Impact of spatiotemporally correlated images on the structure of memory

Alberto Bernacchia[†‡] and Daniel J. Amit[§¶]

[†]Dipartimento di Fisica, Istituto Nazionale di Fisica della Materia, and [‡]Dipartimento di Fisiologia, Dottorato in Neurofisiologia, Universita di Roma ''La Sapienza'', Rome, Italy; and [§]Racah Institute of Physics, Hebrew Univesrity, Jerusalem, Isreal

Communicated by Giorgio Parisi, University of Rome, Rome, Italy, December 23, 2006 (received for review May 2, 2006)

How does experience modify what we store in long-term memory? Is it an effect of unattended experience or does it require supervision? What role is played by temporal correlations in the input stream? We present a plastic recurrent network in which memory of faces is initially embedded and then, in the absence of supervision, the presentation of temporally correlated faces drastically changes long-term memory. We model and interpret the results of recent experiments and provide predictions for future testing. The stimuli are frames of a morphing film, interpolating between two memorized faces: If the temporal order of presentation of the frame stimuli is random, then the structure of memory is basically unaffected by synaptic plasticity (memory preservation). If the temporal order is sequential, then all image frames are classified as the same memory (memory collapse). The empirical findings are reproduced in the simulated dynamics of the network, in which the evolution of neural activity is conditioned by the associated synaptic plasticity (learning). The results are captured by theoretical analysis, which leads to predictions concerning the critical parameters of the stimuli; a third phase is identified in which memory is erased (forgetting).

forgetting | learning | morphing | neural networks | psychophysics

We study the effect of the unsupervised experience of a set of spatiotemporally correlated stimuli on memory structure. In particular, how well can classification memory fare in the presence of spatiotemporal correlations in a sequence of successive stimuli? How does classification memory depend on the presentation protocol? Temporal correlations affect significantly the structure of working memory and its neurophysiological correlate, i.e., delay activity (1–3). Temporal order of presentation of stimuli during training is reflected in correlations of their corresponding delay activity, as observed in anterior inferotemporal cortex (1). In the pair-associate task, temporal correlations and prospective delay activity are induced by presenting fixed pairs of stimuli in a given order (2, 3). Temporal associations between neural representations of memorized stimuli can be induced by Hebbian plasticity mediated by selective delay activity (4–7). Such delay activity carries the neural representations of stimuli across the delay interval separating successive stimulations. Patterns of activity of two successive stimuli can then overlap, and unsupervised Hebbian plasticity renders the two neural representations correlated. Motivated by psychophysical experiments, we propose another mechanism for the evolution of temporal associations between memorized stimuli on the basis of the spatiotemporal interpolation between them.

It has been found that the order of presentation of a sequence of spatially correlated visual images interpolating (morphing) between two human faces has a notable effect on classification performance (8, 9). In ref. 8, human subjects were trained with a set of morphing sequences with no supervision. In each sequence, one face was simultaneously rotated and morphed between two characters. Subjects were then asked to recognize whether a pair of faces, presented from different angles, belonged to the same person. It turns out that subjects confused the faces within the pairs of the morphing sequences. However, if during training the frames of each sequence were presented in random order (mixed), then subjects discriminated successfully.

In ref. 9, human subjects were trained to classify faces into two classes: F (friends) and NF (nonfriends). In each trial, one face was presented and, in the delay interval, was classified as F or NF. Feedback was provided to reinforce the correct classification. Upon reaching good performance, a morphing sequence was generated from one F to one NF, interpolating between the two familiar faces. Next, the images of the morphing sequence were presented for F–NF classification, interspersed with familiar faces. Two protocols were implemented: mixed, in which successive frames were selected at random, and sequential, in which frames followed the order of the morphing sequence. The statistics of F–NF classifications were monitored from session to session.

Subjects initially divided the morphing sequence equally: Frames of the first half were classified F, and those of the second half were classified NF. The separation threshold was near the middle of the sequence. As sessions proceeded, subjects exposed to the mixed protocol maintained the initial threshold for classification: Memory was preserved. In contrast, with the sequential protocol, the separation threshold moved toward NF: An increasing fraction of frames was classified F. Finally, all morphed frames were classified F: Memory collapsed. Hence, the unsupervised experience of spatially correlated stimuli affected strongly the memorized classification structure, depending on the temporal correlations of the presented stimuli.

We present a simple realization of a plastic recurrent network, in which neural response interacts with synaptic plasticity, probing the interplay between memory and (unsupervised) learning. Input stimuli are represented by a distribution of external afferent currents to the neurons of the network. The current received by a given cell depends on the cell and on the presented frame. Hence, each cell has a ''tuning curve.'' Fig. 1*a* is a schematic illustration of the recurrent network (fully connected), neurons (circles), synaptic connections (arrows), and external input currents (rectangles ''tuning curves'') (see *Stimuli, Protocol, and Behavioral Interpretation*).

The results were observed in autonomous simulations, implementing the protocols of ref. 9 by using dynamical equations (see *Neural Dynamics*) driven by external stimuli with no external intervention. The model accounts for the findings and provides insights into the mechanism, and theoretical analysis captures the results of the simulations and predicts the role of different parameters.

Neural Dynamics. A neuron i ($i = 1, \ldots, N$) is modeled by a binary variable, $S_i = \pm 1$. The collection of all N neural states is a network ''state'' or ''configuration.'' The state of cell i at time $t + 1$ is $+1$ or

Author contributions: A.B. and D.J.A. designed research, performed research, and wrote the paper.

The authors declare no conflict of interest.

Abbreviations: F, friends; NF, nonfriends.

¶To whom correspondence should be addressed at: Dipartimento di Fisica, E. Fermi, Universita di Roma I, Piazzale Aldo Moro 5, Rome, Italy. E-mail: daniel.amit@roma1.infn.it.

This article contains supporting information online at www.pnas.org/cgi/content/full/0611395104/DC1.

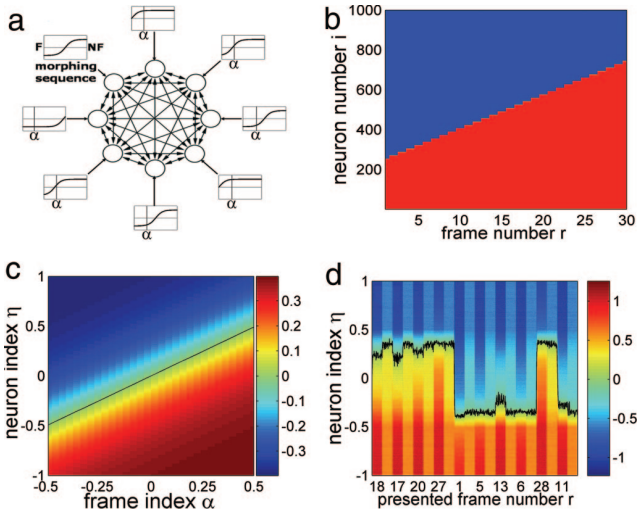

Fig. 1. The network, stimuli, and dynamics. (*a*) Scheme of the network. Circles, eight neurons ($N = 1{,}000$ neurons in simulations); two-headed arrows, recurrent synaptic connections connecting two neurons of the network; one-headed arrows, external connection (input stimulus); rectangles, tuning curves (sigmoidal) of the neurons along the morphing sequence, i.e., external current vs. frame stimulus index; vertical bar in each rectangle, index α of a sample frame stimulus. (*b*) The state of each neuron [+1, (red) or −1 (blue)] vs. frame number r along the morphing sequence. $r = 1$, network state F, has the highest number of neurons in −1 state (blue); $r = 30$: NF, has highest number of +1's (red). Neurons (features) are switched monotonically, proceeding from F to NF. (*c*) External current (color-coded) to all neurons vs. frame index α. Neurons are ordered vertically by increasing values of the neuron index η. The black line corresponds to zero current. (*d*) Sample network dynamics. The first 10 trials of a simulation (mixed protocol, $c = 0.4$, $T = 0.5$) are presented as the total current received by each neuron vs. time. For each presented frame r, the first column is the stimulation interval, and the second column is the corresponding delay interval. Values of the current are color-coded. The Black curve is at zero current ($\eta = \eta^*$). Neurons above this curve are −1, and neurons below are +1. Two (delay) attractor states ($\eta^* = \pm 0.35$) can be observed.

−1, depending on whether the total afferent current at time t is higher or lower than the neural threshold (taken at 0) (10, 11), i.e.,

$$S_i(t+1) = \text{sign}[h_i^{\text{rec}}(t) + h_i^{\text{ext}}(\alpha) + \Gamma_i(t)]. \qquad [1]$$

The current (square brackets) is the sum of the recurrent afferent h_i^{rec} (received from all other network cells); the external current h_i^{ext}, depending on the frame of the morphing sequence presented at time t (indexed α); and a Gaussian noise Γ_i (mean, 0; SD, $\sigma = 0.05$).

The recurrent part is as follows:

$$h_i^{\text{rec}}(t) = \frac{1}{N}\sum_{j=1}^{N} J_{ij}(t)S_j(t), \qquad [2]$$

a sum over the activities of all N neurons of the network ($j = 1, \ldots, N$), modulated by the synaptic weights J_{ij} (from j to i). In the simulations, $N = 10^3$ neurons and $\approx 10^6$ synapses. At $t = 0$, the states of the neurons are initialized once at random ($S_i = \pm 1$ with equal probability) and are updated by Eq. **1** at each time step. Synaptic dynamics is described below in *Synaptic Plasticity and Embedded Memory*.

Stimuli, Protocol, and Behavioral Interpretation. To model the set of morphed frames, we interpret each of the two extreme images, F and NF, as a collection of features; some are shared, and in others they differ. Morphing starts from F, taking off some features and adding others in a monotone, one-way process until NF. In the model with two-state neurons, a +1 neuron implies that a given feature is on; a −1 neuron implies that a another feature is on. When a neuron changes state ($-1 \rightarrow +1$), one feature is taken off and another replaces it. In the model network, the state F has a certain number of neurons in the +1 state, and the rest are −1. Along the morphing sequence, from F to NF, neurons switch states always in the same direction, viz ($-1 \rightarrow +1$). There are more −1's in the F state and more +1's in the NF state.

Because the morphing process is monotonic, we represent each network state in the sequence by a linear ordering of the neurons from −1 to +1 (Fig. 1*b*). Neurons are listed on the vertical axis ($i = 1, \ldots, 1{,}000$); −1 and +1 neural states are depicted in blue and red, respectively. States corresponding to successive frames in the sequence ($r = 1, \ldots, 30$) are in successive columns. The monotonically increasing height of the red part of the column, moving from left ($r = 1$ corresponds to F) to right ($r = 30$ corresponds to NF), are neurons that have switched state (feature) in the morphing process. Neurons maintaining their "color" along the sequence represent features that are common to F and NF. This representation does not imply a linear spatial structure of the network.

The transition of neurons from −1 to +1 is modeled by defining the external current afferent to neuron i upon presentation of a given frame stimulus r. We labeled neurons and frame stimuli by normalized indices, η instead of r and α instead of i. The n (= 30) interpolating frames of the morphing sequence labeled by ($r = 1, \ldots, n$), are labeled by $\alpha = (r-1)/(n-1) - 0.5$; with F by $\alpha = -0.5$; and with NF by $\alpha = +0.5$. With this notation, the first half of the morphing sequence has $\alpha < 0$, and the second half has $\alpha > 0$. Each neuron in the network ($i = 1, \ldots, N$) is assigned a "neuron index," $\eta = 2(i-1)/(N-1) - 1$ ($-1 \le \eta \le 1$). In F, 75% of the cells are −1 and 25% are +1; the opposite holds true for NF (Fig. 1 *b* and *c*).

We write the external current as the following:

$$h_\eta^{\text{ext}}(\alpha) = c\tanh\left(\frac{\alpha-\eta}{T}\right). \qquad [3]$$

This is an analog of the saturating, sigmoidal tuning curves for ocular disparity (12, 13). In Fig. 1*c*, each frame (labeled by α) is described by the variation, from neuron to neuron (labeled by η), of the external current (color-coded), starting from bright red up to deep blue. Each horizontal line is the "tuning curve" of a given neuron as the frame varies. Here, the black diagonal ($\alpha = \eta$) in Fig. 1*c* is the index of the frame stimulus at which the external current to neuron η switches sign. In the absence of recurrents, $S = \text{sign}(h^{\text{ext}})$, and the corresponding neural state flips (Fig. 1*b*). Neurons about to flip in the next morphing step receive small currents, so that a small shift in α results in a sign change of the current. T is the width of the transition region, and c is the saturation amplitude.

In each simulation session, a sequence of values α is prepared according to one of the protocols: mixed, in which the values of α are picked in a random sequence, or sequential, in which the values of α follow the order of the morphing sequence. There are no repetitions in a single session. For each α, the corresponding external current $h^{\text{ext}}(\alpha)$ is injected (Eq. **3**) for the duration of the stimulus presentation (50 steps of the network dynamics), followed by a delay interval (50 steps) in which there is no external current. A single session consists of 3,000 steps (50 + 50 steps × 30 stimuli), and a total of 10 successive sessions (30,000 steps) are simulated for each set of parameters (c, T), for each protocol.

