hamiltonian $\mathcal{H}_0$ along the optimal trajectory; this precisely yields the value of the infinitesimal CoT at time t. Note that the partial time derivative of the value function could be used if preferred.

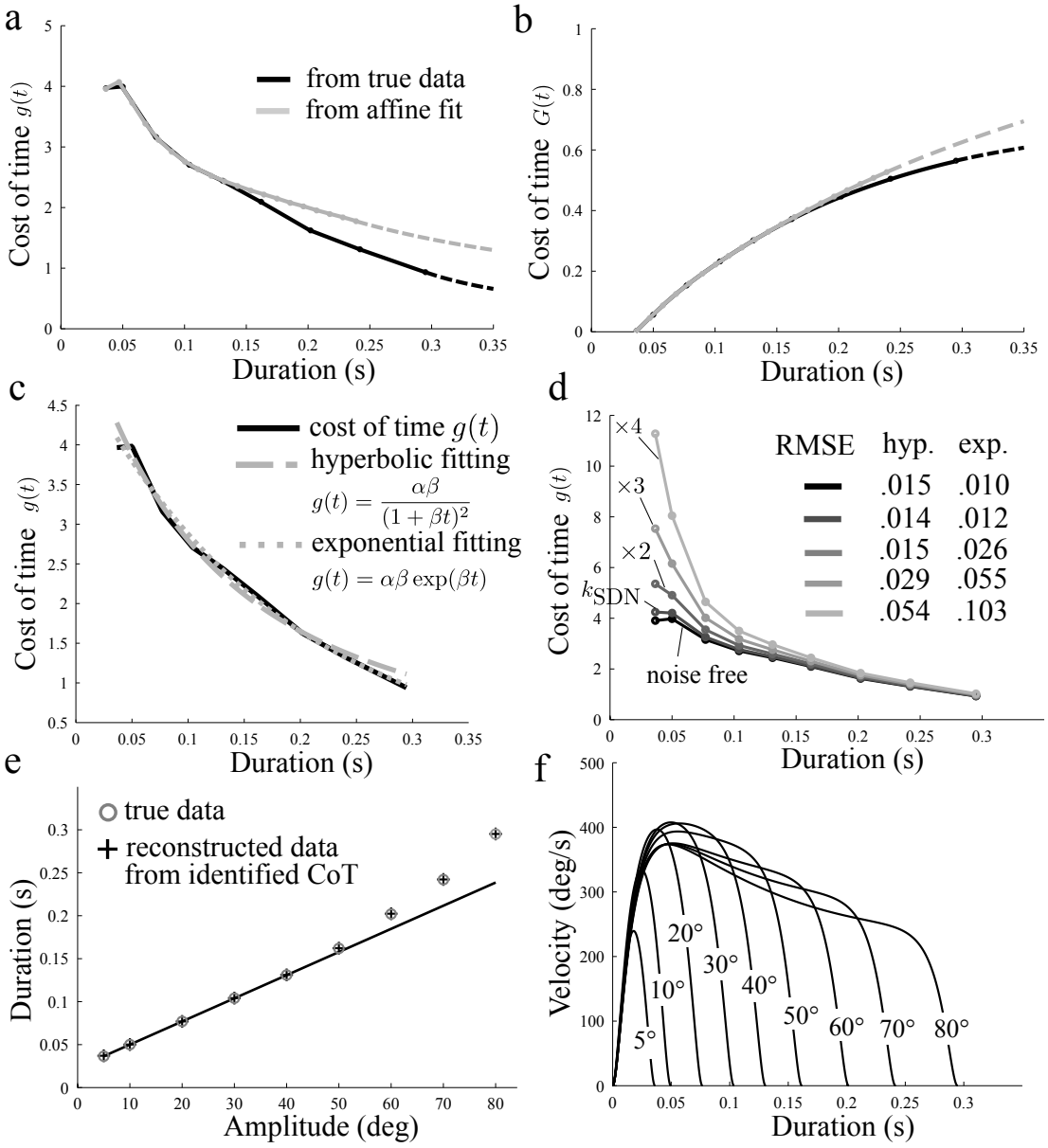

Figure 3: Cost of time underlying saccades, as inferred from data reported in Collewijn et al. (1988). **a**. Infinitesimal CoT $g(t)$. To infer this cost we used either the linear regression or the true data points exhibiting a growth larger than linear for amplitudes greater than 40 degrees (see Collewijn et al., 1988). Dotted traces indicate extrapolated parts of the CoT whereas solid lines are for values inferred using our inverse methodology. Note that our approach does not require neither fitting parameters nor hypothesizing the shape of the time cost. **b**. Integral CoT $G(t)$. The concavity of the CoT is obvious and in particular linear/convex costs can be ruled out (Shadmehr et al., 2010). **c**. Fitting of the infinitesimal CoT. We fitted g on the actual range of durations of observed saccades, i.e. from 36.5 ms to 295 ms. We considered both exponential and hyperbolic candidate functions. Best fitting parameters were: $(\alpha, \beta) = (1.3, 4.3)$ and $(\alpha, \beta) = (0.9, 5.5)$ for the hyperbolic and exponential fits, respectively. **d**. Simulations in the stochastic case with different levels of signal-dependent noise (from k_{SDN} to nk_{SDN} with $n = 2$, 3 and 4, denoted by $\times n$ in the figure). RMSE is reported for hyperbolic (hyp.) and exponential (exp.) functions, respectively. With larger multiplicative noise, the CoT gets closer to the hyperbolic class of CoT. **e**. Verification of the amplitude-duration relationship as obtained from free-time optimal control simulations with the CoT depicted in panel b (black trace). White-filled circles indicate true data points (i.e. target values for the model) and black crosses are for simulated data (i.e. reconstructed values from the model). Both series of points matched perfectly since the methodology provides an exact solution to the problem of identifying the CoT associated with the observed amplitude-duration relationship.

