Phase Transition in Non-Markovian Animal Exploration Model with Preferential Returns – Supplementary Material

Ohad Vilk[a,b,c], Daniel Campos[d], Vicenç Méndez[d], Emmanuel Lourie[b,c], Ran Nathan[b,c], Michael Assaf[a,e,*]
[a] *Racah Institute of Physics, The Hebrew University of Jerusalem, Jerusalem 91904, Israel.*
[b] *Movement Ecology Lab, Department of Ecology, Evolution and Behavior, Alexander Silberman Institute of Life Sciences, Faculty of Science, The Hebrew University of Jerusalem, Jerusalem 91904, Israel.*
[c] *Minerva Center for Movement Ecology, The Hebrew University of Jerusalem, Jerusalem 91904, Israel.*
[d] *Grup de Física Estadística, Dept. de Física, Universitat Autónoma de Barcelona, 08193 Bellaterra (Barcelona), Spain. and*
[e] *Institute for Physics and Astronomy, University of Potsdam, Potsdam 14476, Germany**
(Dated: March 7, 2022)

Here we provide additional details and results to support the derivations presented in the main text. Below, the notations and acronyms are the same as in the main text and the equations, figures and citations refer to those therein.

S1. CUMULATIVE NUMBER OF SITES

In this section we detail the exact solution and the WKB approximation to Eq. (2) of the main text

$$\frac{dP(n,t)}{dt} = q(n-1)^{-\beta}P(n-1,t) - qn^{-\beta}P(n,t). \tag{S1}$$

A. Solution by Laplace transform

Here we derive Eqs. (4) and (5) of the main text by Laplace transforming Eq. (S1) and solving the resulting recurrence equation. First we define $J(n,t) = qP(n,t)/n^{\beta}$, which turns Eq. (S1) into:

$$\frac{n^{\beta}}{q}\frac{dJ(n,t)}{dt} = J(n-1,t) - J(n,t). \tag{S2}$$

Transforming (S2) by Laplace in time and considering the initial condition $P(n,t=0) = \delta_{n,1}$, where $\delta_{a,b}$ is the Kronecker delta, we obtain the recurrence equation

$$\hat{J}(n,s) = \frac{1}{1+\frac{sn^{\beta}}{q}}\hat{J}(n-1,s) + \frac{1}{1+\frac{sn^{\beta}}{q}}\delta_{n,1}, \tag{S3}$$

where s is the Laplace variable and $\hat{J}(n,s)$ stands for the Laplace transform of $J(n,t)$ defined as follows:

$$\hat{J}(n,s) = \mathcal{L}_s\left[J(n,t)\right] = \int_0^{\infty} e^{-st}J(n,t)dt.$$

Multiplying (S3) by $\prod_{j=0}^{n}\left(1+\frac{sj^{\beta}}{q}\right)$ it has the form

$$A(n,s) = A(n-1,s) + \delta_{n,1}\prod_{j=0}^{n-1}\left(1+\frac{sj^{\beta}}{q}\right), \tag{S4}$$

where $A(n,s) = \hat{J}(n,s)\prod_{j=0}^{n}\left(1+\frac{sj^{\beta}}{q}\right)$ has been introduced. Since $P(n\leq 0,t)=0$ one has $A(0,s)=0$ and using (S4), $A(n\geq 1,s)=1$. Inserting this result into (S3) the solution to the master equation in the Laplace domain reads

$$\hat{P}(n,s) = \frac{n^{\beta}}{q\prod_{j=0}^{n}\left(1+\frac{sj^{\beta}}{q}\right)} = \frac{q^{n-1}}{[(n-1)!]^{\beta}}\frac{1}{\prod_{j=1}^{n}\left(s+\frac{q}{j^{\beta}}\right)}. \tag{S5}$$

* To whom correspondence should be addressed: michael.assaf@mail.huji.ac.il

Now we need to invert (S5) by Laplace. To do this we use the Heaviside expansion theorem [59]

$$\mathcal{L}_t^{-1}\left[\frac{1}{f(s)}\right]=\sum_{k=1}^{n}\frac{e^{-a_k t}}{f'(s=-a_k)},\qquad f(s)=\prod_{j=1}^{n}(s+a_j), \tag{S6}$$

where the prime symbol stands for the derivative with respect to s. Making use of the property

$$f'(s=a_k)=\prod_{j=1,j\neq k}^{n}(a_j-a_k),$$

and (S6), one readily obtains for any $\beta\neq 0$

$$\mathcal{L}_t^{-1}\left[\frac{1}{\prod_{j=1}^{n}\left(s+\frac{q}{j^\beta}\right)}\right]=\sum_{k=1}^{n}\frac{e^{-qt/k^\beta}}{\prod_{j=1,j\neq k}^{n}\left(\frac{q}{j^\beta}-\frac{q}{k^\beta}\right)}=(-1)^{n-1}(n!)^\beta q^{1-n}\sum_{k=1}^{n}\frac{k^{-\beta}e^{-qt/k^\beta}}{\prod_{j=1,j\neq k}^{n}\left(\frac{j^\beta}{k^\beta}-1\right)}. \tag{S7}$$

Finally, plugging (S7) into (S5) the exact solution for $\beta\neq 0$ has the form

$$P(n,t)=(-1)^{n-1}n^\beta\sum_{k=1}^{n}\frac{k^{-\beta}e^{-qt/k^\beta}}{\prod_{j=1,j\neq k}^{n}\left(\frac{j^\beta}{k^\beta}-1\right)}, \tag{S8}$$

which is Eq. (4) of the main text, and is valid for any $\beta\neq 0$. For $\beta=0$, Eq. (S5) reduces to

$$\hat{P}(n,s)=\frac{1}{q\left(1+\frac{s}{q}\right)^n},$$

which after inversion by Laplace coincides with Eq. (5) of the main text

$$P(n,t)=\frac{(qt)^{n-1}}{(n-1)!}e^{-qt}. \tag{S9}$$

B. Solution by generating function

In some special cases it is more convenient to solve Eq. (S1) using the generating function approach. Particularly, in the limit of a large number of sites, the equation in the generating function domain becomes a fractional integro-differential equation for $0<\beta\leq 1$, see below.

