

**Generalized priority-based model for smartphone screen touches**Jean-Pascal Pfister<sup>1,\*</sup> and Arko Ghosh<sup>2,†</sup><sup>1</sup>*Institute of Neuroinformatics and Neuroscience Center Zurich, University of Zurich/ETH Zurich, 8057 Zürich, Switzerland and Department of Physiology, University of Bern, 3012 Bern, Switzerland*<sup>2</sup>*Cognitive Psychology Unit, Institute of Psychology, Leiden University, 2333 AK Leiden, The Netherlands*

(Received 7 February 2020; accepted 16 June 2020; published 14 July 2020)

The distribution of intervals between human actions such as email posts or keyboard strokes demonstrates distinct properties at short versus long timescales. For instance, at long timescales, which are presumably controlled by complex process such as planning and decision making, it has been shown that those interevent intervals follow a scale-invariant (or power-law) distribution. In contrast, at shorter timescales—which are governed by different processes such as sensorimotor skill—they do not follow the same distribution and we know little about how they relate to the scale-invariant pattern. Here, we analyzed 9 million intervals between smartphone screen touches of 84 individuals which span several orders of magnitudes (from milliseconds to hours). To capture these intervals, we extend a priority-based generative model to smartphone touching events. At short timescale, the model is governed by refractory effects, while at longer timescales, the intertouch intervals are governed by the priority difference between smartphone tasks and other tasks. The flexibility of the model allows us to capture interindividual variations at short and long timescales, while its tractability enables efficient model fitting. According to our model, each individual has a specific power-law exponent which is tightly related to the effective refractory time constant suggesting that motor processes which influence the fast actions are related to the higher cognitive processes governing the longer interevent intervals.

DOI: [10.1103/PhysRevE.102.012307](https://doi.org/10.1103/PhysRevE.102.012307)**I. INTRODUCTION**

Human actions such as mail correspondences, library loans, or website visits are not equally distributed in time but are typically structured in bursts followed by long periods of inactivity [1–3]. Several types of models have been proposed to capture the power-law structure of interevent time distribution (for a review see Ref. [3]). Priority-based queuing models [4–9] rely on one (or multiple) list(s) of tasks to be executed, where each task is associated with a priority level which directly influences the timing of its execution. This class of models have been pioneered by Barabási [4] and then generalized to multiple interacting queues [7,9], time-varying priorities [10], or priorities which depend on the position within the list of tasks [11]. Those models provide an interesting interpretation for the origin of the power-law scaling for long intervals (they come from prioritizing tasks) but are usually not designed to capture short interevent timings.

Poisson-based models belong to another class [12–15]. They rely on the assumption that the event rate is governed by a Poisson process whose rate can change over time. Because of this dynamic rate assumption, those Poisson-based models can easily accommodate a precise description of short interevent intervals. In the simplest case where the Poisson rates are piecewise constant and stochastically jump at each event time, the power-law exponent can be directly obtained from the distribution of Poisson rates [12]. This approach

has been extended to continuously changing Poisson rates [12–15]. In particular, the framework proposed by Malmgren *et al.* [14] provides a circadian explanation for the origin of power-law distributions. Self-exciting point processes (also called Hawkes processes [16,17]) have also been used to provide a mechanistic interpretation of power law interevent intervals [18,19]. Those models are very flexible (they can accommodate both short and long interevent intervals), but lack the priority-related interpretation. Indeed, it is unclear how those Poisson-based models relate to priority-based models.

Here, we start from a priority-based framework, generalise it on different levels and apply it to smartphone touchscreen interaction data (see Fig. 1). First, unlike the priority-based model from Barabási, our priority-based model aims at predicting interevent distribution instead of response-time distribution. Second, our model is the continuous-time extension of the classical priority-based model. Under this limit, we can compute analytically the interevent distribution and show that our generalized priority-based model can be mapped to Poisson-based models. Third, our priority-based model does not only describe long interevent intervals but also includes a detailed description of short interevent intervals and thereby overcomes the need to define an arbitrary onset of the power-law distribution [20]. In particular, it assumes that the agent remains in a so-called refractory state during a short time after each event, where the probability of generating a new event is reduced.

Finally, because our model is based on arbitrary priority distribution and not on specific priority distribution imposed by the presence of lists (with discrete number of items), it can produce any power-law exponent. We found that for each

\*Corresponding author: [pfister@pyl.unibe.ch](mailto:pfister@pyl.unibe.ch)†[a.ghosh@fsw.leidenuniv.nl](mailto:a.ghosh@fsw.leidenuniv.nl)

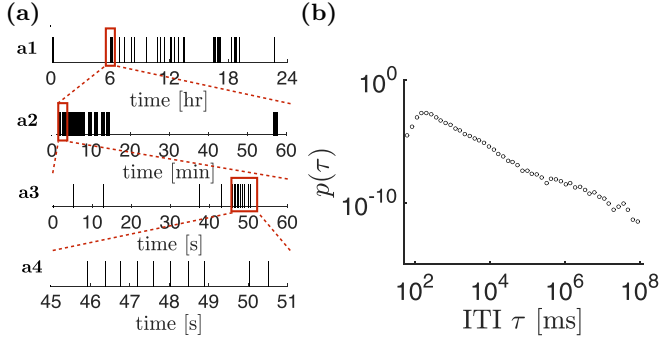


FIG. 1. Smartphone touch data. (a) Smartphone touch events (vertical bars) are characterized by bursts as well as long gaps at timescales of hours (a1), minutes (a2), and seconds (a3). At the sub-second timescale (a4), touches are more regular. (b) The intertouch interval (ITI) distribution is scale free from seconds to hours. Data from one individual.

subject, the intertouch interval (ITI) distribution is different and well captured by the model. We also found that from those fitted parameters, we can quantify the relative priority placed on smartphone actions.

## II. SMARTPHONE TOUCHING MODEL

### A. Discrete-time model

In the first step, we propose a discrete-time generative model for smartphone touches. This model extends existing priority-based models by including refractoriness [4,7]. The output of the model is the set of touch times  $\{t_0, t_1, \dots, t_N\}$ , where  $t_i$  can take discrete values, i.e.,  $t_i = k_i \Delta t$ , with  $\Delta t$  being the bin width and  $k_i \in \mathbb{N}$ . Equivalently, the model output can be described by the touch train  $s_t$  where  $s_t = 1$  denotes the presence of a touch while  $s_t = 0$  indicates the absence of a touch.

Every touch is the result of a decision process. We assume that an individual can perform tasks from only two categories: either a task related to a smartphone screen *touch* or *other* task such as driving a car. In each category, there can be important tasks (such as dialing an emergency number) or less important tasks (such as checking the news). So we will assume that every task can be described by its priority level which is a number between 0 and 1. Let  $x_t \in [0, 1]$  denote the priority associated with a *touch* task at time  $t$  and  $y_t \in [0, 1]$  the priority associated to the *other* task. If at time  $t$  the *touch* task associated to priority  $x_t$  is executed (i.e.,  $s_t = 1$ ), then a new *touch* task is considered and will be attributed a new *touch* priority value drawn from the *touch* priority distribution, i.e.,  $x_{\text{new}} \sim p(x)$ . If the *touch* task is not executed ( $s_t = 0$ ), then its priority remains the same. This can be summarized as

$$x_{t+\Delta t} = x_t(1 - s_t) + x_{\text{new}}s_t \quad x_{\text{new}} \sim p(x). \quad (1)$$

Conversely, the dynamics for the *other* priority  $y_t$  is such that when the screen is not touched at time  $t$  (i.e.,  $s_t = 0$ ), then it is the *other* action that is executed and a new priority  $y_{\text{new}}$  must be drawn from  $q(y)$ . This is summarized as

$$y_{t+\Delta t} = y_t s_t + y_{\text{new}}(1 - s_t), \quad y_{\text{new}} \sim q(x). \quad (2)$$