Behavioral response takes place during delay periods after the offset of each stimulus. Correspondingly, in the model, classification is identified by the state reached by the network in the delay interval. If there is a majority of blue (−1) neurons, the response is F; if there are more red (+1) neurons, then the response is NF (Fig. 1*b*). If the difference in the numbers of the two types of

neurons in the delay state is clear and stable, then so is the behavioral response. The delay state of the network is classified by an index η_d^*, which is the index of the neuron that receives current closest to zero in that state and separates the +1's from the −1's. Hence, the delay state is classified by the index of the frame ($\alpha = \eta_d^*$) that would have produced the same bare stimulus state. If $\eta_d^* < 0$, i.e., the blue majority in the delay state (nearer F), then there is an F response. If $\eta_d^* > 0$, the red majority (nearer NF), then there is an NF response.

Synaptic Plasticity and Embedded Memory. Plasticity is defined by a stochastic Hebbian dynamics of the synaptic efficacies J_{ij} (14–16). Each can assume one of two values: $+J$ or $-J$ (we set $J = 1$, i.e., external current and noise are in units of J). At each time step, after the states of all neurons are updated (Eq. **1**), each synaptic efficacy evolves as follows: A synapse J_{ij} connecting two cells in the same activity state ($S_i = S_j$), becomes $+J$ with probability p if it was $-J$, and it remains $+J$ if it was $+J$. A synapse J_{ij} connecting two cells that are in opposite states ($S_i = -S_j$) becomes $-J$ with probability p if it was $+J$, and it remains $-J$ if it was $-J$.

When the network is in a given state, specifying all S_i ($i = 1, \ldots, N$), the synapses vary stochastically, rendering that state more stable (17). Plasticity is assumed slow: $P = 0.8 \times 10^{-4}$ per available synaptic update; <0.4% of the synapses can change per stimulation interval. Synaptic dynamics takes place only during simulation: This emulates the physiological conditions for synaptic modifications, which relates plasticity to visual response, i.e., high- and low-spiking rates during presentation of stimuli.

In the experiment, the frames of the morphing sequence are presented after the memory of classes F and NF are well internalized (9). This is modeled by initializing the synaptic matrix to have two memories [see supporting information (SI) *Text*]. The network has two stable neural configurations (delay states) close to N and NF that can be read from a sample of the network dynamics in the first 10 trials of a simulation of the mixed protocol (Fig. 1*d*). The total afferent current to every neuron (color-coded) is presented vs. time. For each stimulation interval, the presented frame (r) is indicated along the time axis. The black curve indicates the neuron that receives a total current closest to threshold. This neuron separates +1 from −1 neurons in the network; its index η at time t is defined as $\eta^*(t)$ (= η_s^* during stimulation and = η_d^* during delay).

After a stimulus of the first half of the morphing sequence ($r \leq 15$; $\alpha < 0$), the network relaxes to a configuration close to the F state (majority −1's; $\eta_d^* < 0$). A stimulus of the second half leads to a state near NF (majority +1's; $\eta_d^* > 0$). The former case is interpreted as an F response, whereas the latter is interpreted as an NF response.

Note that the initial synaptic matrix represents the embedded classification memory of only two faces, one face from each of the two classes (used to generate the morphing sequence). Other F and NF faces are not considered (see *Discussion*).

Results

At each time step t (the entire network is updated[¶]), the state of the network is recorded. A sample evolution is exhibited in Fig. 1*d* for the mixed protocol: The total afferent current to each neuron (labeled η) is plotted vs. time. As a preliminary result, to highlight the properties of the initial synaptic structure expressing the embedded memory of F and NF, we present the evolution of the network activity in the first 10 frame presentations (first session; see also Fig. 2*a*). The configuration of the network at each time step is characterized by the neuron at which the network flips from −1 (blue) to +1 (red), i.e., receiving threshold current. The black curve is the evolution of the index $\eta^*(t)$ of that neuron.

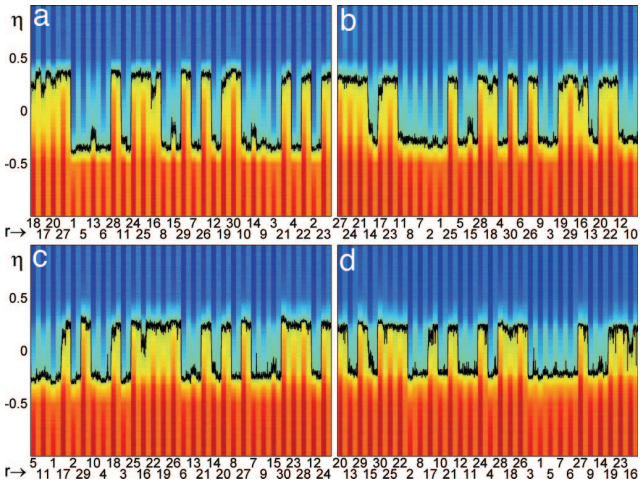

Fig. 2. Network dynamics in the mixed protocol. Total current (color-coded) afferent to each neuron (labeled η) vs. time for 30 trials (one session). Each trial consists of stimulus + delay interval. Abscissa is the frame number (r) of each stimulus. The black curve is the label η^* of the neuron receiving zero current. In each interval, stimulus or delay, η^* fluctuates around a fixed value, respectively η_s^* and η_d^*. (*a*) Session 1. Two delay activity states are present at $\eta_d^* \approx -0.35$ (F response) and $\eta_d^* \approx +0.35$ (NF response). Visual responses vary from frame to frame. (*b–d*): Sessions 4 ($\eta_d^* \approx \pm 0.29$) (*b*), 7 ($\eta_d^* \approx \pm 0.25$) (*c*), and 10 ($\eta_d^* \approx \pm 0.22$). Stimuli from the first (second) half, i.e., $r \leq 15$ ($r > 15$), result in F (NF) responses. ($T = 0.5$; $c = 0.4$).

Because the synapses evolve slowly over a single trial, each time a stimulus is presented or removed, the neural configuration converges rapidly to a stable state, corresponding either to a "visual"[‖] response or to a delay activity configuration. Within each interval, stimulus or delay, the instantaneous label $\eta^*(t)$ fluctuates around its time average for the corresponding interval, i.e., η_s^* and η_d^*, respectively. The visual response varies, depending on the presented frame r, but only two delay states are available, as can be observed by following the black curve in the figure. These two delay states correspond to the memory states.

For instance, the first frame presented ($r = 18$) leads to a visual response $\eta_s^* = +0.25$ and to a delay activity $\eta_d^* = +0.35$. For the fifth stimulus ($r = 1$; i.e., frame F), the visual response is $\eta_s^* = -0.375$, and the delay activity is $\eta_d^* = -0.35$. Each of the other stimuli leads to one of the two memory states, $\eta_d^* \approx \pm 0.35$, which correspond to the two behavioral choices, −0.35 to F and +0.35 to NF. The two memory states express the memory embedded in the synapses.

The first 10 input stimuli of the first session produce the following responses: 18 NF, 17 NF, 20 NF, 27 NF, 1 F, 5 F, 13 F, 6 F, 28 NF, and 11 F. Hence, stimuli of the first half of the morphing sequence ($r \leq 15$) lead to delay states near F (with a blue majority) and consequently to an F response; stimuli of the second half ($r > 15$) lead to delay states near NF (red majority) and to an NF response. This behavior concords with behavior in the early stages of the experiment.

Note that visual responses evoked by frames of the first half of the morphing sequence lead to negative values of η_s^*, whereas frames of the second half lead to positive η_s^*'s.

After the offset of a frame stimulus, the system relaxes to the memory state that is closer to the visual response: Stimuli from the first (second) half of the sequence are followed by delay states close to F (NF), resulting in a F (NF) response.

More specifically, the value of η_s^* for any frame stimulus is in between the index α of the frame, i.e., the value of η_s^* that would

[¶]Simulations were carried out in synchronous mode with noise. We checked that asynchronous updating gave the same results.

[‖]We refer to this response as "visual" for convenience. It is an activation by some abstract feature.

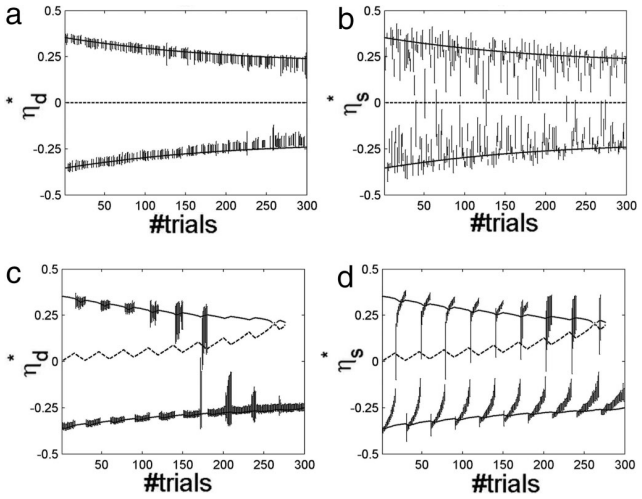

Fig. 3. Delay activity and visual response labels in 10 sessions (300 trials) of the simulations in each of the two protocols. (*a*) η_d^*, label of delay activity states (average over fluctuations in the corresponding interval ± SD) vs. trial number in the mixed protocol. Curves, theoretical predictions; solid lines, labels of stable states; dashed line, label of unstable states. The two solid lines correspond to the two memory states, determining F and NF responses, in agreement with simulation results. The dashed line ($\eta_d^* = 0$) represents the unstable state, the watershed between the two memory attractors. The two memory states are preserved. (*b*) η_s^*, visual response label (average ± SD) vs. trial number in the mixed protocol. The format is the same as in *a*. Labels of visual response states cluster around the memory states. (*c*) Labels of delay activity states vs. trial number in the sequential protocol. The format is the same as in *a*. From session to session (30 trials each), the number of F responses increased, as is evidenced by the fact that the number of delay labels near the upper curve diminishes and that near the bottom curve they increase. Beyond trial 180, there are no NF responses. Next, the line of NF delay activity annihilates in a saddle node. (*d*) Labels of visual response states vs. trial number in the sequential protocol. Like memory states, visual responses tend to leave the NF line and move toward F until all stimuli evoke a state near F. ($T = 0.5$, $c = 0.4$).

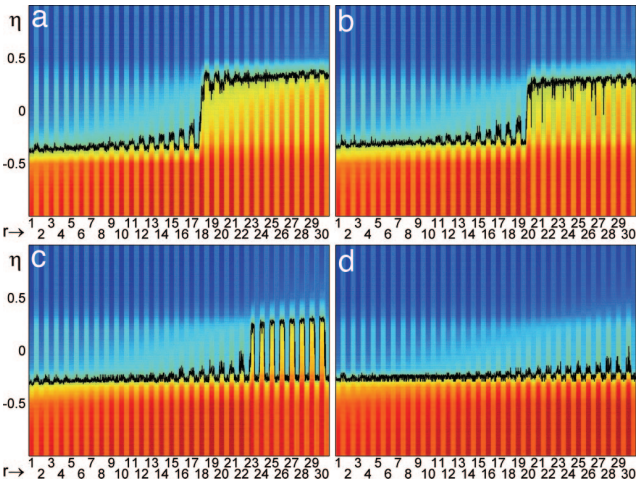

Fig. 4. Network dynamics in the sequential protocol (format is the same as in Fig. 3). Frame stimuli are presented sequentially (r = 1.,,30, abscissa). (*a*) Session 1. The transition of responses from NF (positive η_d^*) to F (negative η_d^*) at $r = 18$ is shown. (*b*) Session 4. Transition at $r = 20$ is shown. (*c*) Session 7. All frames are classified F, but $\eta_s^* > 0$, for r approaching 30. (*d*) Session 10. All $\eta_s^* < 0$, i.e., both visual responses and delay states close to the F state. ($T = 0.5$, $c = 0.4$).

be obtained by the bare stimulus (no recurrents; see, e.g., Fig. 1 *b* and *c*), and the value of η_d^* of the memory state (no external drive). This is a result of the tradeoff between the recurrent afferents modulated by the synaptic matrix of the memory states and the currents injected by the stimulus.

Mixed Protocol. Fig. 2 presents the dynamics of the network along the full simulation for the mixed protocol. It displays the total current afferent to all neurons in 4 of the 10 sessions of the full protocol (sessions 1, 4, 7, and 10 in Fig. 2 *a*, *b*, *c*, and *d*, respectively). The current is color-coded, and is plotted for all neurons (labeled by η) vs. time (same format as in Fig. 1*d*). The indices of frames presented during stimulation intervals (r) are shown on the abscissa; each stimulus is followed by a delay interval. The black curve traces the instantaneous label $\eta^*(t)$ of the neuron receiving zero current, as in Fig. 1*d*. η^* fluctuates within each interval, stimulus or delay, because of the built-in noise.