We define the generating function $G(z,t)=\sum_n z^n P(n,t)$, such that

$$P(n,t)=\frac{1}{n!}\frac{\partial^n G(z,t)}{\partial z^n}|_{z=0}, \tag{S10}$$

with initial condition $P(n,0)=\delta_{n,1}\Rightarrow G(n,0)=z$. Substituting Eq. (S1) into the definition of $G(n,t)$ yields

$$\begin{aligned}\frac{\partial G}{\partial t}&=q\sum_n\left[z^n(n-1)^{-\beta}P(n-1,t)-z^n n^{-\beta}P(n,t)\right]=q(z-1)\sum_n z^n n^{-\beta}P(n,t)\\&\simeq q(z-1)\sum_n D_z^{-\beta}z^{n-\beta}P(n,t)=q(z-1)D_z^{-\beta}\left[z^{-\beta}G(z,t)\right],\end{aligned} \tag{S11}$$

where $D_z^{-\beta}$ is the Riemann-Liouville fractional integral defined by ${}_aD_z^{-\beta}f(z)=\frac{1}{\Gamma(\beta)}\int_a^z(z-\xi)^{\beta-1}f(\xi)d\xi$, and we used the following relations

$$D_z^{-\beta}z^{n-\beta}=z^n\frac{\Gamma(n-\beta+1)}{\Gamma(n+1)}=z^n n^{-\beta}[1+\mathcal{O}(1/n)]. \tag{S12}$$

Here, the equality on the left is valid for $\beta\leq 1$ while the approximation on the right holds for $n\gg 1$ and is exact in the special cases of $\beta=0,1$. Notably, for $\beta<0$, *i.e.*, when the growth rate *increases* in n, Eq. (S11) is a fractional differential equation for G, while for $\beta>0$, *i.e.*, when the growth rate *decreases* in n, it is a fractional integro-differential equation. Although Eq. (S11) is hard to solve analytically for general β, and the direct method presented above is more suited, it can be solved for $\beta=-1,0,1$.

1. Solution for $\beta = 0$

For $\beta = 0$, Eq. (S11) simplifies to $\partial_t G(z,t) = q(z-1)G(z,t)$ and is accordingly solved by $G(z,t) = ze^{-qt(1-z)}$. Using Eq. (S10) we find that $P(n,t)$ follows a Poisson distribution (S9). In particular, in this case we have $\langle n \rangle = qt + 1 \simeq qt$ and $\sigma_n^2 = qt$.

2. Solution for $\beta = -1$

For $\beta = -1$ a similar derivation to Eq. (S11) yields a partial differential equation for the generating function: $\partial_t G(z,t) = q(z-1)z\partial_z[G(z,t)]$, whose solution is $G(z,t) = z/[z + e^{qt}(1-z)]$. Using Eq. (S10) we find

$$P(n,t) = e^{-nqt}\left(e^{qt}-1\right)^{n-1}. \tag{S13}$$

The average number of sites here is $\langle n \rangle = e^{qt}$, *i.e.*, we find exponential growth, as expected for a growth rate that is linear in n. Here, the variance is $\sigma_n^2 = e^{qt}\left(e^{qt}-1\right) \simeq e^{2qt} = \langle n \rangle^2$, which is significantly broader than that of the Poisson distribution. This result also agrees with Eq. (5) of the main text.

3. Solution for $\beta = 1$

For $\beta = 1$, Eq. (S11) is rewritten in explicit integro-differential form:

$$\frac{\partial G}{\partial t} = q(z-1)\int_0^z y^{-1}G(y,t)dy. \tag{S14}$$

Here, we Laplace transform Eq. (S14) in time

$$uG(z,u) = z + q(z-1)\int_0^z y^{-1}G(y,u)dy, \tag{S15}$$

to obtain an integral equation. This equation can be solved iteratively by the Neumann series method [60] to give:

$$G(z,u) = \frac{1}{u}\left[(z-1)e^{\frac{qz}{u}}\left(\frac{qz}{u}\right)^{-\frac{q}{u}}\gamma\left(\frac{p+u}{u},\frac{qz}{u}\right)+z\right], \tag{S16}$$

where $\gamma(\cdot,\cdot)$ is the lower gamma function. Using Eq. (S10) we can inverse Laplace transform Eq. (S16) to obtain

$$P(n,t) = \frac{1}{(n-1)!}\sum_{k=1}^{n}(-1)^{n-k}k^{n-1}\binom{n}{k}e^{-\frac{qt}{k}}, \tag{S17}$$

in agreement with Eq. (5) of the main text.

C. Time dependent WKB approximation

Here we derive Eqs. (7) and (8) of the main text. We employ the time-dependent WKB approximation in the limit of a large number of sites $n \gg 1$ [38, 39]. Substituting the time-dependent ansatz $P(n,t) \sim e^{-S(n,t)}$ into Eq. (S1) and neglecting terms of order $\mathcal{O}(n^{-1})$ we obtain a classical Hamilton-Jacobi equation for the action function $S(n,t)$:

$$\frac{\partial S}{\partial t} = H(n,\frac{\partial S}{\partial n}) \equiv H(n,p)\,, \qquad H(n,p) = q\left(1-e^{-p}\right)n^{-\beta}, \tag{S18}$$

where H is the Hamiltonian and $p = -\partial_n S$ is the conjugate momentum. Instead of directly solving the Hamilton-Jacobi equations, we use the Hamilton approach for the classical equations of motion [Eq. (6)]

$$\dot{n} = qe^{-p}n^{-\beta}\,, \quad \dot{p} = \beta q\left(1-e^{-p}\right)n^{-\beta-1}. \tag{S19}$$

We write the action on a classical trajectory as [38]:

$$S = Et - \int_0^t p\dot{n}dt = Et - \int^n p(n')dn' \tag{S20}$$

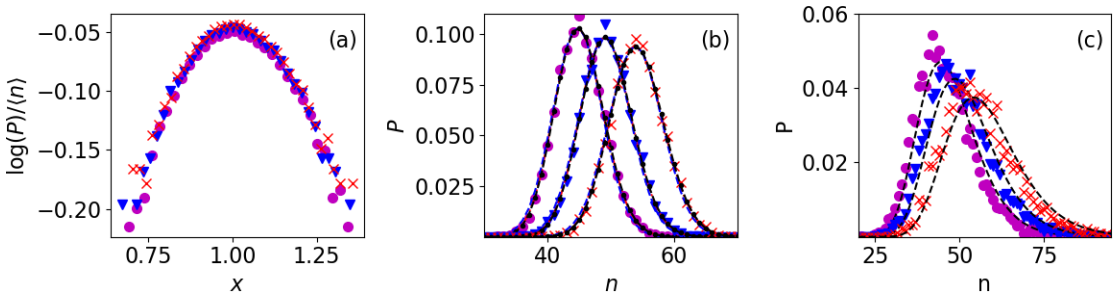

FIG. S1. (a) The action $S(x) = \log(P)/\langle n\rangle$ as a function of $x = n/\langle n\rangle$ and (b) the probability $P(n,t)$ as a function of n. In both panels $\beta = 1$, $q = 1$, $\sigma = 0$ and $t = 1000, 1250$ and 1500 (circles, triangles and x marks respectively). In (a) only simulation results are plotted showing that the exact result and WKB approximation have a similar scaling with time and that pre-exponential corrections are insignificant, see text. In (b) we compare simulations (symbols), exact result [black dashed-dotted line, Eq. (S17)], WKB approximation [red dashed lines, Eq. (S23)], and WKB approximation at low energies [blue dashed lines, Eq. (S26)]. In (c) we compare $P(n,t)$ as a function of n for $\sigma = 0.1$ [all other parameters are similar to (a) and (b)] with numerical solutions of Eq. (S28) (dashed lines).

where the energy $E \equiv H[n(t), p(t)]$ is constant along a dynamical trajectory given by $p(n) = \log\left[q/(q - En^{\beta})\right]$. To find the energy we solve the equation of motion (S19) on this given dynamical trajectory, which yields

$$\dot{n} = qn^{-\beta} - E. \tag{S21}$$

For $n \gg 1$ and $\beta > 0$ the right hand side of Eq. (S21) varies very slowly with time $[\mathcal{O}(n^{-\beta})]$ (as shown below, the energy E also scales as $n^{-\beta}$), such that the solution for Eq. (S21) can be approximated as $n = (qn^{-\beta} - E)t + C$. Here, C is a slowly-varying function of time, and includes constants such that the energy corresponding to the mean-field solution $n = \langle n\rangle$ obeys $E(n = \langle n\rangle) = 0$ [38]. Having shown that $\langle n\rangle \sim t^{1/(1+\beta)}$, we find $C = \langle n\rangle \beta/(1+\beta)$, which indeed varies with time slower than t. Substituting this back into the equation for n and solving for the energy yields

$$E = q\langle n\rangle^{-\beta}\left\{x^{-\beta} + [\beta - (\beta+1)x]\right\} \tag{S22}$$

where we have expressed t in terms of $\langle n\rangle$ and defined $x \equiv n/\langle n\rangle$. Substituting the energy (S22) into Eq. (S20) and solving the integral yields

$$S(n,t) = \langle n\rangle \mathcal{S}(x) \quad , \qquad \mathcal{S}(x) = \frac{f(x)x^{-\beta}}{\beta+1} + xf(x)^{-1/\beta} B\left[f(x); 1 + \frac{1}{\beta}, 0\right] + x\log(1 - f(x)) \tag{S23}$$

where $B(z;a,b)$ is the incomplete beta function, defined as $B(z;a,b) = \int_0^z u^{a-1}(1-u)^{b-1}du$, and we define $f(x) = 1 - x^{\beta}(\beta(x-1) + x)$. This result coincides with Eq. (7) of the main text and is valid in the limit of $n \gg 1$.

1. *Low energy solution*

To get better insight for the Gaussian vicinity of $P_n(t)$, we solve Eq. (S19) in the low energy limit $E \ll 1$. This yields an approximated solution for the energy in the form

$$E \simeq q\langle n\rangle^{-\beta} \frac{(2\beta+1)x^{-2\beta-1}\left(1 - x^{\beta+1}\right)}{\beta+1}, \tag{S24}$$

where we have again expressed t in terms of $\langle n\rangle$. Equation (S24) is indeed small in the limit $|x-1| \ll 1$, which is the Gaussian vicinity of $P_n(t)$. By further approximating the dynamical trajectory as $p(n) \simeq En^{\beta}/q + E^2 n^{2\beta}/(2q^2)$, we substitute this back into Eq. (S20). Performing the integral and substituting Eq. (S24) yields the following action

$$\mathcal{S}(x) \simeq \frac{(2\beta+1)x^{-2\beta-1}\left(x^{\beta+1} - 1\right)^2}{2(\beta+1)^2}. \tag{S25}$$

In the limit of $|x-1| \ll 1$ the action further simplifies to

$$\mathcal{S}(x) \simeq (\beta + 1/2)\,(x-1)^2, \tag{S26}$$

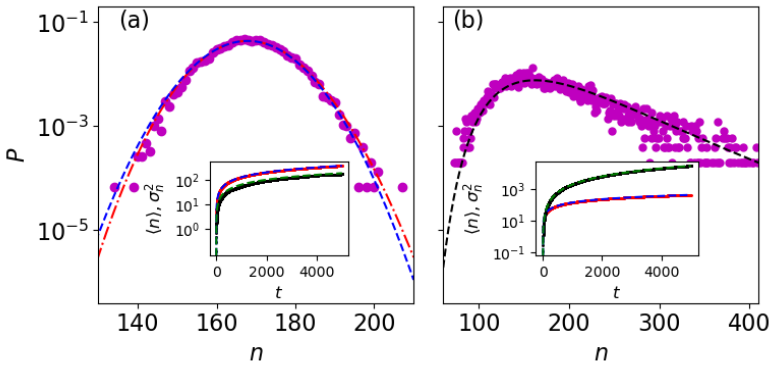

FIG. S2. The probability $P(n,t)$ for $\beta = 0.5$ and $t = 1500$. (a) No variation in β ($\sigma = 0$). We compare simulations (circles), WKB approximation [red dashed line, Eq. (S23)], and WKB approximation at low energies [blue dashed line, Eq. (S25)]. (b) Variability in β with $\sigma = 0.1$, compared to a numerical solution of Eq. (S28) (dashed line). In the insets of both panels are $\langle n\rangle$ and σ_n^2 (red and black marks, respectively) as a function of t, showing very good agreement with the theory (dashed lines).

which coincides with Eq. (8) of the main text and can be shown to solve Eq. (S18) in the limit $|x-1| \ll 1$.