To generate a smartphone touch, two conditions need to be satisfied. First, the priority  $x_t$  of the smartphone action needs to be greater than the priority  $y_t$  of the other action, and second, the individual must be in a nonrefractory state. Formally, the touch variable  $s_t$  is sampled from the following Bernoulli distribution:

$$s_t \sim \text{Bernoulli}(\lambda(x_t, y_t, \tau_t)\Delta t), \quad (3)$$

where the touching intensity  $\lambda$  (probability per time bin  $\Delta t$ ) is given by

$$\lambda(x, y, \tau) = \rho r(\tau)H(x - y), \quad (4)$$

where  $\tau = t - \hat{t}$  is the time since last touch ( $\hat{t} = \max_{t_k} \{t_k < t\}$ ) and  $H$  is the Heaviside step function which guarantees that touches can only be generated when  $x > y$  and  $\rho$  is the touching rate.  $r(\tau) \geq 0$  is the refractory function which includes post-touch effects (i.e., right after a touch, the touch probability can be reduced). A hard refractoriness function takes the following form:

$$r(\tau) = H(\tau - \Delta), \quad (5)$$

where  $H(\tau)$  is the Heaviside step function [i.e.,  $H(\tau) = 1$  if  $\tau \geq 0$  and  $H(\tau) = 0$  else] and  $\Delta$  the hard refractory time (i.e., minimal ITI). If we relax this strong condition and allow touches for any  $\tau > 0$  (but with reduced probability when  $\tau \simeq 0$ ), then we can define a relative refractoriness function as a sum of basis functions,

$$r(\tau) = 1 + \sum_{k=1}^n \gamma_k \exp(-\alpha_k \tau), \quad (6)$$

with logarithmically spaced inverse time constants, i.e.,  $\alpha_k = \alpha_1 \beta^{-(k-1)}$ . We took  $\alpha_1^{-1} = 50$  ms and set  $\beta$  such that  $\alpha_n^{-1} = 1000$  ms. Note that the set  $\{\gamma_k\}_{k=1}^n$  has to be chosen such that the condition  $r(\tau) \geq 0$  is satisfied for all  $\tau \geq 0$ . If  $r(0) \leq 0.5$ , then we define the effective time constant  $\tau^*$  as the time for which refractoriness is half, i.e.,  $r(\tau^*) = 0.5$ ; see Fig. 4(c). In the rare cases where multiple solutions exist for  $\tau^*$  (which can occur when  $r(\tau)$  is nonmonotonic), we took the maximal value of the set of solutions.

The discrete-time model described by Eqs. (1), (2), and (3) is a latent dynamical system. Note that sampling this model is slow since the complexity of this sampling scheme scales with the number of bins. Even more critical is the learning procedure for such a latent dynamical model which can be prohibitively slow for smartphone touching data sets which typically extend over months. A much faster sampling scheme is proposed below.

### B. Continuous-time model

The idea of the continuous-time model is to directly sample the intervals  $\tau$  instead of sampling the touch variable  $s_t$  at each time step. The transition to this continuous model can be done in two steps. First, we observe that when  $\Delta t$  is small, the *other* priorities  $y_t$  constantly change (except at the rare times where  $s_t = 1$ ), i.e., Eq. (2) can be approximated as  $y_t \sim q(y)$ . This means that the priorities  $y_t$  are independent of time and therefore, the probability of generating a touch

can be marginalized over  $y_i$ :

$$\begin{aligned} p(s_i|x_i, \tau_i) &= \int_0^1 p(s_i|x_i, y_i, \tau_i)q(y_i)dy_i \\ &= \text{Bernoulli}(\bar{\lambda}(x_i, \tau_i)\Delta t), \end{aligned} \quad (7)$$

where the average touching intensity  $\bar{\lambda}$  is given by

$$\bar{\lambda}(x, \tau) = \int_0^1 \lambda(x, y, \tau)q(y)dy = \rho r(\tau)\pi(x), \quad (8)$$

and  $\pi(x)$  is the probability of having  $x > y$  for a given  $x$ ,

$$\pi(x) = \int_0^x q(y)dy. \quad (9)$$

In the second step, we take the limit  $\Delta t \rightarrow 0$  and therefore, the intertouch interval distribution conditioned on  $x$  can be expressed as (see also Ref. [21]):

$$p(\tau|x) = \bar{\lambda}(x, \tau) \exp\left(-\int_0^\tau \bar{\lambda}(x, t)dt\right). \quad (10)$$

The unconditioned ITI distribution is obtained by averaging the conditioned ITI distribution over the *touch* priority distribution  $p(x)$ :

$$p(\tau) = \int_0^1 \bar{\lambda}(x, \tau) \exp\left(-\int_0^\tau \bar{\lambda}(x, t)dt\right) p(x)dx. \quad (11)$$

So samples of the continuous-time model can be simply obtained in a two-step procedure. First,  $x$  is sampled from  $p(x)$ , then  $\tau$  is sampled from  $p(\tau|x)$  given by Eq. (10). For this second step, one can use the time rescaling theorem [21]. Note that this continuous-time model describes a renewal process and hence the sampling complexity scales with the number of touches  $N$ .

In the absence of refractoriness [i.e.,  $r(\tau) = 1$ ], the sampling procedure is even simpler. First, a Poisson rate  $\bar{\lambda} = \rho x$  can be drawn from a distribution of rates  $p(\bar{\lambda})$  with maximal rate  $\bar{\lambda}_{\max} = \rho$ , then  $\tau$  is sampled from an exponential distribution  $p(\tau|\bar{\lambda}) = \bar{\lambda} \exp(-\bar{\lambda}\tau)$  and the ITI can be expressed as

$$p(\tau) = \int_0^\rho p(\tau|\bar{\lambda})p(\bar{\lambda})d\bar{\lambda}, \quad (12)$$

which is precisely the ITI one would get from an heterogeneous Poisson model [12]. This shows the equivalence between the priority-based model and the Poisson-based models.

### III. PROPERTIES OF THE MODEL

#### A. Invariance of the model

Before giving a parametric form for all distributions, let us first note an invariant property of the model. In particular, it can be shown (see Appendix A 1) that the ITI distribution given by Eq. (11) remains unchanged if the pair of priority distributions  $[p(x), q(y)]$  is replaced by  $[\tilde{p}(x), \tilde{q}(y)]$  given by

$$\tilde{p}(x) = p[\phi(x)]\phi'(x) \quad \text{and} \quad \tilde{q}(y) = q[\phi(y)]\phi'(y), \quad (13)$$

where  $\phi$  is a differentiable and strictly monotonously increasing function with boundary conditions  $\phi(0) = 0$  and  $\phi(1) = 1$ . This invariance can be understood intuitively by noting that the notion of priority contains some arbitrariness. Indeed,

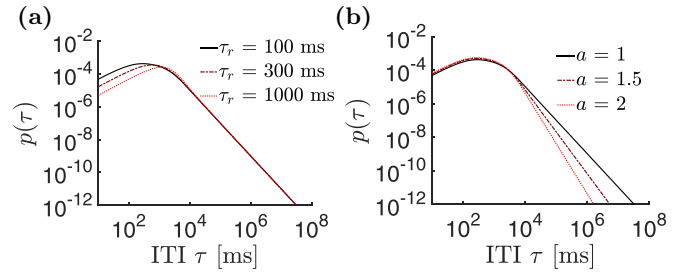


FIG. 2. Properties of the smartphone touching model. (a) The refractory time constant affects the early part of the ITI distribution.  $n = 1$ ,  $\tau_r = \alpha_1^{-1}$ . (b) The parameter  $a$  from the priority distribution affects the power-law exponent of the ITI distribution.

the only element which is relevant in the decision process is whether  $x$  is larger or smaller than  $y$  [see Eq. (4)]. If we define a new priority  $x' = \phi(x)$  (with the above conditions on  $\phi$ ), then we observe that the ordering remains unchanged, i.e.,  $x > y \Rightarrow \phi(x) > \phi(y)$ . This observation can also be made more formally with a change of variable in Eq. (11) (see Appendix A 1). Second, this invariance property of the model means that without loss of generality, we can set one distribution and rescale the other one. For example, without loss of generality, we can set  $q(y) = 1$ . For the *touch* priority distribution, we will assume that it is given by a  $\beta$  distribution:

$$p(x) = \beta(x; a, b) = \frac{x^{a-1}(1-x)^{b-1}}{B(a, b)}, \quad (14)$$

where  $B(a, b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$  is the  $\beta$  function. With the above choice of  $q$ , the ITI distribution in Eq. (11) can be rewritten in a simpler form

$$p(\tau) = \rho r(\tau) \int_0^1 x \exp\left(-x\rho \int_0^\tau r(t)dt\right) p(x)dx. \quad (15)$$

#### B. Scale-free intertouch interval distribution

For short timescales ( $\tau < \alpha_n^{-1}$ ), the ITI distribution is governed by the refractory function  $r$  [see Fig. 2(a)]. However, for longer timescales ( $\tau \gg \alpha_n^{-1}$ ), the ITI distribution follows a power-law distribution. This can be seen in two steps. First, in the limit of large  $\tau$ , we have  $r(\tau) \rightarrow 1$ . Second, in the limit of large  $\tau$ , we know from Eq. (15) that the ITI distribution is only sensitive to the *touch* priority distribution in the vicinity of  $x = 0$  that we denote as  $p_0(x)$ . Note that  $p_0(x)$  is not normalized. For the  $\beta$  distribution, we have  $p(x) \rightarrow p_0(x) = x^{a-1}/B(a, b)$  when  $x \rightarrow 0$ . Therefore, when  $\tau \gg \alpha_n^{-1}$ , the ITI distribution can be approximated as

$$\begin{aligned} p(\tau) &\simeq \rho \int_0^1 x p_0(x) e^{-x\rho\tau} dx \\ &\simeq \frac{\Gamma(a+1)}{B(a, b)\rho^a} \tau^{-(a+1)}, \end{aligned} \quad (16)$$

where  $\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$  is the  $\Gamma$  function. Therefore, the power-law exponent is given by  $a + 1$  [see Fig. 2(b)]. So when  $a$  is large, *touching* tasks are more important—in the sense that there are few low-priority touch tasks—and consequently there are less large ITI.