Along the simulation (10 sessions), the pattern of behavioral responses remains the same as in the first 10 trials (Fig. 1*d*): Stimuli from the first half of the sequence ($r \leq 15$) determine an F response, and stimuli of the second half ($r > 15$) determine an NF response. This is in accordance with the experimental observations.

In a given session, the delay activity states after frame stimuli of the same half of the sequence are approximately independent of the frame. They vary from session to session as a consequence of synaptic plasticity: In session 1, the two delay states correspond to $\eta_d^* \approx \pm 0.35$; in session 4 they are at $\eta_d^* \approx \pm 0.29$; in session 7 $\eta_d^* \approx \pm 0.25$; and in the last session $\eta_d^* \approx \pm 0.22$. The two attractor states are gradually approaching each other along sessions, but theoretical analysis predicts that they remain distinctly separate as the number of sessions increases further. The prediction is that the difference between the two delay states (in terms of η_d^*) converges asymptotically to a fixed value. This prediction is confirmed by the observation that, for both memory states, the variation of η_d^* in successive sessions decreases.

Fig. 3*a* presents the delay activity label η_d^* (each with its respective SD) vs. trial number for all 300 trials of the 10 sessions. Curves are theoretical predictions at $\sigma = 0$ (solid lines for stable states; dashed line for the unstable state). The two lines characterizing the stable memory states are in good agreement with simulation results. The central line at 0 η_d^* determines the unstable state, i.e., the watershed between the two memory states. As discussed above, the two memory states are seen to approach and remain separate.

Fig. 3*b* presents a scatter plot of the "visual" response label η_s^* vs. the trial number, for the same 300 trials, accompanied by theoretical predictions of the delay activity labels η_d^*, which are the same as in Fig. 3*a*. The labels group around the memory states, because beyond the effect of the afferent stimulus, they reflect in part the structure of memory (as observed above). In particular, the labels of most visual response states are in between the labels of the two memory states (see e.g., Fig. 3 *b* and *d*). Because synaptic plasticity tends to reinforce visual response states, the two memory states tend to converge toward each other.

The central result of this section is that the two classification memories are preserved along the 10 sessions, and synaptic plasticity does not change the initial macroscopic picture: The first half of the sequence remains associated with F responses and the second half with NF, in accordance with the experiment in ref. 9.

Sequential Protocol. Fig. 4 presents the analog of Fig. 2 for the sequential protocol (same parameters). The total current afferent to every neuron vs. time is presented in four sessions (sessions 1, 4, 7, and 10 in Fig. 4 *a*, *b*, *c*, and *d*, respectively). The frame numbers (r) on the abscissa follow the sequential presentation order 1, . . . , 30.

In session 1, frames of approximately the first half of the sequence ($r \leq 17$) determine negative values of η_s^*, a delay activity $\eta_d^* \approx -0.35$, and hence F responses. Frames of the second half ($r >$

NEUROSCIENCE

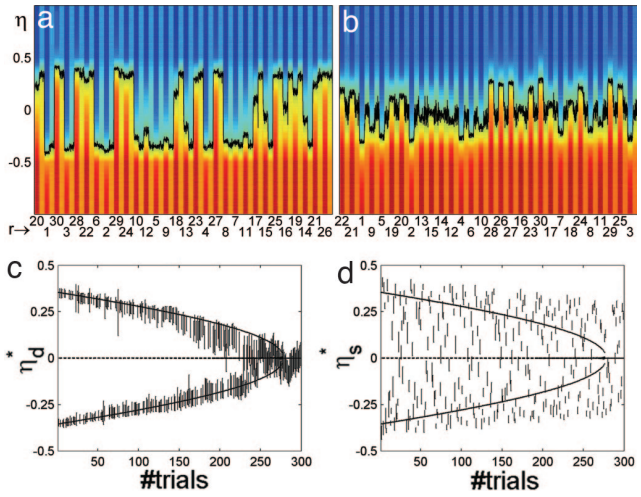

Fig. 5. Forgetting regime. Network dynamics in the mixed protocol, for T/c = 0.6 is shown. (*a* and *b*) Total current (format is the same as in Fig. 2). (*a*) Session 1. This is very similar to Fig. 2*a* (T/c = 1.25), except for the more widely distributed visual responses. (*b*) Session 10. For all delay states, $\eta_d^* \approx 0$, i.e., equal numbers of blue and red neurons, hence random responses. (*c*) η_d^* of delay activity states vs. trial number (format is the same as in Fig. 3). Lines are theoretical predictions for delay activity label. Solid lines, stable; dashed line, unstable. In contrast to the case T/c = 1.25 (Fig. 3*a*), the two memory states are not preserved; they collapse onto a single stable state after 275 trials. (*d*) Scatter plot of η_s^* of visual response states vs. trial number. Visual responses are broadly distributed (compare with Fig. 3*b* for T/c = 1.25). (T = 0.3; c = 0.5).

17) evoke positive η_s^*'s, the memory state $\eta_d^* \approx +0.35$, and NF responses. As a result, apart from the sequential ordering of presentations, there is little difference with respect to the mixed protocol. As the simulation proceeds, things change significantly: Session by session, more and more stimuli of the second half of the sequence evoke delay states with negative η_d^*, producing F responses. In session 4, the transition from F to NF responses occurs after frame 20. By session 7, all frame stimuli determine an F response (see also Fig. 3*c*). For the rest of the simulation (session 8, 9, and 10), all stimuli determine F responses.

Fig. 3*c* presents a scatter plot of η_d^* (delay states) for all (300) trials of this protocol, together with theoretical predictions. The line representing the unstable state approaches the upper memory line (NF responses), until they annihilate in a saddle node. That memory state becomes unstable even before the saddle node because of the noise in the current (Eq. **1**). In fact, in simulations, there are no NF responses after session 6, whereas the saddle node is at session 9. This can be seen noting that there are no delay activity labels near the upper curve in the figure. In between sessions 6 and 9, the memory state has a very small basin of attraction and is destabilized by the noise. Note that the gradual convergence of memory states, observed in the mixed protocol (Fig. 3*a*), is present in the sequential protocol as well, as long as the upper delay state exists.

Fig. 3*d* presents the η_s^* of visual responses for all 300 trials of sequential protocol, together with theoretical prediction for the corresponding η_d^*'s. One observes that as trials proceed, the labels of the visual response states move to negative values, i.e., toward the F state. However, even after the upper memory state (the one closer to NF) has disappeared (session 7, in simulations), frame stimuli can still evoke visual responses with $\eta_s^* > 0$, i.e., close to the NF state.

Forgetting Phase. Theoretical analysis reveals that T/c (ratio of width to contrast of the tuning; see Eq. **3**) is a critical parameter for the network dynamics. The critical point $T/c = 1$ separates two distinct phases: For $T/c > 1$, memory is preserved, whereas for $T/c < 1$, a long exposure to the morphing sequence provokes total memory erasure due to synaptic plasticity (*SI Text*). In the simulations shown in previous figures, T/c = 1.25. Instead, Fig. 5 shows the results of a simulation with the mixed protocol, at T/c = 0.6. Fig. 5 *a* and *b* present the network dynamics in sessions 1 and 10, respectively (as in Figs. 2 and 4 *a* and *d*). The first session is similar to Fig. 2*a*, except for a wider distribution of visual response labels (see also Fig. 5*d*). By session 10, the picture is very different: All delay states correspond to $\eta_d^* \approx 0$.

Fig. 5 *c* and *d* show, respectively, the labels of delay activity states and visual responses in all 300 trials, as in Fig. 3. The stable delay states converge toward each other until they annihilate in a pitchfork after 275 trials (Fig. 5*c*). The visual response labels are not as clustered around the memory states as in Fig. 3*b*. They tend to spread more and more broadly as trials proceed (Fig. 5*d*). Given that in late trials $\eta_d^* \approx 0$, in the delay period after each stimulus, the network relaxes to a state in which the number of blue and red neurons are essentially equal, and excess in any direction is random. The behavioral responses are expected to be random and uncorrelated with the stimulus presented. This situation is identified with total forgetting.

The underlying reason for the appearance of this critical value of T/c is that for a frame stimulus of index α, neurons with $\eta \approx \alpha$ are the ones that have their activity determined mostly by recurrent currents, because the external afferents they receive are near threshold (Eq. **3**). These neurons are the ones sensitive to the structuring of long-term classification memory. Their fraction increases with T/c. If T/c is too small, the effect of memory in the process of plasticity is overcome by the external stimuli and is independent of previous memory structure. The result is that memory is lost.

Discussion

The model (for $T/c > 1$) captures, in a simple and complete way, the phenomenology described in ref. 9 and gives a rich picture of the interaction of memory with continued presentation of a sequence of spatiotemporally correlated stimuli. The resulting phenomenology is due to the interaction of neural activity and synaptic plasticity dynamics, which allows the interplay between memory and unsupervised exposure to the input stimuli. Neither neural nor synaptic dynamics are task-specific.

The underlying scenario is that, initially, each time a stimulus is presented, the corresponding visual response state is close to one of the two extreme frames, F or NF. Hebbian plasticity reinforces that state, increasing its stability, and the other state is correspondingly somewhat weakened. In the sequential protocol, in the middle of a session, i.e., in the middle of the morphing sequence, the recent past is characterized by a systematic presentation of stimuli from the first half of the sequence, all of which have enhanced the stability of the memory state near F, resulting in the expansion of the basin of attraction of that memory and a reduction of the basin of the other one. This is expressed by the drift of the unstable state toward the memory state near NF (Fig. 3*c*). As a result, some stimuli from the second half evoke a visual response near the F memory (frames16 and 17 in Fig. 4*a*). Beyond some position along the second half of the sequence, the states evoked are near NF, and synaptic plasticity allows some recovery of the lost basin, as seen by small back drifts of the unstable state toward the center near the end of every session. However, it does not recover fully: As sessions progress, this state eventually disappears (collapse). In contrast, in the mixed protocol, stimuli from both halves of the sequence alternate randomly, and the two memory basins remain balanced (Fig. 3 *a* and *b*). Consequently, the F and NF responses remain balanced, and classification memory is preserved. Note that the value of p, determining the speed of synaptic plasticity, is crucial: A very small p implies that plasticity is slow, and the two memories could be maintained even in the sequential protocol. For instance, at most 1.5% of synapses are expected to switch along 10 sessions if $P = 10^{-6}$. Conversely, if

p is large, plasticity is fast, and memories are stored and/or lost instantly. If $P = 10^{-2}$, up to 40% of synapses are expected to switch during a single presentation. In that case, the collapse of memory would take place for every stimulus, even in the mixed protocol.

In the memory preservation regime, $T/c > 1$, neural configurations during and after stimulus presentation are related: Visual response states cluster around the delay activity states (Fig. 3 *b* and *d*). The presence of the two clusters of stimulated activity is due to the fact that the memory of F and NF is stored in the synaptic structure, and the visual response to an input stimulus reflects, in part, the embedded memory structure. However, the patterns of visual response can be separated from the patterns of delay activity. This appears in accordance with physiological observations (18).

The fact that in the mixed protocol ($T/c > 1$) two distinct delay activity states remain stable despite persistent synaptic plasticity (Fig. 3*a*) concords also with the physiological observation. In ref. 19, morphing sets of images were generated between two sets of visual stimuli. These images were presented for classification in random (mixed) order. The delay activity was found to maintain the coding for the class as trials went on, irrespective of the level of morphing of the input stimulus.

Also the memory collapse for the sequential protocol leads to a physiological prediction: The model predicts that the patterns of stable activity, where they exist, are only mildly affected by plasticity (Fig. 3*c*). But in well advanced sessions, the responses to frames in the second half of the sequence begin to be F (instead of NF). This implies, on the physiological level, that neurons should be observed that abruptly change their delay activity from F to NF selectivity more and more into the second half of the morphing sequence.

The model predicts also (for $T/c < 1$; Fig. 5) the total erasure (forgetting) of the classification memory (Fig. 5*b*), a phenomenon that could be observed experimentally. Unlike the previous cases, both delay activity states and visual response states are expected to be strongly affected by synaptic plasticity. This would lead to the broadening of the distribution of visual responses (Fig. 5*d*) and the drawing nearer of the two memory states (Fig. 5*c*). The delay activity patterns corresponding to the two distinct memories would become increasingly similar, until they merge. Subjects would then be expected to respond as if they did not remember the classification, i.e., by chance. Both extracellular recordings *in vivo* and psychophysical observations could detect this phenomenon. In particular, neurons should be found that switch their selectivity between NF and F shortly before total forgetting, i.e., before the subject starts responding by chance. The latter would correspond to neurons with $\eta \approx 0$. To induce the "forgetting" effect, the stimuli should be designed by manipulating the smoothness of the transition of features along the morphing sequence. This can be effected, for example, by applying sigmoidal filters.