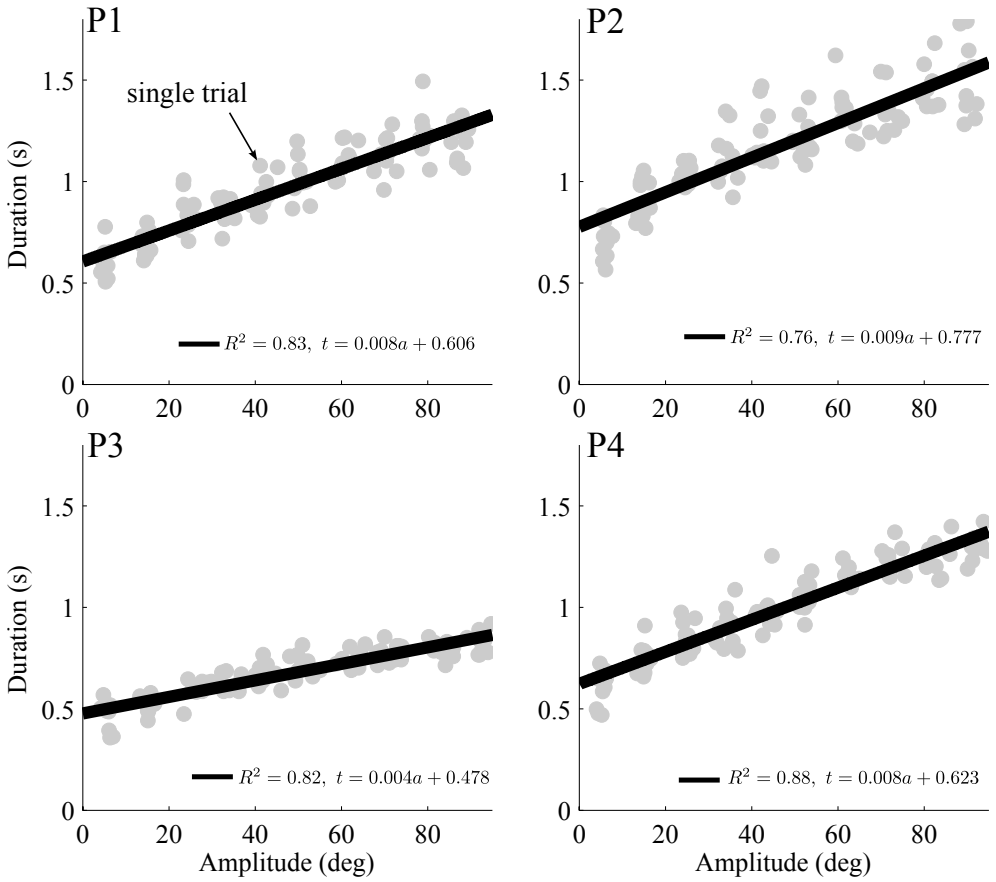

Figure 4: Experimental amplitude-duration relationships for four different subjects (P1-P4) during arm reaching. Each dot represents a single movement (trial) of amplitude lying between 5° and 95°. The result of a linear regression applied to the data is reported, namely an affine fit of the form $t = \alpha a + \beta$, with t being the MT in seconds and a the motion amplitude in degrees. Overall, even though inter-trial variability was noticeable, a clear and significant ($p < .001$) affine trend was observed in all cases, as revealed by the relatively large R^2 values.

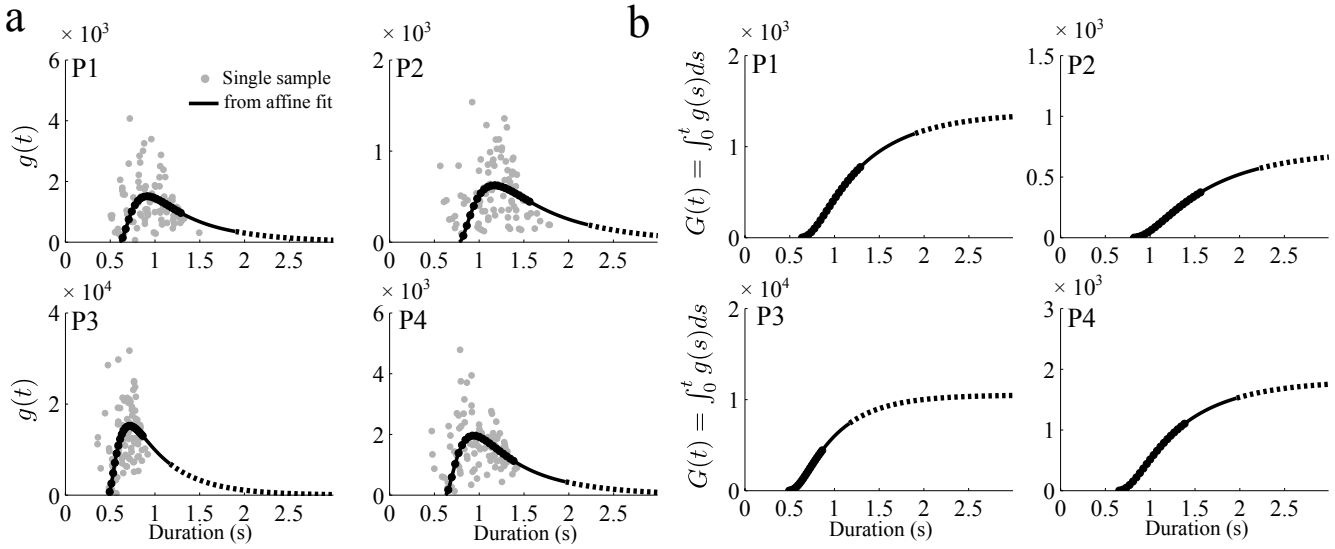

Figure 5: Cost of time underlying single-joint arm reaching, as inferred from data reported in Fig. 4. **a**. Infinitesimal CoT, $g(t)$, for each of the four participants (P1-P4). Solid black traces are values of $g(t)$ inferred from linear regressions. Experimentally, MTs belonged to a certain interval $[t_{\min}, t_{\max}]$ (emphasized by black-filled circles) but it was possible to identify $g(t)$ on a larger time interval by using predictions from the regression equation (black solid line). Dotted lines represent extrapolated values of $g(t)$, i.e. values of $g(t)$ not computed via our methodology. Light gray dots are values of $g(t)$ inferred from single trials: each gray dot corresponds to a dot displayed in Fig. 4. **b**. Integrated CoT, $G(t)$. These curves were obtained by trapezoidal integration of the infinitesimal CoTs for each participant separately. Note that the cost $G(t)$ is only identified up to a constant but in all cases the CoT exhibited a sigmoid-like shape on the range of actual motion durations (approximately between 500 ms to 1500 ms on the present data).