TABLE I. List of models.

Model	Parameters	Number of parameters	Assumptions
$M_1$	$\theta = (a, \rho)$	2	No ref.
$M_2$	$\theta = (a, b, \rho)$	3	No ref.
$M_3$	$\theta = (a, \rho, \Delta)$	3	$r$ : hard, $b = 1$
$M_4$	$\theta = (a, b, \rho, \Delta)$	4	$r$ : hard
$M_5$	$\theta = (a, \rho, \gamma_1, \dots, \gamma_n)$	$n + 2$	$r$ : rel, $b = 1$
$M_6$	$\theta = (a, b, \rho, \gamma_1, \dots, \gamma_n)$	$n + 3$	$r$ : rel
$M_7$	$\theta = (\kappa, \varphi)$	2	No ref.
$M_8$	$\theta = (\kappa, \lambda)$	2	No ref.
$M_9$	$\theta = (\mu, \sigma^2)$	2	No ref.

## IV. RESULTS

### A. Model fitting

For each subject, we fitted six versions of the priority-based model ( $M_1$ – $M_6$ ) and three benchmark models ( $M_7$ – $M_9$ ). Those six versions are obtained by fixing a subset of parameters to specific values in the case of no refractoriness ( $M_1$ – $M_2$ ), hard refractoriness ( $M_3$ – $M_4$ ) as well as relative refractoriness ( $M_5$ – $M_6$ ) (see Table I). The three benchmark models ( $M_7$ – $M_9$ ) are further detailed in Appendix A 4. So, in total, we fitted nine models per subject:

(1) Model  $M_1$  is the simplest model and contains only two parameters:  $\theta = (a, \rho)$ . It is assumed that  $b = 1$  and that there is no refractoriness ( $\Delta = 0$ ).

(2) Model  $M_2$  is the same as model 1 except that the *touch* priority distribution has 2 free parameters:  $a$  and  $b$ . Overall, it contains three parameters:  $\theta = (a, b, \rho)$ .

(3) Model  $M_3$  includes hard refractoriness (with refractory time  $\Delta$ ) but assumes  $b = 1$ . It contains therefore three parameters:  $\theta = (a, \rho, \Delta)$ .

(4) Model  $M_4$  is the same as model 3 except that the *touch* priority distribution is described by both  $a$  and  $b$ . It contains four parameters:  $\theta = (a, b, \rho, \Delta)$ .

(5) Model  $M_5$  uses a relative refractory kernel parametrized by  $n$  basis functions with coefficients  $\gamma_1, \dots, \gamma_n$ . It also assumes that  $b = 1$ . So the model contains  $n + 2$  parameters:  $\theta = (a, \rho, \gamma_1, \dots, \gamma_n)$ .

(6) Model  $M_6$  is the same as model  $M_5$ , but  $b$  is not constrained to be equal to one. The models contains therefore  $n + 3$  parameters:  $\theta = (a, b, \rho, \gamma_1, \dots, \gamma_n)$ .

(7) Model  $M_7$  is the gamma distribution with  $\theta = (\kappa, \varphi)$ .

(8) Model  $M_8$  is the Weibull distribution with  $\theta = (\kappa, \lambda)$ .

(9) Model  $M_9$  is the log-normal distribution with  $\theta = (\mu, \sigma^2)$ .

For each model and for each subject, the model parameters  $\theta$  are fitted from the set  $\mathcal{D} = \{\tau_i\}_{i=1}^N$  of intertouch intervals  $\tau_i = t_i - t_{i-1}$ . To do so, we relied on the continuous-time model which massively simplifies the expression of the log-likelihood. Indeed, the detailed model can be seen as a dynamical latent variable model (where the latent variables are  $x$  and  $y$ ) which can be fitted through EM type algorithm but is known to be very slow. Here, because of the analytical expression of the ITI for the continuous-time model [see Eq. (15)], we can express the following objective

function:

$$\mathcal{L}(\theta) = L(\theta) - \lambda \sum_{k=1}^n \gamma_k^2, \quad (17)$$

which is the log-likelihood  $L(\theta) = \sum_{i=1}^N \log p(\tau_i)$  [see Eq. (A6)] minus a regularization term on the coefficients  $\gamma_k$  to prevent overfitting. This regularization term (with  $\lambda = 1000$ ) is only used in models 5 and 6. Note that this objective function can be seen as the log-posterior with a Gaussian prior (with variance  $1/2\lambda$ ) on the coefficients  $\gamma_k$  and a flat prior for the other parameters.

Because the refractory kernel must remain positive for all time, i.e.  $r(\tau) \geq 0, \forall \tau \geq 0$ , the optimization task can be expressed as

$$\begin{aligned} \theta^* &= \arg \max_{\theta} \mathcal{L}(\theta) \\ \text{s.t.} \quad &\sum_{k=1}^n \exp(-\alpha_k \tau) \gamma_k^* \geq -1 \quad \forall \tau \geq 0. \end{aligned} \quad (18)$$

However, the difficulty of the optimization problem defined in Eq. (18) lies in the fact that the constraints are defined for all  $\tau \geq 0$  (i.e., infinitely many inequality constraints). For a practical numerical implementation, we defined a grid of  $M = 200$  points  $\tau_1, \dots, \tau_M$  where the first 100 points are linearly spaced ( $\tau_i = i\Delta t$  for  $1 \leq i < 100$  and  $\Delta t = 1$  ms) and the subsequent 100 points are logarithmically spaced ( $\tau_i = \kappa^{i-100} \tau_{100}$  for  $100 < i \leq 200$  where  $\kappa$  is set such that  $\tau_{200} = 3\alpha_n^{-1}$ ). So we replace the inequality constraints of Eq. (18) by

$$\sum_{k=1}^n \exp(-\alpha_k \tau_i) \gamma_k^* \geq -1 \quad \forall i = 1, \dots, M.$$

### B. Fitting results

We recorded smartphone touches from 84 individuals for an average duration of 36.5 days (see Appendix B 1 for details on data collection). The average number of smartphone screen touches per day ranged from 285 to 9915 with a median value of 2540 touches per day.

For each individual, the six different models were fitted according to the procedure described above. In particular, we first fitted the models without refractoriness ( $M_1$  and  $M_2$ ) and the models with hard refractoriness ( $M_3$  and  $M_4$ ). We found that the likelihood can be drastically improved by adding the hard refractory time parameter  $\Delta$  [see Figs. 3(a), 3(b) and Fig. 5(a)]. Actually, the optimal value is exactly  $\Delta^* = \tau_{\min}$ , where  $\tau_{\min}$  is the minimal ITI [see Eq. (A15)]. The fitted ITI for model  $M_4$  [see Figs. 3(c) and 3(d)] is decent, but short ITI are not well captured.