This study also suggests that, due to the convergence of the memory states, the perception of the "pure" F and NF faces should evolve from session to session. This could be checked in an experiment in which, at the end of each session, subjects are presented a line-up of several images for perusal and asked to identify the original F or NF. The responses may move toward the center of the morphing sequence.

The fact that each activity pattern is represented by the label of a single neuron (η^*) should not be interpreted as a grandmother cell representation: It is not the activation of one neuron that accounts for one pattern, because the same neuron could be active (or inactive) both in F and in NF states. This type of label serves only for illustrative purposes and is not an intrinsic property of the network. Apart from the (innovative) treatment of external currents, the present model is set up with the ingredients of associative memory models (10, 11, 14, 15), with a large storage (20). In the context of our model, the capacity is not addressed, and only two memories are considered, a single F/NF pair. The present discussion regards the basin of attraction of these two memories. The question of more F and NF stored stimuli should be considered in the context of hierarchically structured attractor networks (21), where memories are considered parts of one large attractor.

Finally, a more realistic modeling of the network dynamics would employ spiking neurons and activity-driven synapses (22, 23), but the step seems rather innocuous. The simplifications of the present model render it aesthetically very attractive and make theoretical analysis accessible.

We thank S. Preminger for communications about the experiment, Dr. Y. Amit and N. Brunel for very helpful comments on a previous version of the manuscript, and S. Romani for helpful discussions and technical assistance. This work was supported by a Center of Excellence Grant from the Israel Science Foundation.

1. Miyashita Y (1988) *Nature* 335:817–820.
2. Sakai K, Miyashita Y (1991) *Nature* 354:152–155.
3. Erickson CA, Desimone R (1999) *J Neurosci* 19:10404–10416.
4. Griniasty M, Tsodyks MV, Amit DJ (1993) *Neural Comput* 5:1–17.
5. Amit DJ, Brunel N, Tsodyks MV (1994) *J Neurosci* 14:6435–6445.
6. Brunel N (1996) *Neural Comput* 8:1677–1710.
7. Mongillo G, Brunel N, Amit DJ (2003) *Eur J Neurosci* 18:2011–2024.
8. Wallis G, Bulthoff HH (2001) *Proc Natl Acad Sci USA* 98:4800–4804.
9. Preminger S, Sagi D, Tsodyks MV (2004) *Neural Plasticity* 12:48.
10. Amari SI (1972) *IEEE Trans Comp C* 21:1197–1206.
11. Hopfield JJ (1982) *Proc Natl Acad Sci USA* 79:2554–2558.
12. Pouget A, Sejnowski TJ (1997) *J Cogn Neurosci* 9:222–237.
13. Dayan P, Abbott LF (2001) *Theoretical Neuroscience* (MIT Press, Cambridge, MA).
14. Tsodyks MV (1990) *Mod Phys Lett B* 4:713–716.
15. Amit DJ, Fusi S (1994) *Neural Comp* 6:957–982.
16. Brunel N, Carusi F, Fusi S (1998) *Network* 9:123–152.
17. Hebb DO, Donderi DC (1994) *Textbook of Psychology* (Kendall/Hunt, Dubuque, IA).
18. Miyashita Y, Chang HS (1988) *Nature* 331:68–70.
19. Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2001) *Science* 291:312–316.
20. Amit DJ, Gutfreund H, Sompolinsky H (1985) *Phys Rev Lett* 55:1530–1533.
21. Virasoro AM (1989) *Phys Rep* 184:301–306.
22. Fusi S, Annunziato M, Badoni D, Salamon A, Amit DJ (2000) *Neural Comp* 12:2227–2258.
23. Mongillo G, Curti E, Romani S, Amit DJ (2005) *Eur J Neurosci* 21:3143–3160.

Journal of the Royal Statistical Society
SERIES B Statistical Methodology

J. R. Statist. Soc. B (2011)
73, *Part* 3, *pp.* 407–422

Self-consistent method for density estimation

Alberto Bernacchia

Yale University, New Haven, USA

and Simone Pigolotti

Niels Bohr Institute, Copenhagen, Denmark

[Received July 2010. Revised December 2010]

Summary. The estimation of a density profile from experimental data points is a challenging problem, which is usually tackled by plotting a histogram. Prior assumptions on the nature of the density, from its smoothness to the specification of its form, allow the design of more accurate estimation procedures, such as maximum likelihood. Our aim is to construct a procedure that makes no explicit assumptions, but still providing an accurate estimate of the density. We introduce the self-consistent estimate: the power spectrum of a candidate density is given, and an estimation procedure is constructed on the assumption, to be released *a posteriori*, that the candidate is correct. The self-consistent estimate is defined as a prior candidate density that precisely reproduces itself. Our main result is to derive the exact expression of the self-consistent estimate for any given data set, and to study its properties. Applications of the method require neither priors on the form of the density nor the subjective choice of parameters. A cut-off frequency, akin to a bin size or a kernel bandwidth, emerges naturally from the derivation. We apply the self-consistent estimate to artificial data generated from various distributions and show that it reaches the theoretical limit for the scaling of the square error with the size of the data set.

Keywords: Binning; Kernel density estimation; Non-parametric statistics

1. Introduction

Every scientist has encountered the problem of estimating a continuous density from a discrete set of data points. This may happen, for example, when determining a probability distribution from a finite Monte Carlo sample (Binder, 1986), rounding off the shape of a galaxy from a collection of stars (Ripley and Sutherland, 1990) or assessing the instantaneous firing rate of a neuron from a discrete set of action potentials (Kass *et al.*, 2005). In all those cases, one can adopt two different approaches: either assuming a given functional form for the density *a priori*, specified by a certain number of parameters, or renouncing any prior knowledge (beyond that a density exists and, in some cases, that it is smooth). These two approaches lead respectively to parametric and non-parametric estimates. We shall focus on the latter approach, although we shall assume knowledge of the density as a reasoning tool, to be released *a posteriori*.

The most popular non-parametric method is simply plotting a histogram, but more sophisticated procedures have been developed. Kernel density estimation (KDE) has been widely studied (Silverman, 1986; Parzen, 1962; Wand and Jones, 1995): instead of counting the number of points in separate bins, KDE constructs a smoothed picture of the data as a superposition of kernel functions centred at the co-ordinates of data points. More formally, given a sample of

Address for correspondence: Alberto Bernacchia, Department of Neurobiology, Yale University, 333 Cedar Street, SHM-C400D, New Haven, CT 06510, USA.
E-mail: alberto.bernacchia@yale.edu

 1369–7412/11/73407

N data points (real numbers), denoted by $\{X_j\}$ ($j=1,\ldots,N$), the kernel density estimate $\hat{f}(x)$ is written as

$$\hat{f}_{\mathrm{KDE}}(x)=\frac{1}{hN}\sum_{j=1}^{N}K\left(\frac{x-X_j}{h}\right) \tag{1}$$

where $K(x)$ is the smoothing kernel and h is the bandwidth. Usually, the choice of $K(x)$ is not crucial (Silverman, 1986), whereas h, which controls the degree of smoothing, must be carefully adjusted: the more concentrated data points are, the less smoothing is necessary to obtain a good estimate of their density. An alternative non-parametric method is maximum penalized likelihood (Good and Gaskins, 1971), which is also known in the physics literature as a regularization of field theory (Bialek *et al.*, 1996; Holy, 1997; Schmidt, 2000): it consists in performing a functional average of densities weighted by their likelihood and by a measure of their smoothness.

In general, each non-parametric method depends on the arbitrary choice of an adjustable parameter, such as the bin size in histograms, the bandwidth in KDE or the cut-off frequency in maximum penalized likelihood and field theory. Each of them regularizes the estimate and avoids overfitting of the data points. In most cases this corresponds to low pass filtering, i.e. cutting the high frequencies that are inherent to the discrete data set, and preventing the estimate from merely reproducing a narrow peak at each data point. However, it would be desirable to devise methods involving the least possible number of parameters, since their determination usually involves some specific assumptions on the distribution to be estimated (for example varying the cut-off parameter in maximum penalized likelihood and field theory precisely corresponds to different choices of the Bayesian prior (Bialek *et al.*, 1996)). Cross-validation techniques have been previously applied for this (Bowman, 1984), but they are computationally expensive and have been seldom applied in the literature.

In this study we show that a self-consistent approach leads to the emergence of a natural cut-off frequency, and an estimate of the density whose performance approaches the theoretical limit for the scaling of the square error with the size of the data set. We start from the observation, which was made in Watson and Leadbetter (1963), that a unique 'optimal' convolution kernel can be derived as a function of the power spectrum of the (unknown) density to be estimated. This result alone is of little use, since the power spectrum of the true density is not known *a priori*. However, in Section 2 we exploit the result by defining the 'self-consistent' estimate as the estimate whose associated optimal kernel, applied to the sample data set, returns the estimate itself. Our main result is to derive the exact expression of the self-consistent estimate for any given data set, and to study its properties. In Section 3 we test the method on three different problems: the estimates of Gaussian, Cauchy and comb distributions. In all cases we show that the self-consistent estimate outperforms existing methods and the mean integrated square error reaches the optimal theoretical scaling of approximately N^{-1}. Technical material is presented in the appendices: in Appendix A we replicate and extend the results on the existence of the optimal kernel. In Appendix B we provide details of the derivation of the self-consistent estimate. In Appendix C we prove that the self-consistent estimates converge almost surely to the true distribution for large N.

The programs that were used to analyse the data can be obtained from

`http://www.blackwellpublishing.com/rss`

2. The self-consistent estimate

In this section, we define the self-consistent estimate, we derive its exact expression and we study

its properties. We start from a result that was derived by Watson and Leadbetter (1963), that we replicate and extend in Appendix A. The basic result is that a unique, optimal convolution kernel can be derived as a function of the power spectrum of the density to be estimated, where 'optimal' is intended as minimizing the mean integrated square difference between the true density and its estimate.

Given a sample of N data points (real numbers), denoted by $\{X_j\}$ $(j=1,\ldots,N)$, each independently drawn from a probability density distribution $f(x)$, we write the estimate as

$$\hat{f}(x)=\frac{1}{N}\sum_{j=1}^{N}K(x-X_j) \tag{2}$$

where we assume that $f, \hat{f}\in L^2$. Note that equation (2) does not depend on any bandwidth h, contrary to the kernel density estimate (1). Instead of choosing an arbitrary shape for the kernel K, and looking for an optimal bandwidth (see for example Silverman (1986)), we rather look for an optimal shape of the kernel. It turns out that the Fourier transform $\kappa_{\mathrm{opt}}(t)$ of the optimal kernel $K_{\mathrm{opt}}(x)$ is equal to (see Appendix A and Watson and Leadbetter (1963))

$$\kappa_{\mathrm{opt}}(t)=\frac{N}{N-1+|\phi(t)|^{-2}} \tag{3}$$

where $\phi(t)$ is the Fourier transform of the true density $f(x)$ (characteristic function). The optimal kernel $K_{\mathrm{opt}}(x)$ is symmetric with respect to $x=0$, where it takes its maximum value. $|\phi(t)|$ in equation (3) requires knowledge of the true density, which is not available; hence equation (3) cannot be used to compute the estimate in equation (2) from the sample observations $\{X_j\}$ alone. We show in what follows how to circumvent this problem by a self-consistent approach. Equation (3) has been previously derived by Watson and Leadbetter (1963) and has been used for assessing the performance of specific kernels (Davis, 1977), as well as for constructing blockwise estimators (Efromovich, 2008).

Although equation (3) cannot be used to compute the density estimate, we make a step further and we write the Fourier transform $\hat{\phi}(t)$ of the density estimate $\hat{f}(x)$ in equation (2), using the transformed kernel (3), as

$$\hat{\phi}(t)=\Delta(t)\,\kappa_{\mathrm{opt}}(t)=\frac{N\,\Delta(t)}{N-1+|\phi(t)|^{-2}}, \tag{4}$$

where $\Delta(t)$ is the empirical characteristic function, i.e.

$$\Delta(t)=\frac{1}{N}\sum_{j=1}^{N}\exp(\mathrm{i}tX_j). \tag{5}$$

Our approach is to construct an iterative procedure based on equation (4), and to determine its exact fixed point. We replace the unknown term ϕ in equation (4) with an initial guess $\hat{\phi}_0$, and we denote the resulting estimate as $\hat{\phi}_1$. Then, we try to obtain an improved estimate $\hat{\phi}_2$ by using a kernel which is optimal for $\hat{\phi}_1$. By iterating this procedure, we construct the following sequence of estimates:

$$\hat{\phi}_{n+1}=\frac{N\Delta}{N-1+|\hat{\phi}_n|^{-2}}. \tag{6}$$

We search for a fixed point of the iteration, namely an estimate $\hat{\phi}_{\mathrm{sc}}$ for which

$$\hat{\phi}_{\mathrm{sc}}=\frac{N\Delta}{N-1+|\hat{\phi}_{\mathrm{sc}}|^{-2}}. \tag{7}$$

This coincides with the density whose corresponding optimal kernel applied to the data sample

gives back the density itself. We call the resulting estimate a 'self-consistent estimate'. We derive in Appendix B the stable solution of equation (7), which is equal to

$$\hat{\phi}_{\mathrm{sc}}(t) = \frac{N\,\Delta(t)}{2(N-1)}\left[1+\sqrt{\left\{1-\frac{4(N-1)}{N^2|\Delta(t)|^2}\right\}}\right] I_A(t) \tag{8}$$

where I is the indicator function ($I_A(t)=1$ if $t\in A$; $I_A(t)=0$ if $t\notin A$) and A is the set of 'accepted' frequencies, i.e. the frequencies giving a non-zero contribution to the estimate. For equation (8) to be a stable solution of equation (7), the set A must be contained in B ($A\subseteq B$), where $t\in B$ if and only if

$$|\Delta(t)|^2 \geqslant \frac{4(N-1)}{N^2}. \tag{9}$$

This condition sets a threshold for the amplitudes of frequencies t below which $\hat{\phi}_{\mathrm{sc}}(t)=0$. Hence, the contribution of small amplitude waves is neglected, and this automatically determines the range of frequencies to be considered for the estimate. In most practical situations, the filter will cut the high frequency bands, but the filter is not constrained to be low pass, and it can rather select different frequency bands.