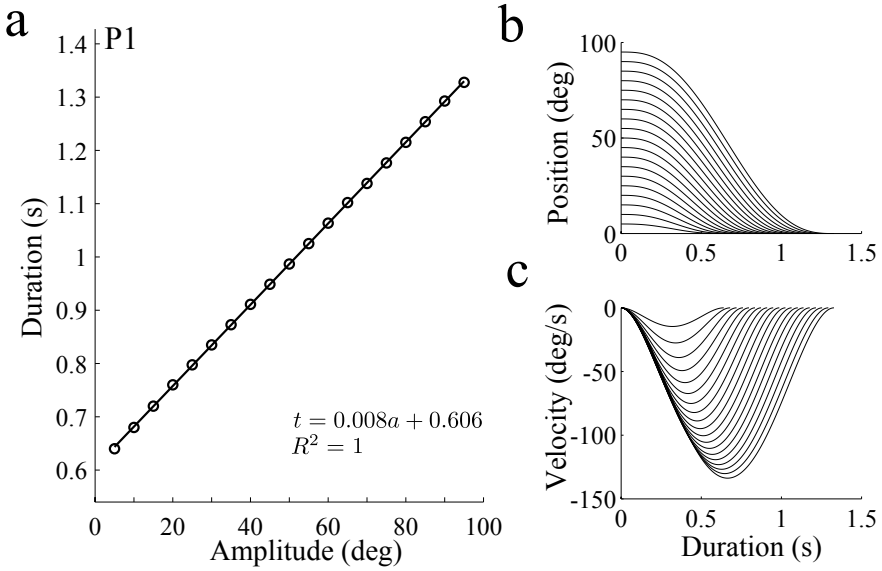

Figure 6: Verification of the methodology via free-time OC simulations. **a**. Amplitude-duration relationship recovered for participant P1 using the CoT identified in Fig. 5b (black trace). Since the methodology is exact, the same regression equation than in Fig. 4 is obtained. Note that MT was left free and emerged automatically for each amplitude in these simulations. **b**. Associated movement kinematics. Velocity profiles remained bell-shaped in accordance with experimental observations and scaled with MT.