We then fitted the models with relative refractoriness ( $M_5$  and  $M_6$ ) and displayed the fitting results of the best model ( $M_6$  with  $n = 21$  basis functions); see Fig 4. We found that for each individual the empirical ITI distribution [see Fig. 4(a)] is well captured by the model both for the short timescales (up to 1 s) which is strongly influenced by the refractory kernel  $r(t)$  [see Fig. 4(c)] as well as the longer ITI which has a typical power-law decay. Note that because of the richness



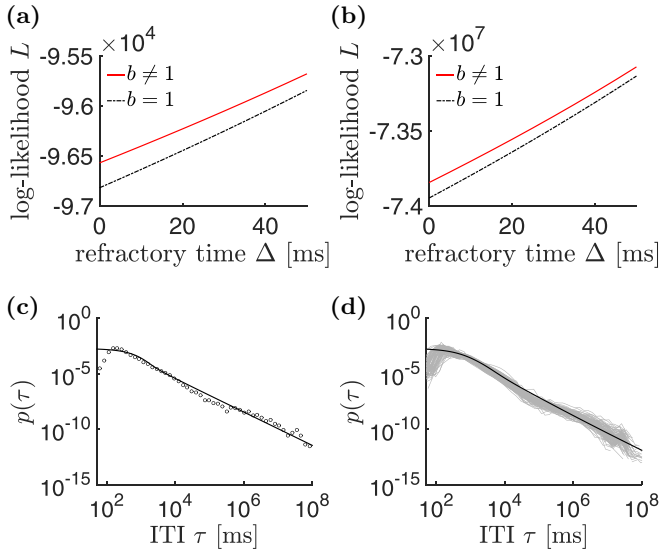


FIG. 3. Fitting results for the models with hard refractoriness ( $M_3$  and  $M_4$ ). (a) Log-likelihood of model  $M_3$  (black, dashed dotted) and  $M_4$  (red, solid) as a function of the refractory time  $\Delta$  for a single subject. The best refractory time is  $\Delta^* = 50$  ms. (b) The log-likelihood summed across subjects also elicits an optimal refractory time at  $\Delta = 50^*$  ms for both model  $M_3$  (black, dashed dotted) and  $M_4$  (red, solid). (c) Intertap interval distribution for one subject (solid line, fit; circles, data). (d) Intertouch interval distribution across the whole population. Each gray line corresponds to the data from one subject. Solid black line denotes the median model ITI distribution.

of the data, the power-law relationship extends over 5 decades (from  $10^3$  to  $10^8$  ms).

The fitted refractory kernel [see Fig. 4(c)] shows a strong reduction of touching rate during the first few hundreds of milliseconds after the last touch. For one subject, it even displays a small increase in touching rate about 300 ms after the last touch [see Fig. 4(d)]. This smooth transition from short ITI to longer ITI removes the need to define an arbitrary onset of the power-law distribution [20].

The fitted touch priority distribution [see Fig. 4(e)] [assuming that the *other* priority distribution is given by  $q(y) = 1$ ] diverges for small priorities (which is the case when  $a < 1$ ). We repeated this fitting procedure for the 84 subjects. The population results are displayed on Figs. 4(b), 4(d) and 4(e). We found that over the population the priority parameter  $a$  is fairly scattered around a median value of  $a = 0.53$  (for model  $M_6$ ) and of  $a = 0.49$  (for model  $M_5$ ). The large interindividual differences is also highlighted in Fig. 4(g) which displays a broad distribution of touching rate  $\rho$  over the population.

### C. Model comparison

To compare the different models (see Table I) for each individual, we can use the Bayesian information criterion (BIC) which is well suited for large data sets (i.e., large number of touching intervals  $N$ ) which is precisely our case. BIC is given by  $\text{BIC} = \log(N)|\theta| - 2\mathcal{L}(\theta^*)$ , where  $|\theta|$  is the number of parameters and  $\mathcal{L}(\theta^*)$  is the objective function given by Eq. (17) and is evaluated at the MAP parameter  $\theta^*$ .

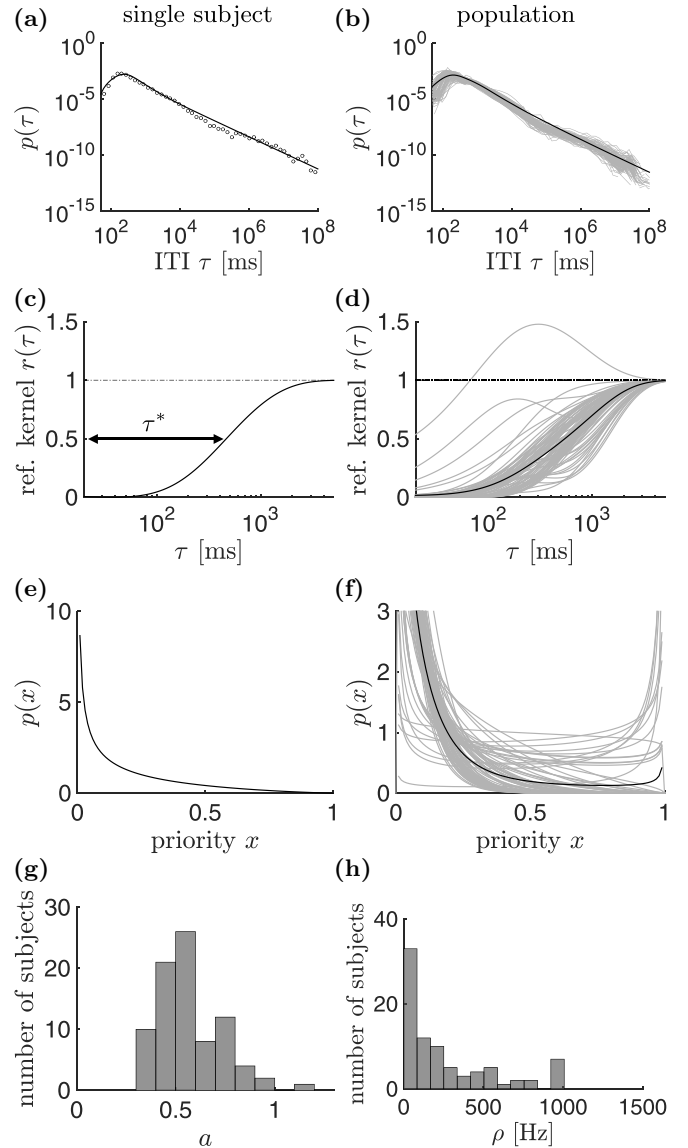


FIG. 4. Fitting results of model  $M_6$  (with  $n = 21$ ) for one subject (a, c, and e) and for the population of 84 subjects (b, d, and f). (a) The ITI distribution for one given subject (open circles) is well captured model (solid line). (c) Refractory kernel. The effective time constant  $\tau^*$  is defined as  $r(\tau^*) = 0.5$ . (e) *touch* priority distribution (with  $q = 1$ ). (b, d, and f) same as in (a, c, and e) but for each of the 84 subjects (gray lines). Solid lines denote the median ITI (b), refractory kernel (d), and priority distribution (f). (g) Distribution of the parameter  $a$ . (h) Distribution of the touching rate  $\rho$ .

To compare the different models for the whole population of  $S = 84$  subjects, we can define a population BIC in an analogously to the individual BIC given above. Let  $N_{\text{pop}} = \sum_{s=1}^S N(s)$  denote the total number of intertouch intervals of the whole population where  $N(s)$  is the number of data points of subject  $s$ . Let  $|\theta_{\text{pop}}| = S|\theta|$  denote the total number of fitted parameters and let  $\mathcal{L}_{\text{pop}}(\theta_{\text{pop}}^*) = \sum_{s=1}^S \mathcal{L}(\theta^*(s))$  denote the population objective function where  $\theta^*(s)$  denotes the fitted parameters of subject  $s$ . The population BIC is therefore

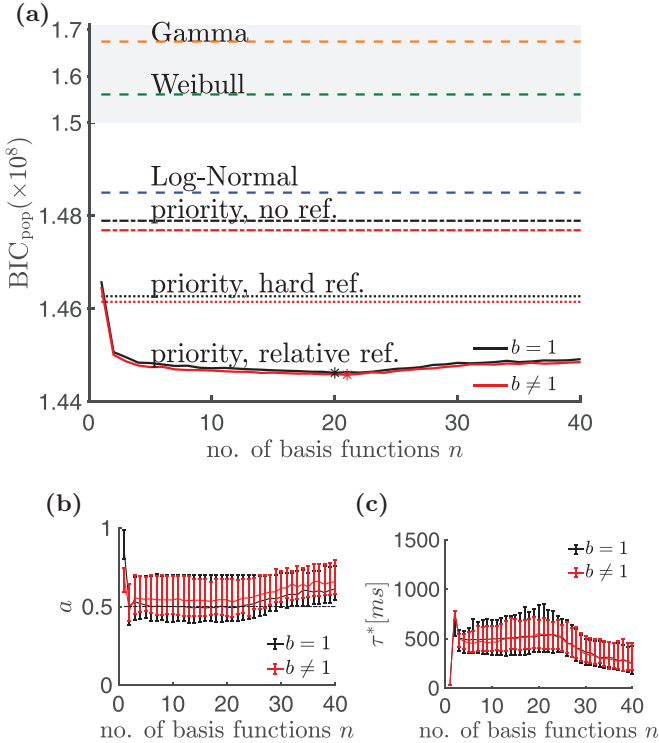


FIG. 5. Model comparison. (a) Comparison of the population BIC for the different models as a function of the number of basis functions  $n$ . Priority-based models (black and red) largely outperform simple models such as the gamma distribution (dashed, orange), the Weibull distribution (dashed, green), and the log-normal distribution (dashed blue). Note that the piecewise linear y-axis is changing at  $\text{BIC}_{\text{pop}} = 1.5 \times 10^8$ . Furthermore, models with relative refractoriness ( $M_5$  and  $M_6$ , solid lines) outperform models with hard refractoriness ( $M_3$  and  $M_4$ , dotted lines;  $M_1$  and  $M_2$ , dot-dashed lines). The best number of basis functions for model  $M_5$  is  $n^* = 20$  (black star). The overall best model is  $M_6$  with  $n = 21$  (red star). Note that the BIC difference between the red and the black star is  $-3.8 \times 10^4$ . (b) Quantile plot for the priority parameter  $a$  for model  $M_5$  (black) and model  $M_6$  (red). Solid lines denote the median across subjects and the error bars denote 25 and 75 percentiles. (c) Same as panel (b) but for the effective time constant  $\tau^*$ .