Notably, Eq. (S26) can alternatively be obtained by using the system-size expansion on the master equation [39]. This leads to the following Fokker-Planck equation

$$\frac{\partial P}{\partial t} = \frac{q}{2}\frac{\partial^2}{\partial n^2}(n^{-\beta}P) - q\frac{\partial}{\partial n}(n^{-\beta}P), \tag{S27}$$

which can be dealt with using the WKB ansatz, $P(x,t) \sim e^{-S(x,t)}$. Performing a similar analysis as done above yields a Hamiltonian which coincides with Eq. (S18) in the limit of $p \ll 1$ [39], eventually leading to the action given by Eq. (S26). That is, the Fokker-Planck approximation captures the Gaussian vicinity of the distribution, but missed its tails, see main text. In Figs. 1(a), S1 and S2(a) we compare the WKB solutions to simulations. Note that, while in Fig. 1(a) (main text) we are able to plot the exact results, in Fig. S2(a), due to the larger values of n, the exact result cannot be plotted with standard computational tools.

D. Individual variability

Here we analytically solve Eq. (9) of the main text in the limit of small variance σ. In the main text we write the probability of having visited n sites at time t as [Eq. (9)]

$$P(n,t) = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} P_\beta(n,t)e^{-\frac{(\beta-\beta_0)^2}{2\sigma^2}}d\beta, \tag{S28}$$

which can be numerically solved (Figs. 1(b), S2(b) and S3). Analytical progress can only be made in limit of small variance, $\sigma \lesssim 1/\sqrt{\langle n\rangle_0}$, where $\langle n\rangle_0 = [(1+\beta_0)qt]^{1/(1+\beta_0)}$ is the mean number of sites given $\beta = \beta_0$. For simplicity we focus on the small energy regime, yet similar calculations can be made with the full expression for the action [Eq. (S23)]. Substituting Eq. (S26) into Eq. (S28) yields

$$P(n,t) \sim \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{-\frac{(\beta-\beta_0)^2}{2\sigma^2} - \langle n\rangle_0 \mathcal{S}_\beta(n/\langle n\rangle_0)}d\beta, \tag{S29}$$

where $\mathcal{S}_\beta(x)$ is given by Eq. (S26). This integral can be solved for $\sigma \lesssim 1/\sqrt{\langle n\rangle_0}$, using the saddle point approximation, which yields

$$P(n,t) \sim e^{-\langle n\rangle_0 \mathcal{S}_{\beta_0}(n/\langle n\rangle_0) + \langle n\rangle_0^2\sigma^2\mathcal{S}_1(n/\langle n\rangle_0)}, \tag{S30}$$

with

$$S_1(x) = \frac{(2\beta_0+1)^2\left((\beta_0+1)\log(\langle n\rangle_0)-1\right)^2}{2(\beta_0+1)^4}(x-1)^2 \tag{S31}$$

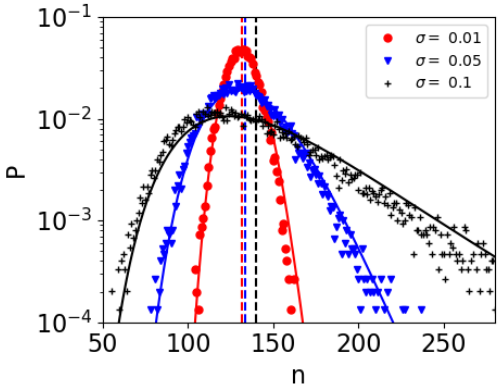

FIG. S3. The probability distribution of number of sites visited at time $t = 1000$ for $\sigma = 0.01, 0.05, 0.1$ (see legend), based on 15000 simulations, compared to Eq. (S29) (solid lines), where the integral is approximated numerically. Note that, the averages $\langle n \rangle$ for each distribution, marked by vertical dashed lines, are only slightly affected by the change in σ.

Here, the mean number of sites obeys $\langle n \rangle = \langle n \rangle_0 \left[1 + \mathcal{O}(\sigma^2)\right]$, whereas the variance obeys

$$\sigma_n = \langle n \rangle_0 \left\{ \frac{1}{2\beta_0 + 1} + \langle n \rangle_0 \sigma^2 \frac{\left[(\beta_0 + 1)\ln(\langle n \rangle_0) - 1\right]^2}{(\beta_0 + 1)^4} + \mathcal{O}(\langle n \rangle_0^2 \sigma^4) \right\}. \tag{S32}$$

Thus, while inter-individual variability will almost not affect the mean number of sites, it does significantly affect the variance of the number of sites (by a factor of $\langle n \rangle_0$ compared to that of the mean), see Fig. S3.

S2. STATISTICS OF NUMBER OF VISITS TO A SITE

Here we provide details on the mean-field equation for the mean number of sites. In particular we explicitly derive and solve this equation for all α values in the limit of $t \gg 1$, and provide evidence of a phase transition at $\alpha = 1$. Our starting point is Eq. (10) of the main text

$$\frac{\partial W_i}{\partial t} = (1 - P_{new}) \left[\Pi_i(m_i - 1) W_i(m_i - 1, t) - \Pi_i(m_i) W_i(m_i, t)\right], \tag{S33}$$

where P_{new} and Π_i are given by Eq. (1) in the main text.

A. The case of $\alpha = 1$

In the main text we assumed that $P_{new} \to 0$ and solved Eq. (S33). Here we provide a solution to the mean-field equation without neglecting P_{new}. In mean field, we obtain an equation for the first moment $\langle m_i \rangle$ by multiplying Eq. (S33) by m_i and summing over all m_i. For $\alpha = 1$ this yields

$$\frac{\partial \langle m_i \rangle}{\partial t} = (1 - P_{new}) \sum_{m_i=1}^{\infty} \frac{m_i}{\sum_{j=1}^{n} m_j} W_i(m_i, t) \simeq \frac{\langle m_i \rangle}{\sum_{j=1}^{\langle n \rangle} \langle m_j \rangle} (1 - q \langle n \rangle^{-\beta}), \tag{S34}$$

where we *a priori* (to be checked *a posteriori*) assume that $\sum_j m_j \gg m_i$ for any site i such that the denominator can be taken out of the sum over m_i, and that $\sum_{j=1}^{n} m_j \simeq \sum_{j=1}^{n} \langle m_j \rangle$. To find the value of $\sum_{j=1}^{\langle n \rangle} \langle m_j \rangle \equiv Q$ we sum over both sides of Eq. (S34) to obtain a differential equation for Q: $\partial_t Q = (1 - q \langle n \rangle^{-\beta})$, an equation which is solved by $Q = t - \langle n \rangle$. Substituting this back into Eq. (S34) gives

$$\frac{d \langle m_i \rangle}{dt} = \frac{\langle m_i \rangle}{t - \langle n \rangle} (1 - q \langle n \rangle^{-\beta}), \tag{S35}$$