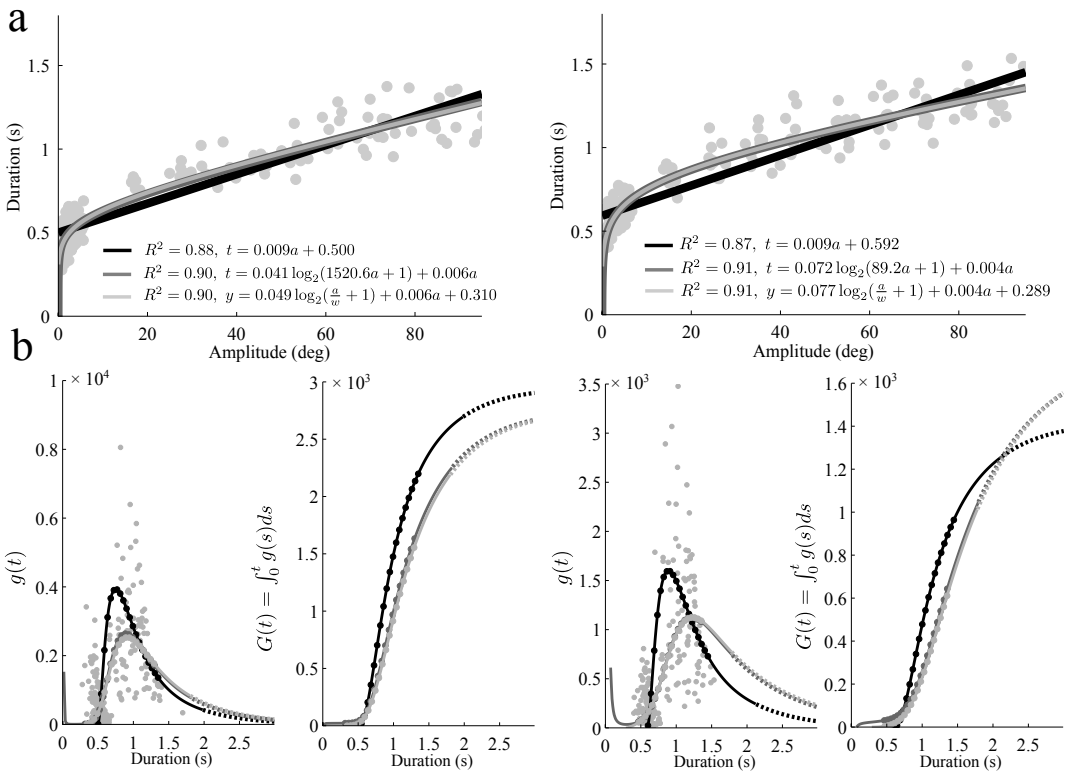

Figure 7: CoT consistency with respect to data fitting. **a.** Amplitude-duration relationship of two participants (left and right panels, respectively) who were asked to perform fully-extended arm movements in the horizontal plane, for amplitudes ranging from 1° to 95°. Besides linear functions, concave ones were tested for regression analysis (gray traces). The first log-based fitting had a zero-intercept constraint and was of the form $t = \alpha \log_2(\beta a + 1) + \gamma a$ (dark gray). The second log-based fitting (light gray) was chosen in the spirit of Fitts (1954) and the same than Young et al. (2009), i.e. of the form $t = \alpha \log_2(a/w + 1) + \beta a + \gamma$ where w denotes the actual width of the target (0.3 cm, i.e. ~0.2 deg here). **b.** On the range of empirical times (approximately between 300 ms and 1500 ms in these data), the CoT systematically exhibited a sigmoidal shape (filled circles on the traces). For smaller and larger times, we had to rely on extrapolations because data were no longer available (solid and dotted lines). Gray dots correspond to values of $g(t)$ identified on a single-trial basis and yield more noisy estimates.