given by

$$\text{BIC}_{\text{pop}} = \log(N_{\text{pop}})S|\theta| - 2 \sum_{s=1}^S \mathcal{L}[\theta^*(s)]. \quad (19)$$

First, we found that the three benchmark models i.e., the gamma, Weibull, and log-normal distributions are outperformed by all the priority-based models (Fig. 5). This is interesting since the exact shape of interevent distribution, although in the context of other data sets such as email correspondence, has been an object of dispute [4,22,23]; see also the review from Karsai *et al.* [3]. In particular, it has been argued that for email correspondence a log-normal distribution can better capture interevent distribution [22] but see also Ref. [23]. In our case, it is actually not a surprise that log-normal distribution is outperformed by a power-law distribution. Indeed, as we have seen in Fig. 4, the power-law exponent is on the order of  $\alpha = a + 1 \simeq 1.5$ , whereas the

log-normal distribution has a power-law decay for large  $\tau$  with an exponent of  $\alpha = 1$  which is significantly different from 1.5.

Second, we found that the simplest priority-based models without any refractoriness ( $M_1$  and  $M_2$ ) or with hard refractoriness ( $M_3$  and  $M_4$ ) are outperformed by models with relative refractoriness ( $M_5$  and  $M_6$ ); see Fig. 5. Indeed, despite their relative large number of parameters which penalizes the BIC, the models with relative refractoriness have a better (i.e., lower) BIC than the other models since they better describe short intervals. In particular, we found that the overall best model is  $M_6$  with  $n = 21$  basis functions. When the priority parameter  $b$  is set to one, then the best model of  $M_5$  is when  $n = 20$ . Note that the difference between the difference in BIC between the best model  $M_6$  and the best model  $M_5$  is  $\Delta \text{BIC}_{\text{pop}} = -3.8 \times 10^4$  which is highly significant.

#### D. Short versus long intervals

Given the fairly broad distribution of fitted power-law exponent  $a$  [Fig. 5(b)] and effective time constant  $\tau^*$  [Fig. 5(c)], one could wonder whether this is a fitting artifact (which would come from a fairly flat landscape of the objective function for every subject) or whether the variability of those parameters actually comes from subject-to-subject variability. To test this, we compared the within-subject variability with the between-subject variability of those parameters and found a high degree of fitting consistency between different instantiations (i.e., different number of basis functions) of the same model [see Figs. 6(c)–6(f)].

We then asked whether the effective refractory time constant  $\tau^*$  is correlated with the power-law exponent across different subjects. Note that from the way the model is constructed, those two parameters are a priori unrelated since the refractory affects short intertouch intervals and the power-law exponent affects longer intervals (see Fig. 2). We found that  $a$  and  $\tau^*$  are indeed inversely correlated [Fig. 6(a)] with an explained variance of 40% for model  $M_5$  and 22% for model  $M_6$  [Fig. 6(b)]. This indicates that subjects that have a fast motor control (i.e., have a small  $\tau^*$ ) also put a higher priority on their smartphone (i.e., higher  $a$ ) in the sense that they have less low priority smartphone tasks.

## V. DISCUSSION

We proposed a generalized priority-based model which is both flexible and tractable. The flexibility comes from the set of basis functions which describe refractory effects at short intertouch intervals, while the tractability stems from the simplified structure of the generative model in continuous-time which enables a fast fitting procedure. The flexibility is essential to capture interindividual differences in touching behavior while the tractability is crucial for fitting large data sets. We found that the intertouch intervals are better captured by the priority-based model than by other reference distributions such as log-normal, gamma, or Weibull distributions. We also found that the interindividual differences in low level motor control ability (reflected by the effective refractory time constant) can be partially explained by the higher-level cognitive processes which attributes priority to specific tasks (reflected by the priority parameter  $a$ ).

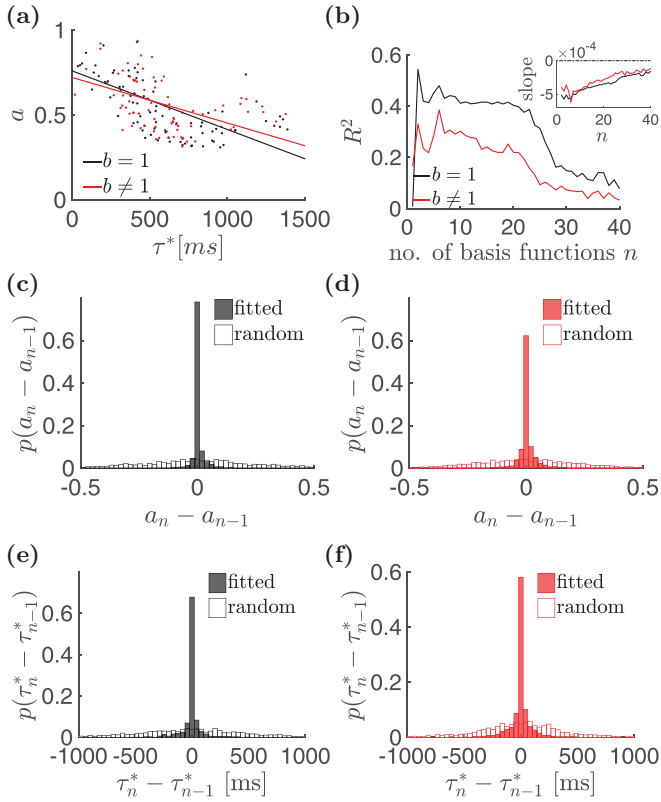


FIG. 6. (a) The scale-free exponent  $a$  is inversely correlated with the effective refractory time constant for model  $M_5$  ( $b = 1$ , black) and  $M_6$  ( $b \neq 1$ , red). (b) Explained variance ( $R^2$ ) for this inverse correlation as a function of the number of basis functions. For the best model  $M_5$ , ( $n = 20$ ),  $R^2 = 40\%$  and for the best model  $M_6$ , ( $n = 21$ ),  $R^2 = 22\%$ . For all numbers of basis function, the slope is negative (see inset). (c–f) consistency of the fitted parameters. (c) The distribution of the difference  $a_n(s) - a_{n-1}(s)$  across subjects  $s = 1, \dots, 84$  and across basis function numbers  $n = 10, \dots, 30$  (black,  $\sigma_{\text{fit}} = 0.022$ ) is much narrower than the distribution of the differences  $a_n(s) - a_{n-1}(s')$  where the indices  $s'$  are randomly permuted (white,  $\sigma_{\text{rand}} = 0.28$ ). (d) Same as panel (c) but for model  $M_6$  ( $\sigma_{\text{fit}} = 0.043$ ,  $\sigma_{\text{rand}} = 0.23$ ). (e) Similarly as in panel (d), the distribution of the difference of effective refractory time constants  $\tau_n^*(s) - \tau_{n-1}^*(s)$  (black,  $\sigma_{\text{fit}} = 88.7$  ms) much narrower than the one obtained when subjects are permuted for each  $n$  (white,  $\sigma_{\text{rand}} = 458$  ms). (f) Same as in panel (e) but for model  $M_6$  ( $\sigma_{\text{fit}} = 100$  ms,  $\sigma_{\text{rand}} = 366$  ms).

Other models describing smartphone activity have been proposed. However, they are aimed at addressing different questions (mostly sleep related) and use different type of data. The model of Cuttone *et al.* [24] aims at predicting sleep patterns and rely on app launch timings binned over 15-min duration and therefore lack the possibility to describe refractory effects on the tens of milliseconds resolution. Abdullah *et al.* [25] only considered screen-on and screen-off events to predict sleep patterns.