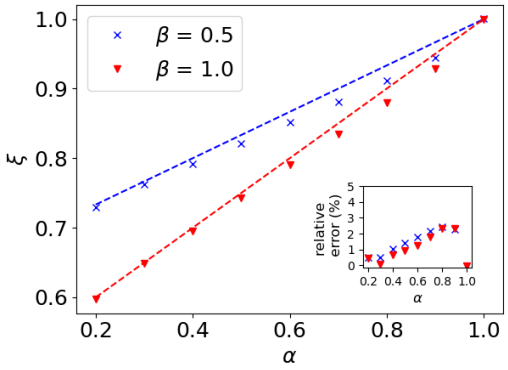

FIG. S4. Comparison between the value of ξ as obtained from 100 simulation of length $t = 10^5$ (symbols) to the theoretical prediction (dashed lines). Plotted for $\beta = 0.5$ (blue crosses) and $\beta = 1$ (red triangles). In the inset we plot the relative error between the predicted value and the one obtained in simulations.

which is solved, assuming site i is first visited at time t_i [*i.e.*, with an initial condition $\langle m_i \rangle (t_i) = 1$], by [24]

$$\langle m_i \rangle = \frac{t - \langle n \rangle}{t_i - \langle n \rangle_{t_i}} \simeq \frac{t}{t_i} \,. \tag{S36}$$

Here $\langle n \rangle_{t_i}$ is the average number of sites at time t_i, and on the right we approximated the solution for $t \gg 1$ and discarded terms of order $\mathcal{O}(t^{1/(1+\beta)})$. This final result agrees with the one found in the main text. Importantly, as all sites have a linear dependence on t, we verify *a posteriori* that $\sum_j m_j \gg m_i$ for any site i, as contribution from all visited sites will not diminish at long times.

B. The case of $\alpha < 1$

For $\alpha < 1$ we solve the mean-field equation at $t \gg 1$ such that $P_{new} \to 0$,

$$\frac{\partial \langle m_i \rangle}{\partial t} \simeq \sum_{m_i=1}^{\infty} \frac{m_i^\alpha}{\sum_{j=1}^n m_j^\alpha} W_i(m_i, t) \simeq \frac{\langle m_i \rangle^\alpha}{\sum_{j=1}^n \langle m_j \rangle^\alpha} \simeq \frac{\langle m_i \rangle^\alpha}{A t^\xi}, \tag{S37}$$

where, similarly to the case of $\alpha = 1$, we *a priori* assume that $\sum_j \langle m_j \rangle^\alpha \gg \langle m_i \rangle^\alpha$ for any site i, and we further assume $\sum_{j=1}^{\langle n \rangle} \langle m_j \rangle^\alpha = A t^\xi$ with $\alpha < \xi < 1$ (to be proved *a posteriori*, see below). Notably, such a scaling was found to hold in numerical simulations. For initial condition $\langle m_i \rangle (t = t_0) = 1$, Eq. (S37) is solved by

$$\langle m_i \rangle \simeq \left[1 + \frac{(\alpha - 1)\left(t^{1-\xi} - t_i^{1-\xi} \right)}{A(\xi - 1)} \right]^{1/(1-\alpha)} . \tag{S38}$$

Note that the asymptotic scaling of this result at $t \gg t_i$ depends on the value of ξ, where for $\xi < 1$, Eq. (S38) predicts an asymptotic scaling of $\langle m_i \rangle \sim t^{(1-\xi)/(1-\alpha)}[1 + \mathcal{O}(t^{\xi-1})]$, for all sites. Now, as all sites scale similarly with t, it is evident that $\sum_j \langle m_j \rangle^\alpha \gg \langle m_i \rangle^\alpha$, thus verifying our initial assumption. Using this solution for $\langle m_i \rangle$, we find that up to some unknown factor $\sum_{j=1}^{\langle n \rangle} \langle m_j \rangle^\alpha \sim t^{1/(1+\beta)} t^{\alpha(1-\xi)/(1-\alpha)}$, entailing that $\xi = \alpha(1-\xi)/(1-\alpha) + 1/(1+\beta) = (1+\alpha\beta)/(1+\beta)$.

In Fig. S4 we show that this prediction agrees with simulations for two different values of β, up to a maximum of 3% relative error. However, we note that this relative error becomes crucial when ξ is substituted back into the scaling $\langle m_i \rangle \sim t^{(1-\xi)/(1-\alpha)}$ found above. Let us denote $\xi_0 = (1 + \alpha\beta)/(1 + \beta)$ such that $\xi = \xi_0(1 - \epsilon)$, where $\epsilon \ll 1$ is a small correction that depends on α and β (see Fig. S4). Substituting ξ into $\langle m_i \rangle \sim t^{(1-\xi)/(1-\alpha)}$ one readily obtains $(1 - \xi)/(1 - \alpha) = \beta/(1 + \beta) + \xi_0\epsilon/(1 - \alpha)$. Here, the approximation is valid only as long as $\beta/(1 + \beta) \gg \epsilon/(1 - \alpha)$, or alternatively $1 - \alpha \gg \epsilon$ [assuming $\beta = \mathcal{O}(1)$]. As we numerically find that $\epsilon = \mathcal{O}(10^{-1})$ this condition is hard to satisfy as α approaches 1, and in this regime it is preferable to find ξ directly from simulations.