The condition $A\subseteq B$ leaves the arbitrary choice of a subset of frequencies among those above the threshold set by condition (9). As demonstrated in Appendix C, the self-consistent estimate converges almost surely to the true density, provided that A is bounded, where the bound grows with N (in addition, the characteristic function is required to be integrable). In practical applications, a bounded interval must be necessarily implemented, and we selected

$$A = B\cap[-t^*, t^*] \tag{10}$$

where, as defined above, $B=\{t:|\Delta(t)|^2\geqslant 4(N-1)/N^2\}$, and t^* is set in such a way that inequality (9) holds in one half of the interval $[-t^*, t^*]$. The estimate is not sensitive to the choice of this interval (in applications, a 50% change in the interval resulted, on average, in a 1% change in the estimate; see Section 3).

Since $\hat{\phi}_{\mathrm{sc}}$ and A are bounded, the self-consistent estimate in Fourier space, equation (8), can be antitransformed back to the estimate in real space, i.e.

$$\hat{f}_{\mathrm{sc}}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-\mathrm{i}tx)\hat{\phi}_{\mathrm{sc}}(t)\,\mathrm{d}t. \tag{11}$$

Applications of the method are considered in the next section; here we describe its properties. A graphical illustration of the filtering properties of the estimate is given in Fig. 1, where the amplitude gain $G=|\hat{\phi}_{\mathrm{sc}}|/|\Delta|$ is plotted.

The self-consistent estimate is normalized, i.e.

$$\int_{-\infty}^{\infty} \hat{f}_{\mathrm{sc}}(x)\,\mathrm{d}x = 1$$

or, equivalently, $\hat{\phi}_{\mathrm{sc}}(0)=1$, as a consequence of the normalization of the empirical density ($\Delta(0)=1$). Beside the zero frequency that is kept intact (corresponding to the normalization condition), the self-consistent estimate attenuates all other frequencies ($|\Delta|\leqslant 1$ implies that $|\hat{\phi}_{\mathrm{sc}}|\leqslant|\Delta|$; see Fig. 1). Because $\hat{\phi}_{\mathrm{sc}}(t)$ is continuous and infinitely differentiable at $t=0$, all the moments of the self-consistent estimate $\hat{f}_{\mathrm{sc}}(x)$ exist. The mean and variance of $\hat{f}_{\mathrm{sc}}(x)$ are equal to (see Appendix B)

$$E(x) = \frac{1}{N}\sum_{j=1}^{N} X_j, \tag{12}$$

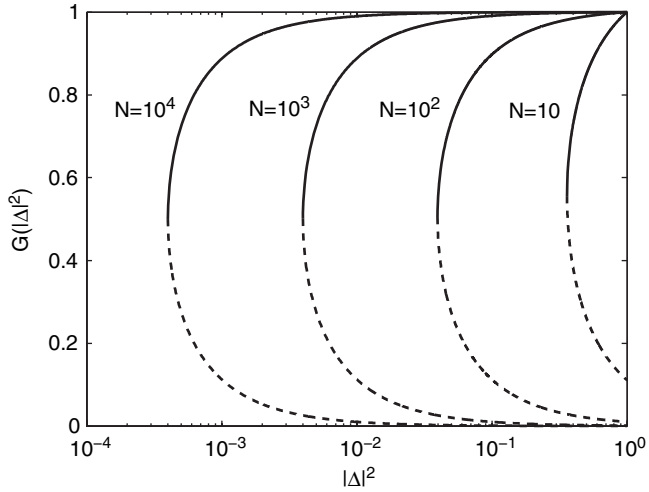

Fig. 1. Amplitude gain of the self-consistent estimate (———), $G = |\hat{\phi}_{sc}|/|\Delta|$, as a function of the squared input amplitude $|\Delta|^2$ and for various values of the sample size N, and the unstable solution (-------) (see Appendix B): the amplitude gain exists only if inequality (9) is satisfied, and is always smaller than 1, implying that the self-consistent estimate attenuates the amplitudes in the input; the exception is for $|\Delta|^2 = 1$, for which $G = 1$, corresponding to the normalization condition; $G \to 1$ for large values of N (whereas the unstable solution vanishes), implying that the estimate tends to reproduce all frequencies of the input in that case

$$\mathrm{var}(x) = \frac{1}{N-2} \sum_{j=1}^{N} \{X_j - E(x)\}^2. \tag{13}$$

Whereas the mean is equal to the sample mean, the variance is larger than the sample variance as well as larger than its unbiased estimator, which is normalized by $1/(N-1)$ instead of $1/(N-2)$.

A drawback of the self-consistent estimate is that it is not guaranteed to be non-negative, whereas the true density is non-negative (note that $|\hat{\phi}_{sc}(t)|^2 \leqslant 1$ holds from equation (8), because $|\Delta|^2 \leqslant 1$, but that is a necessary and not sufficient condition for $\hat{f}_{sc}(x)$ to be non-negative). However, we seldom observed negative values in simulations. Those can be corrected without any error cost by translating the estimate downwards until the positive part is normalized to 1, and setting to 0 the negative part (Glad *et al.*, 2003; Ushakov, 1999). In general, the restriction to a strictly non-negative estimate has a cost, in terms of the mean integrated square error $E(I)$, quantified by the decay exponent α of the error as a function of the sample size, $E(I) \propto N^{-\alpha}$. Among the estimation procedures that properly give non-negative results, histograms, maximum penalized likelihood and field theories have $\alpha = \frac{2}{3}$ (Bialek *et al.*, 1996; Holy, 1997), which is also the limit of KDE when the density is discontinuous (Ushakov, 1999). For a continuous density, KDE improves to $\alpha = \frac{4}{5}$ (Silverman, 1986). Applications of estimates that are allowed to be negative reach better performance, like 'mth-order' kernels (Wand and Jones, 1995; Hall and Marron, 1987; Berlinet, 1993), that have $\alpha = 2m/(2m+1)$ (provided that the density is $(m-1)$th differentiable), whereas infinite order kernels (Devroye, 1992), and the sinc kernel (Davis, 1977; Schmidt, 2000; Glad *et al.*, 2007), yield $\alpha = 1$ (for infinitely differentiable densities, besides logarithmic terms). Hence, releasing the requirement of a non-negative density estimate enables an improvement in the performance. The optimal scaling $\alpha = 1$ is also reached by parametric estimators, such as maximum likelihood, which are, however, strictly non-negative. We present below numerical results suggesting that the self-consistent estimate (8) also reaches $\alpha = 1$ for infinitely differentiable densities.

3. Applications to artificial data

In this section we apply the self-consistent estimate to artificial data. The estimate $\hat{f}_{\mathrm{sc}}(x)$ is constructed from a sample of N data points $\{X_j\}$ $(j=1,\ldots,N)$, by using equations (5) and (8)–(11). As an illustrative example, we present in Fig. 2 an application of the self-consistent estimate to a Gaussian sample of $N=200$ points with respect to a histogram (where we chose the optimal bin width; see Scott (1979)).

We test more systematically the performance of the self-consistent estimate on artificial data sets generated from three distributions: Gaussian, Cauchy and comb (Marron and Wand, 1992). We compare the performance of the self-consistent estimate SC with two kernel estimators: the Gaussian kernel KG (see Silverman (1986)) and the trapezoidal kernel KT (see Politis (2003)), the latter being representative of non-strictly-positive estimators. For the Cauchy distribution we also show the results for a kernel with a locally adaptive bandwidth APT (see Silverman (1986) and Hossjer (1996)). For each density $f(x)$ and each estimator, we generate 100 samples, each composed of N points randomly and independently drawn from $f(x)$. The performance is evaluated in terms of the mean integrated square error, equation (22) in Appendix A, where the mean is calculated over the 100 sample realizations. We repeat the above procedure for various values of N, ranging from 10^2 to 10^6. In addition to the simulation results, we also present the theoretical bound OPT given by the optimal kernel (3), i.e. the best that an estimate of the type of equation (2) can achieve (see also equation (28) in Appendix A). In the case of the Gaussian density we also show the theoretical bound given by maximum likelihood, ML. The bandwidths for the Gaussian and trapezoidal kernels are chosen from the sample data following established empirical rules: $h=0.79\,\mathrm{iq}N^{-1/5}$ for the Gaussian kernel (where iq is the interquartile range; see equation (3.29) in Silverman (1986)), whereas for the trapezoidal kernel

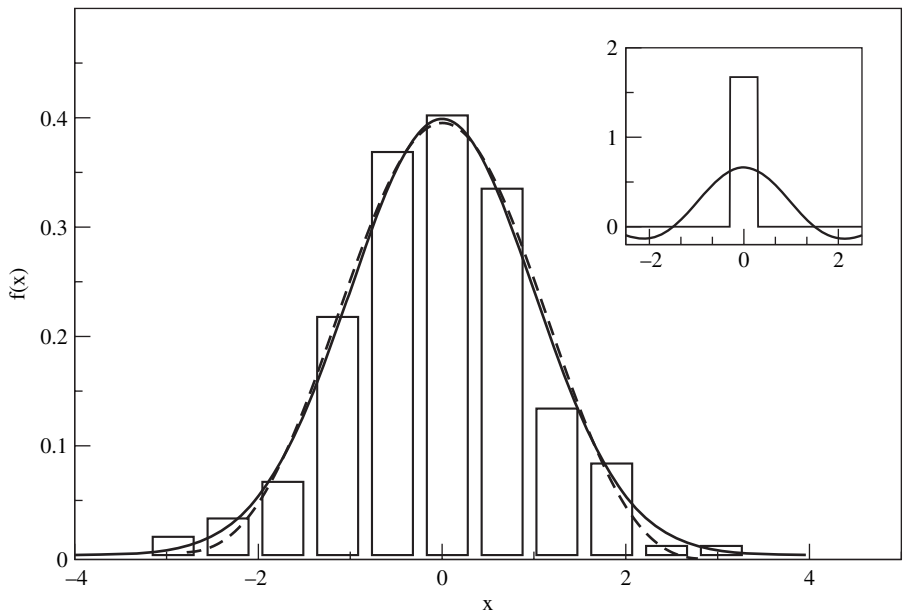

Fig. 2. Illustrative example of an estimate of a Gaussian density from $N=200$ sample points (the inset shows the self-consistent kernel and the bin width of the histogram (Scott, 1979)): ———, true density; – – –, self-consistent estimate, compared with a histogram having optimal bin width

$h^{-1} = 2 \min\{m : |\Delta(m+s)|^2 < c^2 \log(N)/N, \forall s \in (0, K_N)\}$, where we chose $K_N = \log(N)$ and we averaged results over $2^{-2} \leqslant c^2 \leqslant 2^2$ (see Politis (2003)).