More generally, circadian rhythms have received recently a lot of attention [26–28] and the question has been addressed whether circadian rhythms could explain heavy tail distribution [14] has argued that a cascade of Poisson processes can give rise to power-law distribution. Interestingly, Jo *et al.* [19]

showed that even if the data is deseasoned (i.e., the circadian and weekly patterns are removed from the time series of mobile phone events), the heavy-tails remain.

Closer to the present study, the priority model proposed by Refs. [2,4,7] already predicts a power-law distribution. However, our model deviates significantly from their approach in several aspects.

First, and most importantly, the priority-based model from Barabási is aimed at predicting the response time distribution (e.g., the time interval between the reception of a letter and its response), whereas in our case, our priority-based model predicts interevent times (e.g., the time interval between two consecutive posted letters—or in our case intertouch intervals). It is, however, interesting to note that even though the response-time and the interevent-time metrics are clearly different, they can still be seen within the same umbrella. On the formal level, the response-time priority-based model with a list of  $L = 2$  is equivalent to the discrete-time emission model described in Sec. II A. Furthermore, on the interpretation level, since a new smartphone touching task arrives just after every smartphone touching event [see Eq. (1)], one could still reinterpret the following interevent time as a response time. For a further discussion on the relationship between response-time distribution and interevent distribution, see Ref. [3].

Second, on the conceptual level, those authors stress the universality of the various behaviors. Their claim is that the interevent intervals for certain activities such as browsing the Web, sending emails or loaning books fall into a specific universality class with power-law exponent of  $\alpha = 1$  while the response time for other activities such as writing mails follow another universality class with exponent of  $\alpha = 3/2$ . Actually, more recently, Formentin *et al.* [29] elegantly showed that various response-time datasets (on sms [30], email [29], and mail [1]), with apparently different power-law exponents can be cast within the same  $\alpha = 3/2$  universality class provided that the events are properly relocked [31]. Here, we found that the power-law exponent (averaged over the population) is  $\alpha = a + 1 \simeq 1.56 \pm 0.16$  which is indeed close to the rational exponent of  $3/2$ . However, it should be noted that the power-law exponent of individuals are fairly spread ranging from  $\alpha = 1.31 \pm 0.04$  to  $\alpha = 2.19 \pm 0.04$  which are clearly different from  $\alpha = 3/2$ . Capturing those nonuniversal exponents is possible in our model since the power-law exponent is given by  $a + 1$  where  $a$  can take any real positive value. In contrast, in the work of Oliveira *et al.* [7], the exponent is determined by the length of the list of tasks [32].

The third difference w.r.t. the studies of Refs. [2,4,7] is that our model has been actually fitted to the whole set of event times (the touch times) such that we did not neglect small interevent intervals by defining a (somewhat artificial) onset of the power-law distribution [20]. This is possible in our model since short intervals are captured by the refractory kernel. Note that even though refractory kernels have been used in other fields (e.g., in spiking neuron models, the probability of generating a spike just after a first one is also modulated by a refractory kernel [33–35]), the particularity here is that the specific form of the refractory kernel is such that its integral can be computed analytically which boosts the computational efficiency.

Finally, fitting our model to the touching data has been possible because we considered the continuous-time priority

model. Indeed, the marginal likelihood can be expressed analytically for the continuous-time model (and not for the discrete time model) which makes the maximum-likelihood parameter learning extremely efficient.

A separate line of research based on biological signals has also encountered scale-invariant relationships referred as  $1/f$  pink noise [36,37]. However, those studies compute the power-spectrum density and not the interevent distribution. Actually, if we do compute the power-spectrum density for the smartphone touching model, then we find that in the limit of large frequencies, the power-spectrum density remains constant and does not decrease as  $1/f$ .

Here, this generalized priority-based model has been applied to smartphone touching data but could be applied to other event-based data sets which display power-law property for large interevent intervals such as surface mails, emails, or even foraging patterns.

The simulations have been performed using Matlab. The code and the smartphone touching data are available online [38].

**ACKNOWLEDGMENTS**

We thank Simone Carlo Surace for very helpful discussions. J.P.P. was supported by the Swiss National Science Foundation Grants No. PP00P3\_150637 and No. PP00P3\_179060. A.G. was supported by the Society in Science, the Branco Weiss Fellowship, and a research grant from the Holcim Foundation.

J.P.P. and A.G. designed the study, developed the model used here, and drafted the manuscript. J.P.P. analyzed the data and formulated the mathematical model, aided by A.G., who collected the data and conceived the project with J.P.P.

A.G. is an inventor of the patent-pending technology used to track touchscreen interactions in this study. A.G. and J.P.P. are cofounders of QuantActions GmbH, a company focused on quantifying human behavior through smartphone interactions.

**APPENDIX A**

**1. Invariance of the model**

In this section, we will show that that the ITI distribution remains unchanged if the pair of priority distribution  $[p(x), q(y)]$  is replaced by  $[\tilde{p}(x), \tilde{q}(y)]$ , where  $\tilde{p}(x)$  and  $\tilde{q}(y)$  are given by Eq. (13).

Let us consider the following change of variable:  $x = \phi(x')$ . The ITI distribution can be therefore expressed as

$$p(\tau) = \int_0^1 p^q[\tau|\phi(x')]p[\phi(x')]\phi'(x')dx', \tag{A1}$$

where the conditional ITI distribution  $p^q[\tau|\phi(x')]$  depends on the *other* priority distribution  $q(y)$  via the instantaneous rate  $\bar{\lambda}^q[\phi(x'), \tau]$  which can be expressed as

$$\begin{aligned} \bar{\lambda}^q[\phi(x'), \tau] &= \rho r(\tau) \int_0^{\phi(x')} q(y)dy \\ &= \rho r(\tau) \int_0^{x'} q[\phi(y)]\phi'(y)dy = \bar{\lambda}^{\tilde{q}}(x', \tau), \end{aligned} \tag{A2}$$

where  $\tilde{q}$  is given by Eq. (13). Note that the dependence on  $q$  is included only here for the clarity of the argument, but is omitted otherwise for the simplicity of the notation. Therefore, the ITI distribution is invariant under the change of variable  $\phi$  for both  $x$  and  $y$ . Indeed, we have

$$\begin{aligned} p(\tau) &= \int_0^1 p^q(\tau|x)p(x)dx \\ &= \int_0^1 p^{\tilde{q}}(\tau|x)\tilde{p}(x)dx. \end{aligned} \tag{A3}$$

For example, if the *touch* priority distribution is given by  $p(x) = \beta(x; a, 1)$  and the *other* priority distribution is given by  $q(y) = \beta(y; a', 1)$ , then the function  $\phi(x) = x^k$  allows us to generate a family of equivalent pairs of priority distributions  $[\tilde{p}(x), \tilde{q}(y)] = [\beta(x; ka, 1), \beta(y; ka', 1)]$ . Therefore, the ITI remains unchanged as long as  $a/a'$  remains constant.

Note that this argument can be generalized to arbitrary smooth distribution  $\tilde{p}$  and  $\tilde{q}$ . Let  $\tilde{A}(x)$  denote the ratio of the logarithm of both cumulative density functions  $\tilde{P}(x) = \int_0^x x' \tilde{p}(x')dx'$  and  $\tilde{Q}(x) = \int_0^x x' \tilde{q}(x')dx'$ :

$$\tilde{A}(x) = \frac{\log[\tilde{P}(x)]}{\log[\tilde{Q}(x)]}. \tag{A4}$$

Since  $\tilde{p}$  and  $\tilde{q}$  are smooth, when  $x \rightarrow 0$ , we can express those priority distribution as  $p(x) \simeq c_1x^{a-1}$  and  $q(y) = c_2y^{a'-1}$ . Also, since  $\phi(x)$  is smooth, it can be approximated as  $\phi(x) = x^k$  in the vicinity of  $x = 0$ . Now we can show that the  $\tilde{A}(0)$  is independent of the change of variable function  $\phi$ . Indeed,

$$\begin{aligned} \tilde{A}(0) &= \lim_{x \rightarrow 0} \frac{\log[x\tilde{p}(x)]}{\log[x\tilde{q}(x)]} \\ &= \lim_{x \rightarrow 0} \frac{\log\{p[\phi(x)]\} + \log[\phi'(x)] + \log(x)}{\log\{q[\phi(x)]\} + \log[\phi'(x)] + \log(x)} \\ &= \lim_{x \rightarrow 0} \frac{\log(c_1) + ka \log(x)}{\log(c_2) + ka' \log(x)} = \frac{a}{a'} \end{aligned} \tag{A5}$$

is independent of  $k$ .