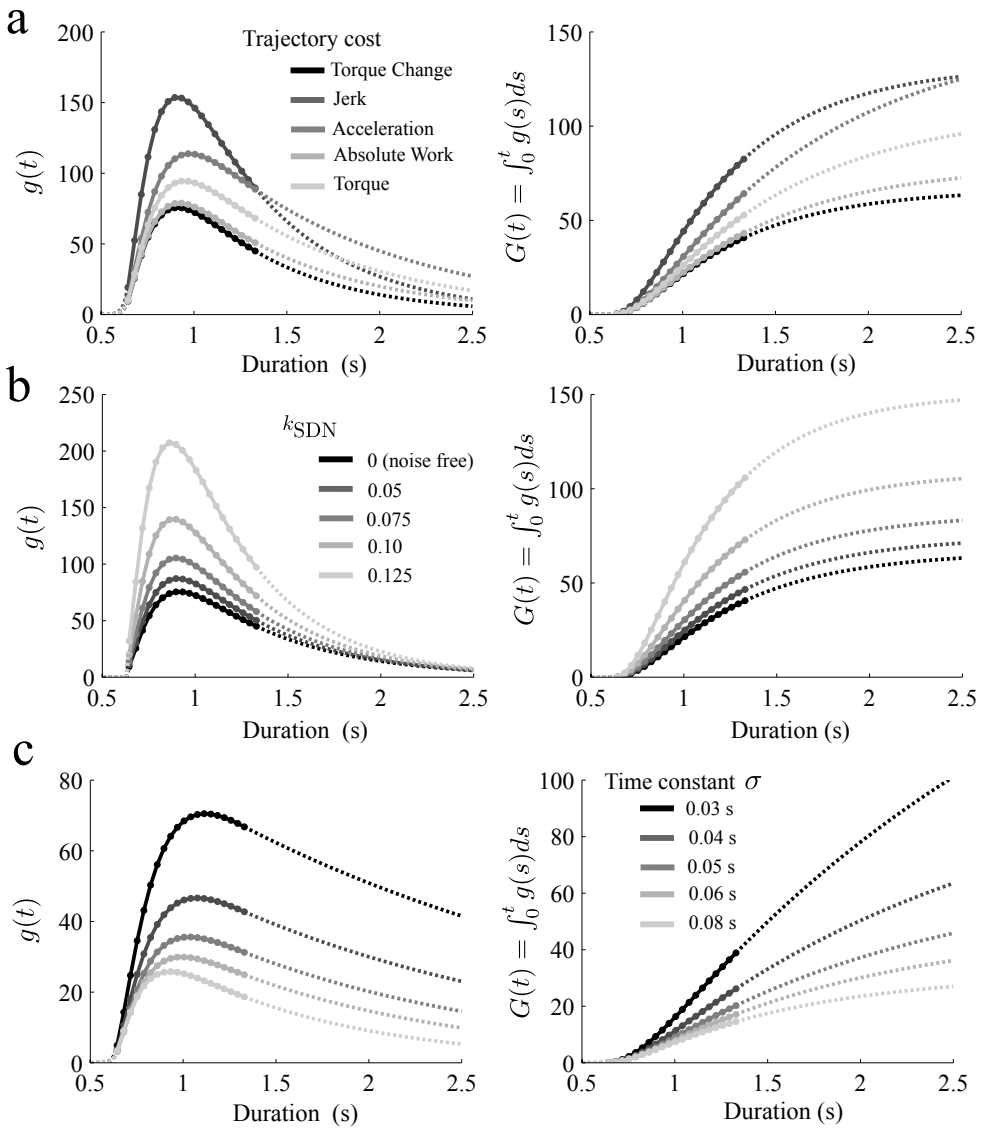

Figure 8: Cost of time consistency (for participant P1, with the affine fit of the amplitude-duration relationship of Fig. 4). **a**. CoT when assuming various trajectory costs. In black, the CoT for the torque change model, here scaled by a 0.05 factor for visualization purpose (infinitesimal CoT in *left* panel, integrated CoT in *right* panel). In gray scale, the CoT identified from various trajectory costs (jerk, acceleration, absolute work and torque). Although the specific values of the CoT necessarily changed with respect to the trajectory cost, its sigmoidal shape was nevertheless very robust. **b**. CoT when considering signal-dependent noise (k_{SDN}) and a stochastic OC formulation. Increasing multiplicative noise essentially induced an increase of the CoT but, again, the shape of the CoT remained sigmoidal. **c**. CoT when modeling muscle dynamics as 2nd-order low-pass filters with different time constants and optimizing a neural effort cost. Note that when still optimizing the torque change, no effects of the muscle dynamics was observed (result not depicted). When optimizing effort, the exact shape of the CoT depended on the underlying muscle dynamics. However, the shape of the infinitesimal CoT still showed a stiff increase after some time until a peak and followed by a slower decrease toward zero. Hence the systematic non-monotonicity of $g(t)$ proves that the CoT is neither concave nor convex but possesses a generalized logistic shape. Notably, this conclusion can be drawn from the truly identified part of the curves, which is depicted via solid lines with filled markers. Dotted lines still indicate extrapolated values of the CoT outside of the range of actual measurements (using exponential functions here).