In Fig. S5 we plot the number of visits to the five most visited site for different values of α. We show that for $\alpha < 1$ all sites converge to the same number of visits, while for $\alpha > 1$ the most visited site diverges from all other sites (see

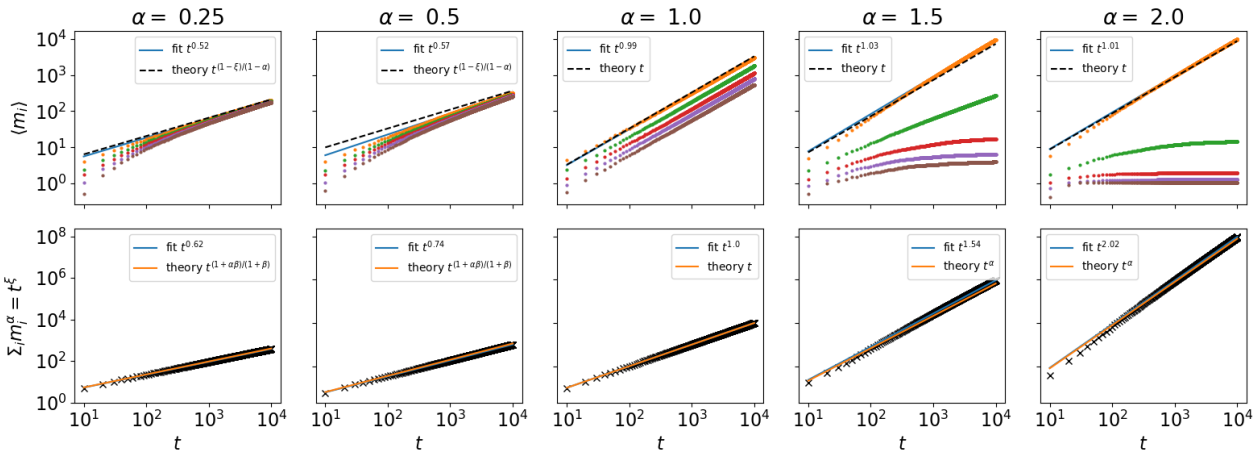

FIG. S5. Upper panels: the mean number of visits to the five most visited sites for different values of α (different colors mark different sites). We fit the most visited site to a power law (blue solid lines, see legend) and plot the theoretical prediction (black dashed line, see legend). Lower panels: measuring ξ from simulations (black crosses), fit (blue solid line) and theory (orange solid line).

below). All α values show good agreement with the theory presented above. Similarly, in the bottom panels of S5 we show evidence for our *a priori* assumption that $\sum_{j=1}^{\langle n \rangle} \langle m_j \rangle^{\alpha} = At^{\xi}$, again with good agreement to the theory.

C. The case of $\alpha > 1$

Here, in contrast to the previous cases, in the limit of $t \gg 1$ we *a priori* assume that $\langle m \rangle_1^{\alpha} \gg \sum_{j=2}^{\langle n \rangle} \langle m_j \rangle^{\alpha}$, *i.e.* at long times the most visited site dominates and contributions from all other sites diminish. We again obtain an equation for the first moment $\langle m_i \rangle$ by multiplying Eq. (S33) by m_i and summing over all m_i:

$$\frac{\partial \langle m_i \rangle}{\partial t} \simeq \sum_{m_i=1}^{\infty} \frac{m_i^{\alpha}}{\sum_{j=1}^{n} m_j^{\alpha}} W_i(m_i, t) \simeq \begin{cases} 1 & i = 1 \\ \frac{\langle m_i \rangle^{\alpha}}{\sum_{j=1}^{n} \langle m_j \rangle^{\alpha}} & i > 1, \end{cases} \tag{S39}$$

where we have separated the most visited site $i = 1$ from all other sites, in accord with the above assumption. For $i = 1$, Eq. (S39) with initial conditions $m_1(0) = 1$ is solved by $\langle m_1 \rangle \simeq 1 + t \simeq t$, *i.e.* we predict a linear scaling with time. For all other sites we assume that $\sum_{j=1}^{\langle n \rangle} \langle m_j \rangle^{\alpha} \simeq \langle m_1 \rangle^{\alpha} \sim t^{\alpha}$. Plugging this into Eq. (S39) yields

$$\langle m_i \rangle \simeq \begin{cases} t & i = 1 \\ \text{const}[1 + \mathcal{O}(t^{1-\alpha})] & i > 1, \end{cases} \tag{S40}$$

where const $\sim (1 - t_i^{1-\alpha})^{1/(1-\alpha)}$. Importantly, for $\alpha > 1$ and $\beta > 0$ it follows that $\alpha > 1/(1+\beta)$, such that $\langle m_1 \rangle^{\alpha} \sim t^{\alpha} \gg t^{1/(1+\beta)} \sim \sum_{j=2}^{\langle n \rangle} \langle m_j \rangle^{\alpha}$, thus verifying our initial assumption. As discussed in the main text, these results suggest a phase transition at $\alpha = 1$, see Sec. S2 F.

D. Statistics of the variance of the number of visits to a site

In the main text we focus on the average number of visits to site i, $\langle m_i \rangle$ (see Figs. 2, S5 and S8), and based on the analysis shown in the main text, we compare between the results for the fruit bats and theory. In Fig. 3(c-f) we find that during summer $\langle m_1 \rangle \sim t^{0.97}$ and $\langle m_2 \rangle \sim t^{0.99}$, as well as $\sigma^2_{m_1} \sim t^{1.94}$ and $\sigma^2_{m_2} \sim t^{2.16}$. In contrast, during winter $\langle m_1 \rangle \sim t^{0.89}$ and $\langle m_2 \rangle \sim t^{0.87}$, as well as $\sigma^2_{m_1} \sim t^{1.96}$, $\sigma^2_{m_2} \sim t^{1.80}$. To further compare these scalings to theory we measured the scalings numerically from simulations. These are summarized, for a specific example ($\beta = 1$), in Fig. S6, while very similar results are obtained for other β values.

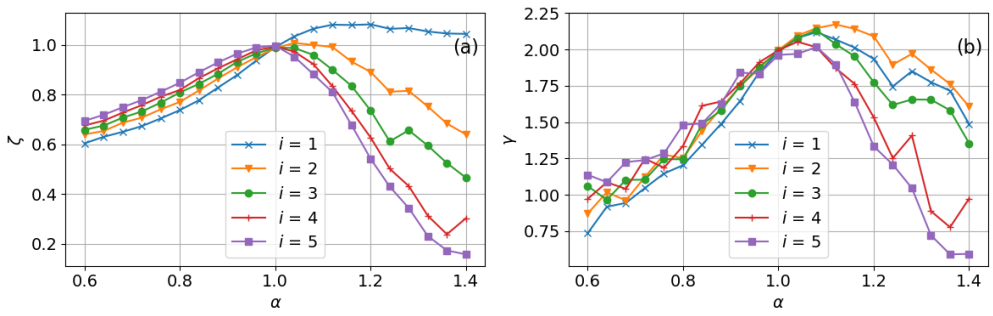

FIG. S6. (a) The value of the exponent ζ in the scaling of the mean number of visits to site i, given by $\langle m_i \rangle \sim t^{\zeta}$. (b) The value of the exponent γ in the scaling of the variance of the number of visits to site i, given by $\sigma^2_{m_i} \sim t^{\gamma}$. In both panels we performed 100 simulations with $t = 10^4$, $\beta = 1$ and $p = 1$, and the scaling exponents are plotted for sites with different ranks (see legends).