First, we study the Gaussian distribution,

$$f(x) = \frac{\exp(-x^2/2)}{\sqrt{(2\pi)}}.$$

Fig. 3 shows the mean integrated square error $E(I)$ as a function of the sample size N, for the self-consistent estimate SC, the Gaussian and trapezoidal kernels KG and KT, the optimal bound OPT and the maximum likelihood ML. For the Gaussian kernel with a fixed bandwidth h, the exact expression for the error is

$$E(I_{\mathrm{KG}}) = \frac{1}{2\sqrt{\pi}} \left\{ \frac{1}{Nh} - \frac{1}{N\sqrt{(1+h^2)}} + 1 - \frac{2}{\sqrt{(1+h^2/2)}} + \frac{1}{\sqrt{(1+h^2)}} \right\}. \tag{14}$$

The average value of the bandwidth is $h = 1.06N^{-1/5}$, which gives an approximate value of the error $E(I_{\mathrm{KG}}) \simeq 0.33N^{-4/5}$. The error of the optimal estimate is given by equation (28), and for a Gaussian density is equal to

$$E(I_{\mathrm{OPT}}) = \frac{\{N/2(N-1)\sqrt{\pi}\}\mathrm{Li}_{1/2}(1-N) - 1}{N-1} \simeq \frac{\sqrt{\log(N)}}{\pi N} \tag{15}$$

where Li is the polylogarithm function, defined by $\mathrm{Li}_s(z) = \Sigma_{k=1}^{\infty} z^k/k^s$. The error for estimator ML is

$$E(I_{\mathrm{ML}}) = \frac{7}{16\sqrt{\pi}} N^{-1}. \tag{16}$$

Fig. 3 shows that the error of the SC estimator is consistently smaller than that for both kernel estimates, KG and KT. Both SC and KT approach the theoretical OPT scaling of approximately

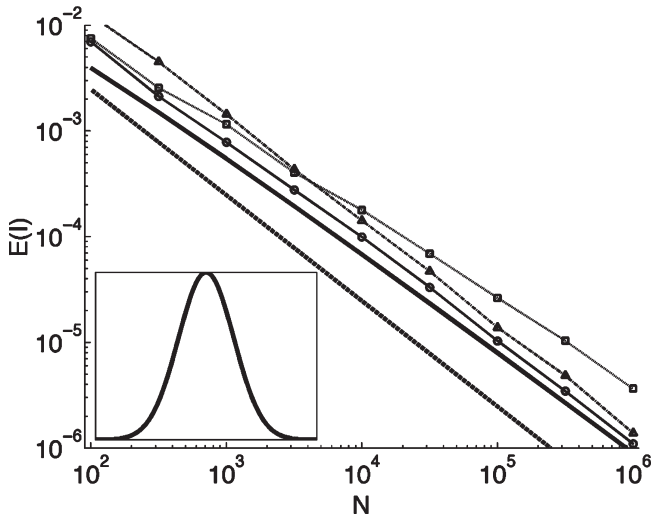

Fig. 3. Mean integrated square error $E(I)$ for the estimate of a Gaussian distribution (inset), as a function of the sample size N (standard errors are about 5% and smaller than the size of the symbols; each point is an average over 100 realizations of the sample): □, Gaussian kernel KG; Δ, trapezoidal kernel KT; ○; self-consistent estimate SC (which is applied without any prior knowledge); ———, optimal kernel OPT (which requires the power spectrum of the true density in advance); · · · · · · ·, maximum likelihood bound ML (which assumes that the density is Gaussian)

$N^{-1}\sqrt{\log(N)}$ (see Davis (1977)), whereas ML scales as approximately N^{-1}. We stress that both ML and OPT require prior knowledge of the density to be estimated. The former needs to know that the density is Gaussian and the latter needs its spectrum in advance, whereas the self-consistent method achieves the same scaling without any prior assumption.

The second application is the estimate of a Cauchy distribution, $f(x) = \{\pi(1+x^2)\}^{-1}$. The interest of this case comes from the difficulties of binning long-tailed distributions, especially when the variance diverges. For the Gaussian kernel with a fixed bandwidth, the error is

$$E(I_{\mathrm{KG}}) = \frac{1}{2Nh\sqrt{\pi}} + \frac{1}{2\pi} + \frac{N-1}{2Nh\sqrt{\pi}} \exp(h^{-2}) \operatorname{erfc}(h^{-1}) - \frac{\sqrt{2}}{h\sqrt{\pi}} \exp(2h^{-2}) \operatorname{erfc}(h^{-1}\sqrt{2}). \quad (17)$$

The average bandwidth is $h = 1.58N^{-1/5}$, for which $E(I_{\mathrm{KG}}) \simeq 0.55N^{-4/5}$. The error of the optimal estimate, equation (28), in the case of the Cauchy distribution, yields

$$E(I_{\mathrm{OPT}}) = \frac{\{N/2\pi(N-1)\}\log(N) - 1/2\pi}{N-1} \simeq \frac{\log(N)}{2\pi N}. \quad (18)$$

To cope with long-tailed distributions, kernel methods have been generalized to the 'adaptive' kernel, which allows the bandwidth to vary locally according to a first estimate of the density (Silverman, 1986; Hossjer, 1996). In Fig. 4 we show the mean-square error as a function of N for the SC, KG and KT estimators, the adaptive kernel estimator APT and the OPT bound. For small data sets, the adaptive kernel method performs best. However, the self-consistent method still shows better scaling with N, and its error is lower for large sample sizes, with a crossover occurring for N between 10^4 and 10^5. Again, the SC estimator performs better than both kernel estimator KG and KT. Both SC and KT approach the OPT scaling. We stress that the adaptive method requires some prior knowledge: it is used when we know that the distribution is long tailed and, again, the optimal estimate requires prior knowledge of its power spectrum. Conversely, we applied the self-consistent method blindly, in the same way as we did in the Gaussian case.

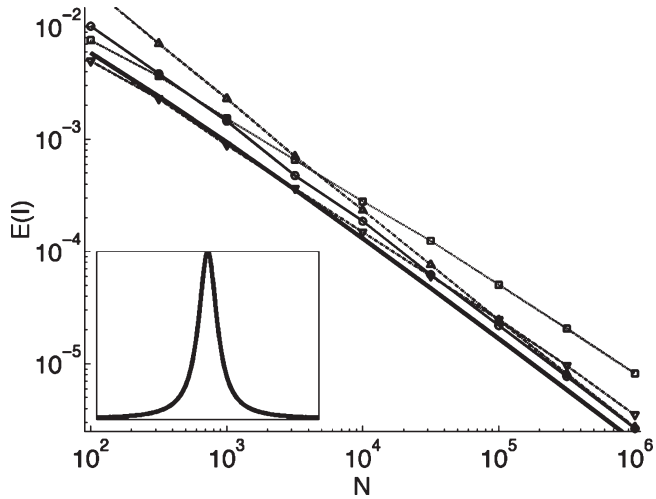

Fig. 4. Mean-square error $E(I)$ for the estimate of a Cauchy distribution (inset) as a function of the sample size N (standard errors are about 5% and smaller than the size of the symbols; each point is an average over 100 realizations of the sample): □, Gaussian kernel KG; Δ, trapezoidal kernel KT; ○, self-consistent estimate SC (which is applied without any prior knowledge); ———, optimal kernel OPT (which requires the power spectrum of the true density in advance); ∇, bound APT (which applies to long-tailed distributions)

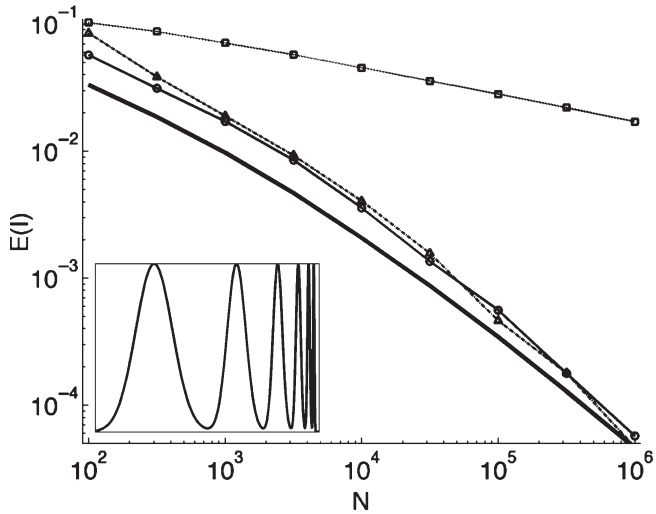

Fig. 5. Mean-square error $E(I)$ for the estimate of a comb distribution (inset) as a function of the sample size N (standard errors are about 1% and smaller than the size of the symbols; each point is an average over 100 realizations of the sample): □, Gaussian kernel KG; Δ, trapezoidal kernel KT; ○, self-consistent estimate SC (which is applied without any prior knowledge); ———, optimal kernel OPT (which requires the power spectrum of the true density in advance)

The last application is the comb distribution (Marron and Wand, 1992), which is multimodal, where different modes have different widths, and its transform is affected by a large interval of frequencies. Fig. 5 shows the results for the estimators SC, KG and KT and the OPT bound (which is computed numerically by using equation (28)). Note that KG performs poorly in this case, because the empirical bandwidth cannot capture the large interval of frequencies of the comb distribution. Instead, both SC and KT embrace an appropriate interval of frequencies, estimated from the empirical characteristic function, they perform much better than KG and approach the OPT scaling.

Finally, we remark that the self-consistent estimator may perform poorly when applied to particular families of density functions. As explained in Appendix C, if the density is not square integrable or its characteristic function is not integrable, then the self-consistent estimate is not guaranteed to converge to the true density for large N. Simulations (which are not shown) suggest that the self-consistent estimator does not perform well when applied to the box function (whose characteristic function is not integrable) and to the χ^2-distribution with 1 degree of freedom (which is not square integrable).

4. Discussion

We presented a method that estimates a density in a self-consistent way from a finite sample. This approach produces a unique estimate where no parameters have to be adjusted, and the only prior is the belief in the self-consistent procedure. In contrast, the cases in which there is a widely accepted theoretical framework for the system at hand would rather point to a Bayesian, parametric approach in which a specific form for the density is postulated.

The self-consistent estimate converges to the true density for large N and it has a better performance and scaling than existing methods for all the examples that we studied. Together with the simplicity of its implementation, these features make it preferable for applications,

especially when large data sets are available. However, if the true density is not square integrable or its characteristic function is not integrable, the self-consistent estimate is not guaranteed to converge and it may perform poorly.

A frequency filter emerges naturally in the derivation of the self-consistent estimate, as a function of the data sample, and prevents overfitting of the data. When long tails or power law behaviour is suspected, one is tempted to use logarithmic binning, but that has been shown to be highly inaccurate (Clauset *et al.*, 2009; Goldstein *et al.*, 2004) and sometimes leads to the conclusion that the density has power law tails when it has not. The self-consistent method constitutes a good alternative when there are no solid theoretical grounds to assume (or discard) a power law.

For small values of N, adaptive bandwidth kernels may perform better than the self-consistent kernel (see Fig. 4). If the bandwidth is allowed to vary locally, the performance of the estimator is not bounded by the optimal kernel performance, since it does not belong to estimates of the type of equation (2), and can in principle perform better. Future studies will be devoted to extend the self-consistent approach to estimates in which the kernel may vary locally. For example, a generic non-adaptive estimate can be converted into an adaptive estimate by a transformation of x leading to a space with uniform measure (Ruppert and Kline, 1994; Periwal, 1997).

The self-consistent method can be applied concretely to any problem in which a density is sampled and there is no prior knowledge of its functional shape. For instance, the mode of the instantaneous firing rate of a cortical neuron may indicate whether the choice of a behaving monkey is triggered by a memorized object or a spatial cue (Olson *et al.*, 2000). Another application could be the density estimate of the spacings between 0s of the Riemann zeta function, based on a very large numerical data set, which could help to investigate the related mathematical conjectures (Odlyzko, 1987). Finally, the method can be used to analyse samples obtained from Monte Carlo simulations (Binder, 1986).

We remark that the method could not be applied to the case of integer numbers. In general, when data points are uniformly spaced by multiples of any constant length, the Fourier transform of their density is periodic. An amplitude threshold and a cut-off frequency would be meaningless in that case, and this would preclude any practical application. However, when data points are integers there is usually no need for filtering, and it is sufficient to use a histogram, counting the occurrences of each number.

Finally, a straightforward generalization is to consider a finite interval $[a, b]$ instead of the entire line $(-\infty, \infty)$, by using Fourier series instead of Fourier transforms. Another possibility is to apply the method to high dimensional distributions. In our derivation, the relevant variables are scalar quantities, and the d-dimensional analogy is obtained by using the d-dimensional Fourier transform, i.e. by performing the integral $\{1/(2\pi)^d\} \int_{\mathbb{R}^d} \mathrm{d}t^d$ instead of $(1/2\pi) \int_{\mathbb{R}} \mathrm{d}t$. Among many possible implementations, the method could then be applied to the analysis of multielectrode neuronal recordings (Brown *et al.*, 2004), multivariate financial data (Breymann *et al.*, 2003) and reconstruction of Ramachandran angle distributions (Kleywegt and Jones, 1996).

Acknowledgements

AB thanks Alfonso Sutera for introducing him to the issues of non-parametric density estimation. The authors further thank Yali Amit, Massimo Cencini, Rishidev Chaudhuri, Andrew Jackson, John Murray, Umberto Picchini and Angelo Vulpiani for their useful comments on a preliminary version of the manuscript.

Appendix A: Derivation of the optimal kernel

In this section, we show that a unique optimal convolution kernel can be derived as a function of the power spectrum of the density to be estimated. A similar result has been presented in Watson and Leadbetter (1963). Given a sample of N data points (real numbers), denoted by $\{X_j\}$ $(j=1,\ldots,N)$, each independently drawn from a probability density distribution $f(x)$, we write the estimate as

$$\hat{f}(x)=\frac{1}{N}\sum_{j=1}^{N}K(x-X_j). \tag{19}$$

The true density f is assumed to be normalized, i.e. $\int_{-\infty}^{\infty}f(x)\,\mathrm{d}x=1$.