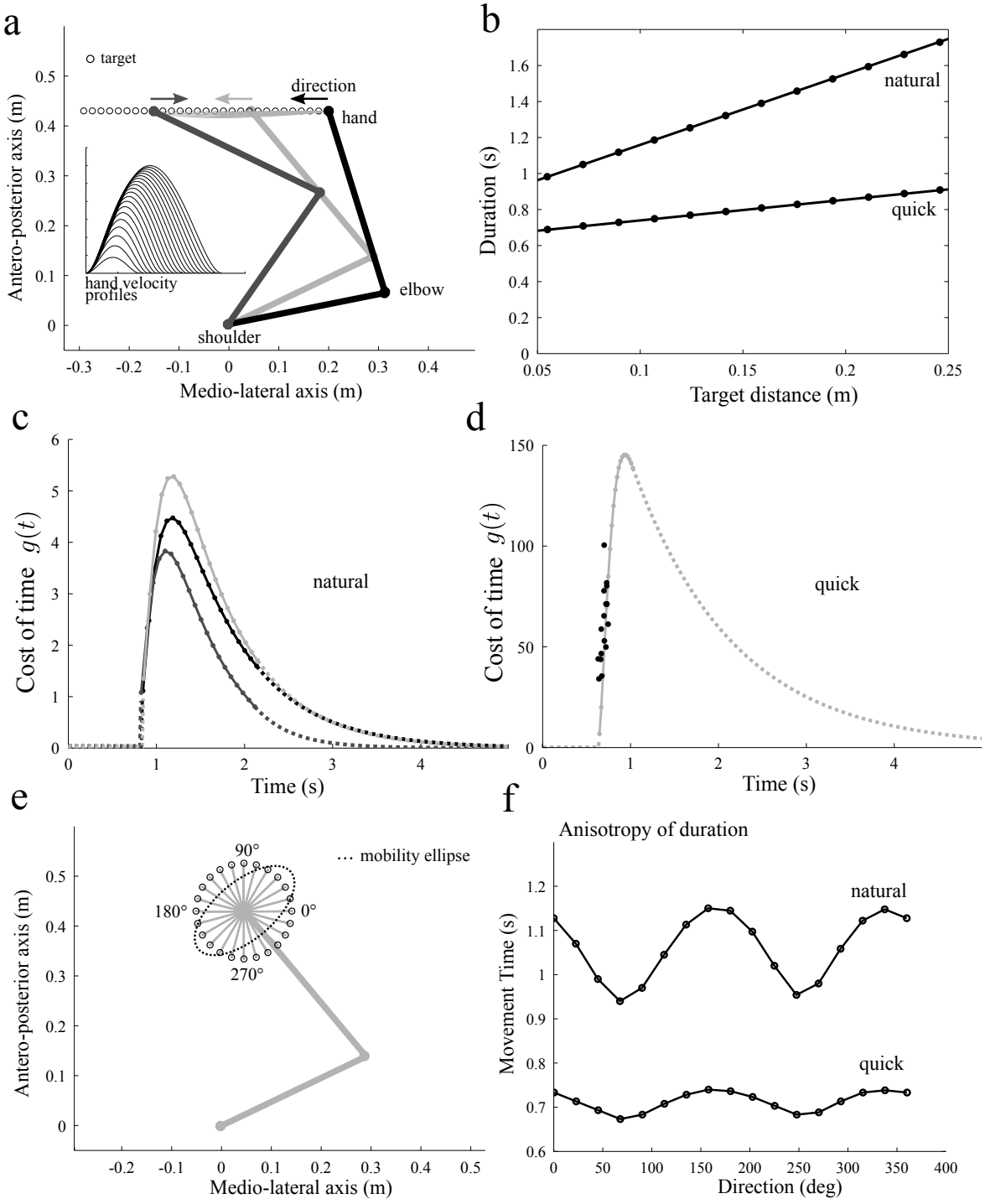

Figure 9: Cost of time underlying multijoint reaching, as inferred from data reported in Young et al. (2009) and van Beers et al. (2004). **a**. Illustration of the task and optimal trajectories as predicted by the model. **b**. Affine amplitude-duration relationships reported in Young et al. (2009) for natural and quick speeds. **c**. Infinitesimal CoT $g(t)$ identified for natural speed for rightward (dark gray) and leftward movements starting from different initial positions (black and light gray). **d**. Infinitesimal CoT identified for quick speed for the central starting position (light gray). Black dots represent pointwise evaluations of $g(t)$ when considering the center-out reaching task (panel e) and the corresponding MTs reported in van Beers et al. (2004). **e**. Center-out reaching task with the hand mobility ellipse depicted. **f**. MTs predicted by the model when solving free-time OCPs using the previously identified CoT (i.e. the light gray traces in panels c and d).