E. Statistics of initial site

Here we numerically study the statistics of the visits to the first visited site (that is, the one visited at $t = 0$), and its correlation to the most visited site in the asymptotic limit $t \to \infty$. Due to preferential returns, the first visited site has a significant probability to become eventually the most visited site; we will denote that probability as P_F. This probability will be largely influenced by the dynamics at the initial stage of the process; so, higher values of P_{new} will promote the exploration of new sites, reducing P_F. Likewise, larger values of α strongly reinforce revisits to known sites and will then increase P_F.

While the explicit dependence of P_F on P_{new} and α cannot be analytically determined, it is possible to compute its behavior in some limiting cases. For instance, for $\alpha \to \infty$ and finite P_{new} one can assume that the most visited site will be that one which is visited first twice (then the return probability to that site will be always arbitrarily large compared to the other). Using again $P_{new} = qn^{-\beta}$ one finds

$$\lim_{\alpha\to\infty} P_F = (1-q) + \frac{1}{2}q\left(1 - \frac{q}{2^\beta}\right) + \frac{1}{3}q\frac{q}{2^\beta}\left(1 - \frac{q}{3^\beta}\right) + \ldots = \sum_{i=1}^{\infty} \frac{i-q}{i^{1+\beta}} \frac{q^{i-1}}{[\Gamma(i)]^\beta}, \tag{S41}$$

where $\Gamma(x)$ denotes the Gamma function. In Fig. S7 (left) we show the explicit behavior of P_F as a function of α, checking that the value predicted by (S41) (dotted line) is asymptotically approached as α increases.

The phase transition reported in the main text has also a clear counterpart in the dynamics of the first visited site. Fig. S7 (right) shows the probability density of the random variable m_F, defined as the number of visits to the first visited site, after normalization with respect to time t (such that $0 \leq m_F/t \leq 1$). As observed, the regime $\alpha < 1$ will lead to a distribution peaked at an intermediate value of m_F/t, meaning that the first visited sites will be progressively revisited together with many other sites available. For $\alpha > 1$, instead, the distribution becomes suddenly peaked at $m_F/t = 0$, such that in many realizations of the process there is another site which becomes the most visited at large times. However, as α increases, P_F increases as well, and eventually the distribution will exhibit a second peak at $m_F/t = 1$, representing the case where the first visited site receives all revisits.

F. Evidence of a phase transition

Here we prove that there is a phase transition at $\alpha = 1$, with no *a priori* assumptions on the solution (see previous sections and Figs. 2 and S8). To this end, we define $Q \equiv \sum_{i=1}^{\langle n \rangle} m_i^\alpha$, so the probability that the most visited site will be visited again in the next time step t will be $p_1(t) = m_1^\alpha/Q$, and for any site in general we have $p_i(t) = m_i^\alpha/Q$. Next, we write an expression for the expected value of $p_1(t)$ in the next time step:

$$\langle p_1(t+1) \rangle = \frac{m_1^\alpha + f_\alpha(m_1)}{Q + f_\alpha(m_1)} p_1(t) + \sum_{i=2}^{\langle n \rangle} \frac{m_1^\alpha}{Q + f_\alpha(m_i)} p_i(t), \tag{S42}$$

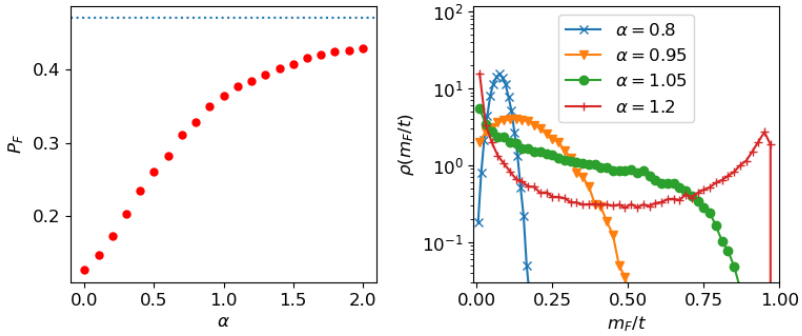

FIG. S7. Left: Probability P_F that the first visited and the most visited site coincide asymptotically, as a function of α. The dotted line shows the value predicted from Eq. (S41). The rest of the parameter values are $\beta = 1$ and $q = 0.9$. Right: Probability distribution function of m_F/t, the number of visits to the first visited site normalized with respect to the total time. The legend shows the different values of α considered, with a transition observed between the behavior for $\alpha < 1$ and $\alpha > 1$. The number of simulations is 5×10^4.

where we have defined $f_\alpha(m_i) \equiv (m_i + 1)^\alpha - m_i^\alpha$. The first term on the right hand side represents the case for which the most visited site is visited in the next time step $t + 1$, and the second term corresponds to the case where any other site is chosen instead.

In the case $\alpha = 1$ we have $f_1(m_i) = 1$ for any m_i. Taking into account that $p_1(t) = m_1^\alpha/Q$, Eq. (S42) leads after some algebra to $\langle p_1(t+1)\rangle = p_1(t)$ independently of the specific set of values $\{m_i\}$ we have. Thus, the probability of revisiting the most visited site will be kept constant through time (and the same can be proved for any other site).

For $\alpha > 1$ we note that $f_\alpha(m_i)$ increases monotonically with m_i. This, together with the fact that $p_1(t) = 1 - \sum_{i=2}^{\langle n\rangle} p_i(t)$ allow us to write the inequality

$$\langle p_1(t+1)\rangle > \frac{m_1^\alpha + f_\alpha(m_1)}{Q + f_\alpha(m_1)} p_1(t) + \frac{m_1^\alpha}{Q + f_\alpha(m_2)}(1 - p_1(t)). \tag{S43}$$

Finally, introducing $p_1(t) = m_1^\alpha/Q$ into the previous inequality, after some algebra we obtain

$$\langle p_1(t+1)\rangle > \left[1 + \frac{(f_\alpha(m_1) - f_\alpha(m_2)(Q - m_1^\alpha)}{(Q + f_\alpha(m_1))(Q + f_\alpha(m_2))}\right] p_1(t). \tag{S44}$$

We thus conclude that $\langle p_1(t+1)\rangle > p_1(t)$ regardless of the specific set $\{m_i\}$ we have. The probability of revisiting the most visited site thus always increases with time on average, leading eventually to its dominance over the others.