We look for a kernel $K(x)$ such that the estimate (19) minimizes the mean integrated square error

$$E(I)=E\left[\int_{-\infty}^{\infty}\{\hat{f}(x)-f(x)\}^2\,\mathrm{d}x\right] \tag{20}$$

where $E(\cdot)$ denotes an average over all the possible realizations of the data sample $\{X_j\}$. To minimize expression (20), we follow a procedure for signal deconvolution (the Wiener filter (Wiener, 1949)). We introduce the Fourier transform of the unknown density (characteristic function)

$$\phi(t)=\int_{-\infty}^{\infty}\exp(\mathrm{i}tx)\,f(x)\,\mathrm{d}x. \tag{21}$$

The normalization condition now reads $\phi(0)=1$. We also call $\hat{\phi}(t)$ and $\kappa(t)$ the Fourier transforms of the estimate $\hat{f}(x)$ and of the kernel $K(x)$ respectively. The mean integrated square error (20) corresponds to the mean-square distance between the true density f and the estimate $\hat{f}$, in terms of the Euclidean metric in the Hilbert space L^2 (we assume that $f,\hat{f},\phi,\hat{\phi}\in L^2$). By means of Parseval's theorem, we rewrite equation (20) in Fourier space as

$$E(I)=\frac{1}{2\pi}E\left\{\int_{-\infty}^{\infty}|\hat{\phi}(t)-\phi(t)|^2\,\mathrm{d}t\right\}. \tag{22}$$

It is straightforward to perform the average in Fourier space. Applying the convolution theorem to equation (19), the transformed estimate is equal to $\hat{\phi}(t)=\kappa(t)\,\Delta(t)$, where

$$\Delta(t)=\frac{1}{N}\sum_{j=1}^{N}\exp(\mathrm{i}tX_j) \tag{23}$$

is the empirical characteristic function. Note that $\Delta\notin L^2$ for any finite value of N, but $\hat{\phi}\in L^2$ by assumption. Using $E(\Delta)=\phi$, and $E(|\Delta|^2)=|\phi|^2+N^{-1}(1-|\phi|^2)$, we can rewrite the error as

$$E(I)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\{N^{-1}|\kappa|^2(1-|\phi|^2)+|\phi|^2|1-\kappa|^2\}\,\mathrm{d}t. \tag{24}$$

Since $f(x)$ is a density, it is normalized and non-negative, which implies that $|\phi|^2\leqslant 1$. Then, the first term in the integral, proportional to $1/N$, is non-negative: it corresponds to the error due to the finite size of the sample, whereas the second term does not depend on N. These two sources of error are known as the error variance and error bias respectively (Silverman, 1986). Among the possible choices of the kernel, we search for the one minimizing the mean integrated square error. Since equation (24) is quadratic in κ, it is straightforward to find its global minimum, by setting to 0 the functional derivative of $E(I)$ with respect to κ, i.e.

$$2\pi\frac{\delta E(I)}{\delta\kappa^*}=N^{-1}\kappa(1-|\phi|^2)-|\phi|^2(1-\kappa)=0 \tag{25}$$

where the asterisk denotes complex conjugate. This yields a unique optimal kernel, which in Fourier space reads

$$\kappa_{\mathrm{opt}}(t)=\frac{N}{N-1+|\phi(t)|^{-2}}. \tag{26}$$

The optimal kernel satisfies the normalization condition $\kappa_{\mathrm{opt}}(0)=1$, because $\phi(0)=1$, and is a real function. Since the density f is real, then $|\phi(t)|=|\phi(-t)|$, which implies that $\kappa_{\mathrm{opt}}(t)$ is an even function. Then its antitransform $K_{\mathrm{opt}}(x)$, the optimal kernel in the real space, is also real and even and, because expression

(26) is non-negative, $K_{\mathrm{opt}}(x)$ takes the maximum value at $x=0$, i.e. at the co-ordinate of each data point in equation (19).

Using the expression for the transformed optimal kernel, equation (26), we rewrite the estimate, equation (19), in Fourier space, as

$$\hat{\phi}(t)=\Delta(t)\frac{N}{N-1+|\phi(t)|^{-2}}, \tag{27}$$

which we call the 'optimal estimate'. The optimal estimate satisfies the normalization condition, $\hat{\phi}(0)=1$, because $\Delta(0)=1$ and $\phi(0)=1$. For infinite sample size ($N\to\infty$), the optimal estimate reduces to the true density with probability 1, i.e. $\hat{\phi}(t)\to\phi(t)$, because $\Delta(t)\to\phi(t)$ (see Csorgo and Totik (1983) and Ushakov (1999) for the detailed sufficient conditions), whereas the fractional term, the kernel, tends to 1. This is because an infinite sample would reproduce the true density itself, without the need of any transformation of the data. For finite N, the optimal estimate cuts the frequencies that have less power in the true density, and hence are more subject to noise, i.e. frequencies t whose power is of the order $|\phi(t)|^2\simeq 1/N$ or less.

The above procedure is analogous to the derivation of the Wiener filter for signal deconvolution (Wiener, 1949): prior knowledge of the power spectrum of the signal, the unknown density and the noise establishes a unique criterion for optimal signal-to-noise separation (note that, once the signal spectrum is given, $|E(\Delta)|^2=|\phi|^2$, the assumption of independence of data points allows the noise spectrum to be written as a function of the signal, i.e. $E(|\Delta|^2)-|E(\Delta)|^2=N^{-1}(1-|\phi|^2)$).

We conclude this section by deriving the minimum square error obtained by the application of the optimal kernel, which will be useful for assessing the performance of practical applications of the method. By substituting the expression (26) of the optimal kernel in equation (24), the associated minimum square error can be written, after some algebra, as

$$\min_K\{E(I)\}=\frac{K_{\mathrm{opt}}^{(N)}(0)-K_{\mathrm{opt}}^{(1)}(0)}{N-1} \tag{28}$$

where we made explicit the dependence of the optimal kernel on the sample size N, by writing $K_{\mathrm{opt}}=K_{\mathrm{opt}}^{(N)}(x)$.

Finally, note that $|\phi(t)|^2\leqslant 1$ and $|\Delta(t)|^2\leqslant 1$, which imply from equation (27) that also the optimal estimate satisfies $|\hat{\phi}(t)|^2\leqslant 1$. Although this is a necessary condition for the antitransform $\hat{f}(x)$ to be non-negative, it is not sufficient, and $\hat{f}(x)$ is not guaranteed to be a non-negative density.

Appendix B: Derivation of the self-consistent estimate

In this section we derive the expression for the self-consistent estimate $\hat{\phi}_{\mathrm{sc}}$, equation (8), and we study its stability. We start from equation (6), the iterative map that we rewrite here

$$\hat{\phi}_{n+1}=\frac{N\Delta}{N-1+|\hat{\phi}_n|^{-2}}. \tag{29}$$

We search for a fixed point of the iteration, namely $\hat{\phi}_{\mathrm{sc}}$, such that

$$\hat{\phi}_{\mathrm{sc}}=\frac{N\Delta}{N-1+|\hat{\phi}_{\mathrm{sc}}|^{-2}}. \tag{30}$$

We derive in what follows the two solutions (beyond the null solution) of equation (30) and we show that only one solution is stable with respect to the iteration, equation (29). We start by taking the absolute value of equation (30), to obtain an equation for the single unknown variable $|\hat{\phi}_{\mathrm{sc}}|$ (note that $\hat{\phi}_{\mathrm{sc}}$ is complex valued). Then, we multiply the expression by the denominator and by $|\hat{\phi}_{\mathrm{sc}}|$ (leaving the null solution $\hat{\phi}_{\mathrm{sc}}=0$), obtaining a simple quadratic equation

$$(N-1)|\hat{\phi}_{\mathrm{sc}}|^2+1=N|\Delta||\hat{\phi}_{\mathrm{sc}}|. \tag{31}$$

Provided that $|\Delta|^2\geqslant 4(N-1)/N^2$, this equation has the following two solutions, denoted by the superscript '$\pm$'

$$|\hat{\phi}^{\pm}|=\frac{N|\Delta|}{2(N-1)}\left[1\pm\sqrt{\left\{1-\frac{4(N-1)}{N^2|\Delta|^2}\right\}}\right]. \tag{32}$$

This solution gives the absolute value of $\hat{\phi}^{\pm}$. By replacing this expression back in the right-hand side of equation (30), we obtain the solution for $\hat{\phi}^{\pm}$, given by

$$\hat{\phi}^{\pm} = \frac{N\Delta}{2(N-1)}\left[1 \pm \sqrt{\left\{1 - \frac{4(N-1)}{N^2|\Delta|^2}\right\}}\right]. \tag{33}$$

Whereas $\hat{\phi}^{+}$ is normalized, $\hat{\phi}^{-}$ is not, i.e., when $t=0$, $\Delta(0)=1$ implies that $\hat{\phi}^{+}(0)=1$, whereas $\hat{\phi}^{-}(0)=1/(N-1)$. The two solutions are of very different magnitudes: for large N the solution $\hat{\phi}^{-}$ vanishes ($|\hat{\phi}^{-}| \simeq 1/N|\Delta|$), whereas $\hat{\phi}^{+}$ stays finite ($|\hat{\phi}^{+}| \simeq |\Delta|$). For all values of $|\Delta|$ for which the solutions $\hat{\phi}^{\pm}$ exist, the following relationship holds: $|\hat{\phi}^{+}||\hat{\phi}^{-}| = 1/(N-1)$.

We show that $\hat{\phi}^{+}$ is a stable solution of the iteration, whereas $\hat{\phi}^{-}$ is unstable. This can be seen by taking the absolute value of equation (29) and computing the derivative

$$\left.\frac{\mathrm{d}|\hat{\phi}_{n+1}|}{\mathrm{d}|\hat{\phi}_n|}\right|_{|\hat{\phi}_n|=|\hat{\phi}^{\pm}|} = 1 \mp \sqrt{\left\{1 - \frac{4(N-1)}{N^2|\Delta|^2}\right\}}. \tag{34}$$

This implies that, provided that the two solutions exist, i.e. provided that $|\Delta|^2 > 4(N-1)/N^2$, then $\hat{\phi}^{+}$ is stable (derivative smaller than 1) and $\hat{\phi}^{-}$ is unstable (derivative larger than 1). When $|\Delta|^2 = 4(N-1)/N^2$, the two solutions annihilate in a saddle node bifurcation. For $|\Delta|^2 < 4(N-1)/N^2$ only the null solution, $\hat{\phi}_{\mathrm{sc}} = 0$, is available. That is always stable, as can be checked by computing

$$\lim_{|\hat{\phi}_n| \to 0} (\mathrm{d}|\hat{\phi}_{n+1}|/\mathrm{d}|\hat{\phi}_n|) = 0. \tag{35}$$

In summary, when $|\Delta|^2 \geqslant 4(N-1)/N^2$, the iteration (29) reaches $\hat{\phi}_{\mathrm{sc}} = 0$ for $|\hat{\phi}_0| < |\hat{\phi}^{-}|$ and $\hat{\phi}_{\mathrm{sc}} = \hat{\phi}^{+}$ for $|\hat{\phi}_0| \geqslant |\hat{\phi}^{-}|$, whereas, for $|\Delta|^2 < 4(N-1)/N^2$, the unique, globally stable solution is $\hat{\phi}_{\mathrm{sc}} = 0$. As described in the main text, we define the set B of t-values as

$$B = \left\{t : |\Delta(t)|^2 \geqslant \frac{4(N-1)}{N^2}\right\}. \tag{36}$$

Hence, $\hat{\phi}_{\mathrm{sc}}(t) = 0$ when $t \notin B$. However, $t \in B$ does not guarantee that $\hat{\phi}_{\mathrm{sc}}(t) = \hat{\phi}^{+}(t)$, since a small initial guess $|\hat{\phi}_0(t)| < |\hat{\phi}^{-}(t)|$ would determine $\hat{\phi}_{\mathrm{sc}}(t) = 0$. Then, we define A as the set of t-values for which $|\hat{\phi}_0(t)| \geqslant |\hat{\phi}^{-}(t)|$, and hence $\hat{\phi}_{\mathrm{sc}}(t) = \hat{\phi}^{+}(t)$. The arbitrary choice of the initial guess translates into the arbitrary choice of the set A, provided that $A \subseteq B$, as required by the existence of the non-zero solutions. This is summarized in equation (8) and concludes the derivation.