**2. Log-likelihood gradient**

For the models with relative refractoriness, we fitted the parameters  $\theta = (a, b, c, \gamma_1, \dots, \gamma_n)$  by performing maximum likelihood with a suitable regularization for the parameters  $\gamma_i$ . Note that for a practical implementation, it is easier to learn  $c = \log(\rho)$  instead of  $\rho$  itself. For a set of intertouch intervals  $\mathcal{D} = \{\tau_i\}_{i=1}^N$ , the log-likelihood can be expressed as

$$L(\theta) = Nc + \sum_{i=1}^N \log[r(\tau_i)] + \log(\langle xE_i(x) \rangle), \tag{A6}$$

where the expectation  $\langle \cdot \rangle$  is w.r.t.  $p(x) = \beta(x; a, b)$  and the function  $E_i(x)$  is given by

$$E_i(x) = e^{-\rho x R(\tau_i)}, \tag{A7}$$

and  $R(\tau_i)$  is given by

$$R(\tau_i) := \int_0^{\tau_i} r(t)dt = \tau_i + \sum_{k=1}^n \frac{\gamma_k}{\alpha_k} (1 - e^{-\alpha_k \tau_i}). \tag{A8}$$



By noting that

$$\frac{\partial \log[p(x)]}{\partial a} = \log(x) - \langle \log(x) \rangle, \quad (\text{A9})$$

we can compute the log-likelihood gradient w.r.t.  $a$ :

$$\frac{\partial L}{\partial a} = \sum_{i=1}^N \frac{\text{cov}[xE_i(x), \log(x)]}{\langle xE_i(x) \rangle}. \quad (\text{A10})$$

By symmetry, the gradient of  $L$  w.r.t. to  $b$  yields

$$\frac{\partial L}{\partial b} = \sum_{i=1}^N \frac{\text{cov}[xE_i(x), \log(1-x)]}{\langle xE_i(x) \rangle}. \quad (\text{A11})$$

The gradient of  $L$  w.r.t.  $c$  is given by

$$\frac{\partial L}{\partial c} = N - \rho \sum_{i=1}^N \frac{\langle x^2 E_i(x) \rangle}{\langle xE_i(x) \rangle} R(\tau_i), \quad (\text{A12})$$

Finally, the gradient of  $L$  w.r.t.  $\gamma_k$  can be expressed as

$$\begin{aligned} \frac{\partial L}{\partial \gamma_k} &= \sum_{i=1}^N \frac{\partial r(\tau_i)/\partial \gamma_k}{r(\tau_i)} - \rho \frac{\langle x^2 E_i(x) \rangle}{\langle xE_i(x) \rangle} \frac{\partial R(\tau_i)}{\partial \gamma_k} \\ &= \sum_{i=1}^N \frac{e^{-\alpha_k \tau_i}}{r(\tau_i)} - \rho \frac{\langle x^2 E_i(x) \rangle}{\langle xE_i(x) \rangle} \frac{(1 - e^{-\alpha_k \tau_i})}{\alpha_k}. \end{aligned} \quad (\text{A13})$$

For the models with hard refractoriness, the integral over the refractory kernel is given by

$$R(\tau_i) = \tau_i - \Delta \quad (\text{A14})$$

when  $0 \leq \Delta < \tau_{\min}$  where  $\tau_{\min} = \min_i \tau_i$  is the minimal inter-tap interval. Under this condition, the log-likelihood gradient yields

$$\frac{\partial L}{\partial \Delta} = \rho \frac{\langle x^2 E_i(x) \rangle}{\langle xE_i(x) \rangle}, \quad (\text{A15})$$

which is positive as long as  $\Delta \leq \tau_{\min}$ . When  $\Delta > \tau_{\min}$ , the log-likelihood goes to  $-\infty$ . Therefore, the optimal refractory time constant is  $\Delta^* = \tau_{\min}$ .

### 3. Computing the integrals

Both the log-likelihood  $L$  as well its gradient w.r.t. to the parameters  $\theta$  contain integrals that are delicate to evaluate. Indeed, the integrand of all those integrals depend on the beta distribution  $\beta(x; a, b)$  which can diverge at  $x = 0$  or  $x = 1$  depending on the parameters  $a$  and  $b$ . So whenever possible, we compute those integrals analytically. This can be done for the following integrals:

$$\langle \log(x) \rangle_{a,b} = \frac{d}{da} B(a, b) = \psi(a) - \psi(a+b), \quad (\text{A16})$$

where  $\psi(z) = d \log \Gamma(z)/dz$  is the digamma function and  $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$  is the beta function. By symmetry, we have

$$\langle \log(1-x) \rangle_{a,b} = \psi(b) - \psi(a+b). \quad (\text{A17})$$

By Taylor expanding the exponential in the expression of  $E_i(x)$ , the integral  $\langle xE_i(x) \rangle_{a,b}$  can be expressed as

$$\begin{aligned} \langle xE_i(x) \rangle_{a,b} &= \frac{a}{a+b} \langle E_i(x) \rangle_{a+1,b} \\ &= \frac{a}{a+b} {}_1F_1[a+1, a+b+1; -\rho R(\tau_i)], \end{aligned} \quad (\text{A18})$$

where  ${}_1F_1$  is the hypergeometric function defined as

$${}_1F_1(a, b; z) = \sum_{k=0}^{\infty} \frac{z^k}{k!} \frac{(a)_k}{(b)_k} \quad (\text{A19})$$

and  $(a)_k = \prod_{i=0}^{k-1} (a+i)$  for  $k \geq 1$  [and  $(a)_0 = 1$ ] is the rising factorial (also called Pochhammer function). Similarly,  $\langle x^2 E_i(x) \rangle_{a,b}$  can be expressed as

$$\langle x^2 E_i(x) \rangle_{a,b} = \frac{B(a+2, b)}{B(a, b)} {}_1F_1[a+2, a+b+2; -\rho R(\tau_i)]. \quad (\text{A20})$$

When it is not possible to compute the integrals analytically, the idea is to express the integral as a sum of two integrals where the first one is well suited for a numerical integration and the second one can be performed analytically. For example,  $\langle xE_i(x) \log(1-x) \rangle_{a,b}$  can be computed as

$$\begin{aligned} &\langle xE_i(x) \log(1-x) \rangle_{a,b} \\ &= \frac{a}{a+b} \{ \langle (E_i(x) - E_i(1)) \log(1-x) \rangle_{a+1,b} \\ &\quad + E_i(1) \langle \log(1-x) \rangle_{a+1,b} \}, \end{aligned} \quad (\text{A21})$$

where the first term of the r.h.s can be computed numerically and the second term can be computed with Eq. (A17).

Finally, it should be noted that the integral  $\langle xE_i(x) \log(x) \rangle_{a,b}$  can be computed numerically straightforwardly since the integrand does not diverges when  $x = 0$  nor when  $x = 1$ .

### 4. Benchmark distributions

We benchmarked the generalized priority-based model is benchmarked with three simple models: the gamma distribution, the Weibull distribution, and the log-normal distribution. The gamma distribution is given by

$$p(\tau) = \frac{e^{-\tau/\vartheta} \tau^{\kappa-1}}{\Gamma(\kappa) \vartheta^\kappa}, \quad (\text{A22})$$

where  $\kappa$  is the shape parameter and  $\vartheta$  is the scale parameter. The log-likelihood is therefore expressed as a function of a two-dimensional parameter  $\theta = (\kappa, \vartheta)$  and is given by

$$L(\theta) = N \left( (\kappa - 1) \langle \log(\tau) \rangle_\tau - \frac{\langle \tau \rangle_\tau}{\vartheta} - \kappa \log(\vartheta) - \log(\Gamma(\kappa)) \right), \quad (\text{A23})$$

where the expectation symbol here denotes the average over the data points, i.e.,  $\langle f(\tau) \rangle_\tau \equiv N^{-1} \sum_{i=1}^N f(\tau_i)$ . The maximum-likelihood parameters  $\theta^*$  can be found by maximising Eq. (A23). This can be done easily by taking advantage of the fact that the optimal scale parameter has to obey  $\vartheta^* = \langle \tau \rangle_\tau / \kappa$ . So the shape parameter can be found by evaluating

the gradient of the log-likelihood w.r.t. to  $\kappa$  at  $\vartheta^*$  and setting the gradient to zero. This amounts to finding the root of

$$\langle \log(\tau) \rangle_\tau - \log \left( \frac{\langle \tau \rangle_\tau}{\kappa} \right) - \psi(k) = 0, \quad (\text{A24})$$

where  $\psi(k)$  is the digamma function. The Weibull distribution is given by

$$p(\tau) = \frac{\kappa}{\lambda} \left( \frac{\tau}{\lambda} \right)^{\kappa-1} e^{-\left(\frac{\tau}{\lambda}\right)^\kappa}, \quad (\text{A25})$$

where  $\kappa$  is the shape parameter and  $\lambda$  is the scale parameter. The log-likelihood is therefore given by

$$L(\theta) = N \left( \log \frac{\kappa}{\lambda} + (\kappa - 1) \left\langle \log \left( \frac{\tau}{\lambda} \right) \right\rangle_\tau - \frac{\langle \tau^\kappa \rangle_\tau}{\lambda^\kappa} \right), \quad (\text{A26})$$

with  $\theta = (\kappa, \lambda)$ . By observing that the maximising scale parameter is given by  $\lambda^* = \sqrt[\kappa]{\langle \tau^\kappa \rangle_\tau}$ , we can obtain the shape parameter  $k$  by computing the root of

$$\frac{\langle \tau^\kappa \log(\tau) \rangle_\tau}{\langle \tau^\kappa \rangle_\tau} - \langle \log(\tau) \rangle_\tau - \frac{1}{\kappa} = 0. \quad (\text{A27})$$