For $\alpha < 1$ we proceed in a similar manner. Here, $f_\alpha(m_i)$ will decrease monotonically with m_i, so we can write

$$\langle p_1(t+1)\rangle < \left[1 - \frac{(f_\alpha(m_2) - f_\alpha(m_1)(Q - m_1^\alpha)}{(Q + f_\alpha(m_1))(Q + f_\alpha(m_2))}\right] p_1(t). \tag{S45}$$

This leads to $\langle p_1(t+1)\rangle < p_1(t)$, such that for $\alpha < 1$, on average the probability of revisiting the most visited site will decrease with time, thus leading to a much more homogeneous distribution of revisits among all sites available.

S3. DATA COLLECTION AND ANALYSIS

The Egyptian fruit bat (*Rousettus aegyptiacus*, EFB) is a long-lived, widely distributed Old World fruit bat [61]. Like other fruit bats, individual EFBs tend to feed on a small subset of available trees and repeatedly revisit them for weeks and even months [13] affirms that EFBs rely heavily on individual memory. Additionally, it has recently been shown that EFBs obtain a "cognitive map," which encompasses information about a large number of tree locations, suggesting that memory expands beyond the trees used at a given time [13]. Moreover, in order to justify the use of a non-spatial model to adequately describe movement patterns of fruit bats, we have checked (unpublished) that spatial distance between trees does not explain transition rates between trees. That is, bats can perform a transition with equal probability to close or distant trees. Bats were tracked at a 0.125Hz sampling rate for an average tracking period of 23.7 nights and up to 131 nights. The data also includes nearly all fruit trees in the study area (14,314 trees and 18,111 orchard trees), which enabled us to identify specific tree visits.

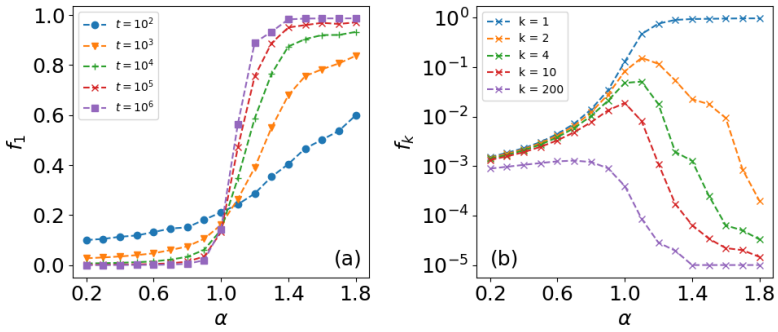

FIG. S8. (a) The average frequency of visits to the most visited site f_1 versus α, for $\beta = 0.5$ (simulations). Each curve corresponds to a given number of visits t (see legend). (b) f_k for different sites (see legend) for $\beta = 0.5$ and $t = 10^5$.

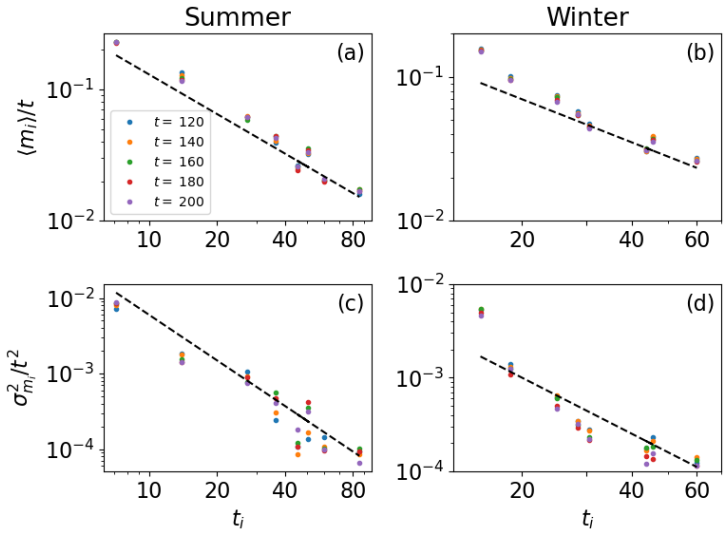

FIG. S9. (a-b): the mean number of visits over time $\langle m_i \rangle / t$ as a function of the average t_i (averaged over different individual bats), at different times t (see legend in top left panel) for summer and winter. Black dashed lines $\langle m_i \rangle / t \sim t_i^{-1}$ correspond to the theoretical prediction for $\alpha = 1$. (c-d): variance of number of visits over time squared $\sigma^2_{m_i}/t^2$ at different times (see legend above) for summer and winter. Black dashed lines $\sigma^2_{m_i}/t^2 \sim t_i^{-2}$ correspond to the theoretical prediction for $\alpha = 1$.

To segment the data into movements and tree visits, we first filtered raw EFB tracks for localization errors based on the covariance matrices attributed to each ATLAS fix [62]. Localization that exceeded the highest realistic speed threshold for this species (20 m/s) were removed. Visits to trees were defined based on track segmentation using the first-passage algorithm to determine the center of a "cloud of fixes" where the animal has spent a specified number of observations within a certain radius (for source code and details see https://github.com/ATLAS-HUJI/R/tree/master/ AdpFixedPoint). Finally, the median coordinates of each cloud were related to the closest tree in the dataset.

To make the seasonal classification most relevant for bats' foraging, we defined winter and summer based on the known peak of fruiting periods of the main seasonal tree species the bats frequently visit in the study area. These are the mulberry (Morus negra) and common fig (Ficus carica) species during May-September (summer) and Chinaberry (Melia azedarach) during November-February (winter). During each fruiting period we used 10 day periods for each bat to ensure sufficient statistics (many of the bats did not have longer tracks during a single season). Based on the field work it is reasonable to assume that during such 10 day periods no significant resource depletion occurs. To fit between the data and the theory we used a standard least square fit procedure from python scipy package.

In addition to the dependence of $\langle m_i \rangle$ and $\sigma^2_{m_i}$ on t for the fruit bats (Fig. 3), we also checked the dependence of these quantities on the the average t_i for any site i, see Fig. S9. Notably the results in the summer are in good agreement with the theory presented in the main text for $\alpha = 1$, whereas for the winter the agreement is not as good. This could be explained by our assessment presented in the main text that during winter α is slightly lower than 1.