Finally, the log-normal distribution is given by

$$p(\tau) = \frac{1}{\tau \sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{[\log(\tau) - \mu]^2}{2\sigma^2} \right\}, \quad (\text{A28})$$

where the maximum likelihood parameters are simply given by

$$\mu^* = \langle \log(\tau) \rangle_\tau, \quad (\text{A29})$$

$$(\sigma^*)^2 = \langle (\log(\tau) - \mu^*)^2 \rangle_\tau. \quad (\text{A30})$$

## APPENDIX B

### 1. Smartphone data collection

A custom-designed software application called Touchometer (now available as TapCounter, QuantActions, Lausanne) that could record the touchscreen events with a maximum error of 5 ms [39] was installed on each participant's phone. To determine this accuracy, controlled test touches were done at precisely 150, 300, and 600 ms while the Touchometer recorded at 147, 301, and 600 ms, respectively, with standard deviations less than 15 ms (interquartile range less than 5 ms). The app posed as a service to gather the timestamps of touchscreen events that were generated when the screen was in an unlocked state. The operation was verified in a subset of phones by using visually monitored tactile events. The data were stored locally and transmitted by the user at the end of the study via secure email. One subject was eliminated as the app intermittently crashed after a software update. The smartphone data were processed by using MATLAB (MathWorks, USA).

- 
- [1] J. G. Oliveira and A.-L. Barabási, *Human Dynamics: Darwin and Einstein Correspondence Patterns*, Vol. 437 (Nature, Berlin, 2005).
- [2] A. Vázquez, J. G. Oliveira, Z. Dezső, K.-I. Goh, I. Kondor, and A.-L. Barabási, *Phys. Rev. E* **73**, 036127 (2006).
- [3] M. Karsai, H.-H. Jo, and K. Kaski, *Bursty Human Dynamics*, SpringerBriefs in Complexity (Springer International Publishing, Cham, 2018).
- [4] A.-L. Barabási, *Nature* **435**, 207 (2005).
- [5] G. Grinstein and R. Linsker, *Phys. Rev. Lett.* **97**, 130201 (2006).
- [6] G. Grinstein and R. Linsker, *Phys. Rev. E* **77**, 012101 (2008).
- [7] J. G. Oliveira and A. Vazquez, *Physica A* **388**, 187 (2009).
- [8] N. Masuda, J. S. Kim, and B. Kahng, *Phys. Rev. E* **79**, 036106 (2009).
- [9] B. Min, K. I. Goh, and I. M. Kim, *Phys. Rev. E* **79**, 056110 (2009).
- [10] O. Mryglod, Y. Holovatch, and I. Mryglod, *Scientometrics* **91**, 101 (2011).
- [11] S. Vajna, B. Tóth, and J. Kertész, *New J. Phys.* **15**, 103023 (2013).
- [12] C. A. Hidalgo R., *Physica A: Stat. Mech. Appl.* **369**, 877 (2006).
- [13] A. Vázquez, *Physica A: Stat. Mech. Appl.* **373**, 747 (2007).
- [14] R. D. Malmgren, D. B. Stouffer, A. E. Motter, and L. A. N. Amaral, *Proc. Natl. Acad. Sci. USA* **105**, 18153 (2008).
- [15] J. L. Guo, C. Fan, and Z. H. Guo, *Eur. Phys. J. B* **81**, 341 (2011).
- [16] A. Hawkes, *J. Roy. Stat. Soc. Ser. B (Methodol.)* **33**, 438 (1971).
- [17] M. Gilson and J.-P. Pfister, *SIAM J. Appl. Dynam. Syst.* **19**, 828 (2020).
- [18] N. Masuda, T. Takaguchi, N. Sato, and K. Yano, *Temporal Networks*, Understanding Complex Systems, edited by P. Holme and J. Saramäki (Springer, Berlin, Heidelberg, 2013), p. 245.
- [19] H.-H. Jo, M. Karsai, J. Kertész, and K. Kaski, *New J. Phys.* **14**, 013055 (2012).
- [20] A. Clauset, C. R. Shalizi, and M. E. J. Newman, *SIAM Rev.* **51**, 661 (2009).
- [21] P. Brémaud, *Point Processes and Queues*, Vol. 30 (Springer, Berlin, 1981).
- [22] D. B. Stouffer, R. D. Malmgren, and L. A. Amaral, [arXiv:physics/0605027](https://arxiv.org/abs/physics/0605027).
- [23] A. L. Barabási, K. I. Goh, and A. Vazquez, [arXiv:physics/0511186](https://arxiv.org/abs/physics/0511186).
- [24] A. Cuttone, P. Bækgaard, V. Sekara, H. Jonsson, J. E. Larsen, and S. Lehmann, *PLoS ONE* **12**, e0169901 (2017).
- [25] S. Abdullah, M. Matthews, E. L. Murnane, G. Gay, and T. Choudhury, in *Proceedings of the ACM International Joint Conference* (ACM Press, New York, 2014), pp. 673–684.
- [26] T. Aledavood, E. López, S. G. B. Roberts, F. Reed-Tsochas, E. Moro, R. I. M. Dunbar, and J. Saramäki, *PLoS ONE* **10**, e0138098 (2015).
- [27] T. Aledavood, S. Lehmann, and J. Saramäki, *Front. Phys.* **3**, 15602 (2015).
- [28] T. Aledavood, S. Lehmann, and J. Saramäki, *EPJ Data Sci.* **7**, 46 (2018).
- [29] M. Formentin, A. Lovison, A. Maritan, and G. Zanzotto, *Phys. Rev. E* **90**, 012817 (2014).
- [30] Y. Wu, C. Zhou, J. Xiao, J. Kurths, and H. J. Schellnhuber, *Proc. Natl. Acad. Sci. USA* **107**, 18803 (2010).

- [31] Note also that this reclocking has the additional benefit of reducing the bimodality present in the response-time distribution which can be seen especially in the sms data [30] and in the email data set [29] as well as in our own intertouch interval data; see Fig. 4(b).
- [32] Technically, Ref. [7] assumes that an event occurs only if the interacting task for both agents A and B have the highest priority compared to all other tasks of length  $L_A$  (for agent A) and  $L_B$  for agent B. If  $L = L_A = L_B$ , then the power-law exponent is given by  $1 + 1/(L - 1)$ .
- [33] J. W. Pillow, J. Shlens, L. Paninski, A. Sher, A. M. Litke, E. J. Chichilnisky, and E. P. Simoncelli, *Nature* **454**, 995 (2008).
- [34] W. Gerstner, W. M. Kistler, R. Naud, and L. Paninski, *Neuronal Dynamics, From Single Neurons to Networks and Models of Cognition* (Cambridge University Press, Cambridge, UK, 2014).
- [35] S. C. Surace and J.-P. Pfister, *PLoS ONE* **10**, e0142435 (2015).
- [36] D. L. Gilden, T. Thornton, and M. W. Mallon, *Science* **267**, 1837 (1995).
- [37] Y. Chen, M. Ding, and J. A. Scott Kelso, *Phys. Rev. Lett.* **79**, 4501 (1997).
- [38] [github.com/jeanpascalpfister/SmartphoneTappingModel/](https://github.com/jeanpascalpfister/SmartphoneTappingModel/).
- [39] M. Balerna and A. Ghosh, *npj Dig. Med.* **1**, 4 (2018).