We conclude this section by calculating the mean and variance of the self-consistent estimate $\hat{f}_{\mathrm{sc}}$, equation (11). In a finite neighbourhood of $t=0$, $\hat{\phi}_{\mathrm{sc}}(t) = \hat{\phi}^{+}(t)$, which is continuous and infinitely differentiable at $t=0$. Then, the mean and variance can be computed by the derivatives of $\hat{\phi}^{+}$ and $|\hat{\phi}^{+}|^2$ at $t=0$, namely

$$E(x) = -\mathrm{i}\left.\frac{\mathrm{d}\hat{\phi}^{+}}{\mathrm{d}t}\right|_{t=0} = \left.\frac{\mathrm{d}\hat{\phi}^{+}}{\mathrm{d}\Delta}\right|_{\Delta=1}\left(-\mathrm{i}\left.\frac{\mathrm{d}\Delta}{\mathrm{d}t}\right|_{t=0}\right), \tag{37}$$

$$\mathrm{var}(x) = -\frac{1}{2}\left.\frac{\mathrm{d}^2|\hat{\phi}^{+}|^2}{\mathrm{d}t^2}\right|_{t=0} = \left.\frac{\mathrm{d}|\hat{\phi}^{+}|^2}{\mathrm{d}|\Delta|^2}\right|_{|\Delta|^2=1}\left(-\frac{1}{2}\left.\frac{\mathrm{d}^2|\Delta|^2}{\mathrm{d}t^2}\right|_{t=0}\right) \tag{38}$$

where we used the chain rule and the fact that $|\Delta|^2$ is an even function; hence

$$\left.\frac{\mathrm{d}|\Delta|^2}{\mathrm{d}t}\right|_{t=0} = 0.$$

It is straightforward to show that the terms in round brackets in equations (37) and (38) are respectively equal to the sample mean and sample variance. Because $|\Delta|$ is even, then

$$\left.\frac{\mathrm{d}\hat{\phi}^{+}}{\mathrm{d}\Delta}\right|_{\Delta=1} = 1$$

and, by differentiating equation (31), we obtain

$$\left.\frac{\mathrm{d}|\hat{\phi}^{+}|^2}{\mathrm{d}|\Delta|^2}\right|_{|\Delta|^2=1} = \frac{N}{N-2}.$$

Finally, equations (37) and (38) can be rewritten as

$$E(x) = \frac{1}{N}\sum_{j=1}^{N} X_j, \tag{39}$$

$$\mathrm{var}(x) = \frac{1}{N-2}\sum_{j=1}^{N}\{X_j - E(x)\}^2. \tag{40}$$

Appendix C: Asymptotic convergence of $\hat{f}_{\mathrm{sc}}$

In this section we investigate the asymptotic (large N) behaviour of the self-consistent estimator. We study the sufficient conditions for the estimate $\hat{f}_{\mathrm{sc}}$ to converge to the true density f for $N \to \infty$. In particular, we prove the following theorem.

C.1. Theorem 1

If the true density $f(x)$ is square integrable and its transform is integrable, then the self-consistent density estimate $\hat{f}_{\mathrm{sc}}(x)$, which is defined by equations (5) and (8)–(11), converges almost surely to the true density for large N, under the additional assumptions

$$\lim_{N\to\infty}(t^*) = \infty, \tag{41}$$

$$\lim_{N\to\infty}(t^*/\sqrt{N}) = 0. \tag{42}$$

C.2. Proof

Because the true density $f(x)$ and the self-consistent estimate $\hat{f}_{\mathrm{sc}}(x)$ are both square integrable, we can express them as Fourier transforms of respectively $\phi(t)$ and $\hat{\phi}_{\mathrm{sc}}(t)$. By assumption, the characteristic function is integrable, i.e.

$$\int |\phi(t)|\,\mathrm{d}t < \infty. \tag{43}$$

In the following sequence of inequalities, we find an upper bound for the difference between the true density and its estimate, and we use the fact that $\hat{\phi}_{\mathrm{sc}}(t) = 0$ for $|t| > t^*$. To prove theorem 1, we show that the upper bound tends to 0 for large N.

$$\begin{aligned}
|\hat{f}_{\mathrm{sc}}(x) - f(x)| &= \left|\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(-\mathrm{i}tx)\{\hat{\phi}_{\mathrm{sc}}(t) - \phi(t)\}\,\mathrm{d}t\right| \\
&\leqslant \frac{1}{2\pi}\int_{-\infty}^{\infty}|\exp(-\mathrm{i}tx)||\hat{\phi}_{\mathrm{sc}}(t) - \phi(t)|\,\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{\infty}|\hat{\phi}_{\mathrm{sc}}(t) - \phi(t)|\,\mathrm{d}t \\
&= \frac{1}{2\pi}\int_{-t^*}^{t^*}|\hat{\phi}_{\mathrm{sc}}(t) - \phi(t)|\,\mathrm{d}t + \frac{1}{2\pi}\int_{|t|>t^*}|\phi(t)|\,\mathrm{d}t \\
&= \frac{1}{2\pi}\int_{-t^*}^{t^*}|\hat{\phi}_{\mathrm{sc}}(t) - \Delta(t) + \Delta(t) - \phi(t)|\,\mathrm{d}t + \frac{1}{2\pi}\int_{|t|>t^*}|\phi(t)|\,\mathrm{d}t \\
&\leqslant \frac{1}{2\pi}\int_{-t^*}^{t^*}|\hat{\phi}_{\mathrm{sc}}(t) - \Delta(t)|\,\mathrm{d}t + \frac{1}{2\pi}\int_{-t^*}^{t^*}|\Delta(t) - \phi(t)|\,\mathrm{d}t + \frac{1}{2\pi}\int_{|t|>t^*}|\phi(t)|\,\mathrm{d}t
\end{aligned}$$

where $\Delta(t)$ is the empirical characteristic function; see equation (5). For increasing N, t^* increases following the limiting bounds given by equations (41) and (42). Then, the second integral in the last expression tends to 0 because of theorem 1 in Csorgo and Totik (1983), whereas the third integral tends to 0 because of the integrability of the characteristic function $\phi(t)$, equation (43). To prove the theorem, we must demonstrate that also the first integral tends to 0 for large N. We use the expression for $\hat{\phi}_{\mathrm{sc}}$, (equation (8)) and we denote by Δ^+ and Δ^- the set of values of t for which $|\Delta(t)|^2$ is respectively above or below the threshold set by equation (9), i.e. $|\Delta(t)|^2 \geqslant 4(N-1)/N^2$ or $|\Delta(t)|^2 < 4(N-1)/N^2$. Then, the first integral is equal to

$$\frac{1}{2\pi}\int\limits_{(-t^*,t^*)\cap\Delta^+}|\Delta(t)|\left(1-\frac{N}{2(N-1)}\left[1+\sqrt{\left\{1-\frac{4(N-1)}{N^2|\Delta(t)|^2}\right\}}\right]\right)\mathrm{d}t+\frac{1}{2\pi}\int\limits_{(-t^*,t^*)\cap\Delta^-}|\Delta(t)|\,\mathrm{d}t$$

where the term in large parentheses is non-negative. The first term can be expanded by means of the inequality $\sqrt{(1-x)}\geqslant 1-\sqrt{x}$ for $x\in(0,1)$, whereas in the second term we substitute the integrand with its maximum value in the interval, and we use the fact that the length of the interval is smaller than $2t^*$. Then, the above expression is smaller than

$$\begin{aligned}\frac{1}{2\pi}\int\limits_{(-t^*,t^*)\cap\Delta^+}&\left\{\frac{1}{\sqrt{(N-1)}}-\frac{|\Delta(t)|}{N-1}\right\}\mathrm{d}t+\frac{2t^*\sqrt{(N-1)}}{\pi N}\\&\leqslant\frac{1}{2\pi}\int\limits_{(-t^*,t^*)\cap\Delta^+}\left\{\frac{1}{\sqrt{(N-1)}}+\frac{|\Delta(t)|}{N-1}\right\}\mathrm{d}t+\frac{2t^*\sqrt{(N-1)}}{\pi N}\\&\leqslant\frac{t^*}{\pi}\left\{\frac{1}{\sqrt{(N-1)}}+\frac{1}{N-1}+\frac{2\sqrt{(N-1)}}{N}\right\}\end{aligned}$$

where we used $|\Delta(t)|\leqslant 1$ in the last inequality. Because of equation (42), the last expression tends to 0 for large N, thus proving the theorem.

C.3. Comment

Theorem 1 assumes that the true density is square integrable and its transform integrable. We expect the self-consistent estimate to perform poorly when applied to densities that do not meet those criteria, and its asymptotic consistency is not guaranteed. Examples of such distributions include the box function and the χ^2-distribution with 1 degree of freedom.

Condition (42) gives the limit on the frequency bound t^*, setting the maximum increase of t^* with N. In simulations, the magnitude of the bound t^* depends on the threshold of the empirical characteristic function equation (9), in such a way that its amplitude is larger than the threshold in half of the interval $(-t^*, t^*)$. Now we argue that this choice satisfies the assumptions of theorem 1, without giving a formal derivation.

To simplify the argument, we assume that the distribution is sufficiently smooth that the following limit exists:

$$\lim_{|t|\to\infty}\{t|\phi(t)|\}=0; \tag{44}$$

if this limit exists, then it must be equal to 0 as a consequence of the integrability of the characteristic function equation (43). We also approximate the empirical characteristic function $\Delta(t)$ with the true characteristic function $\phi(t)$, which is reasonable provided that N is large and t does not increases exponentially with N (see Csorgo and Totik (1983)). Then, the threshold condition (9) becomes $|\phi(t)|^2\geqslant 4(N-1)/N^2$. Hence, t^* increases in such a way that, for large N, the order of magnitude of the characteristic function is $|\phi(t^*)|\sim 1/\sqrt{N}$. Substituting this expression in equation (44) we obtain

$$\lim_{t^*\to\infty}\{t^*|\phi(t^*)|\}=\lim_{N\to\infty}(t^*/\sqrt{N})=0 \tag{45}$$

which corresponds to condition (42). However, since the true cumulative distribution is continuous, then $|\phi(t)|>0$ for any finite value of t, and this implies that $t^*\to\infty$ for large N, as required by condition (41).

In conclusion, theorem 1 gives the sufficient conditions for the asymptotic consistency of the estimate, especially concerning the finite interval of frequencies set by the bound t^*. From these arguments, we expect the recipe for t^* that was used in simulations to guarantee the asymptotic convergence of the estimate.

References

Berlinet, A. (1993) Hierarchies of higher order kernels. *Probab. Theor. Reltd Flds*, **94**, 489–504.

Bialek, W., Callan, C. G. and Strong, S. P. (1996) Field theories for learning probability distributions. *Phys. Rev. Lett.*, **77**, 4693–4697.

Binder, K. (1986) *Monte Carlo Methods in Statistical Physics*. New York: Springer.

Bowman, A. W. (1984) An alternative method of cross-validation for the smoothing of density estimates. *Biometrika*, **71**, 353–360.

Breymann, W., Dias, A. and Embrecht, P. (2003) Dependence structures for multivariate high frequency data in finance. *Quant. Finan.*, **3**, 1–14.
Brown, E. N., Kass, R. E. and Mitra, P. P. (2004) Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nat. Neursci.*, **7**, 456–461.
Clauset, A., Shalizi, C. R. and Newman, M. E. J. (2009) Power-law distributions in empirical data. *SIAM Rev.*, **51**, 661–703.
Csorgo, S. and Totik, V. (1983) On how long interval is the empirical characteristic function uniformly consistent? *Acta Sci. Math.*, **45**, 141–149.
Davis, K. B. (1977) Mean integrated square error properties of density estimates. *Ann. Statist.*, **5**, 530–535.
Devroye, L. (1992) A note on the usefulness of superkernels in density estimation. *Ann. Statist.*, **20**, 2037–2056.
Efromovich, S. (2008) Adaptive estimation of and oracle inequalities for probability densities and characteristic functions. *Ann. Statist.*, **36**, 1127–1155.
Glad, I. K., Hjort, N. L. and Ushakov, N. G. (2003) Correction of density estimators that are not densities. *Scand. J. Statist.*, **30**, 415–427.
Glad, I. K., Hjort, N. L. and Ushakov, N. G. (2007) Density estimation using the sinc kernel. *Prepr. Statist.*, **2**, 1–20.
Goldstein, M. L., Morris, S. A. and Yen, G. G. (2004) Problems with fitting to the power-law distribution. *Eur. Phys. J.* B, **41**, 255–258.
Good, I. J. and Gaskins, R. A. (1971) Nonparametric roughness penalties for probability densities. *Biometrika*, **58**, 255–277.
Hall, P. and Marron, J. S. (1987) Choice of kernel order in density estimation. *Ann. Statist.*, **16**, 161–173.
Holy, T. E. (1997) Analysis of data from continuous probability distributions. *Phys. Rev. Lett.*, **79**, 3545–3548.
Hossjer, O. (1996) Asymptotic bias and variance for a general class of varying bandwidth density estimators. *Probab. Theor. Reltd Flds*, **105**, 159–192.
Kass, R. E., Ventura, V. and Brown, E. N. (2005) Statistical issues in the analysis of neuronal data. *J. Neurphysiol.*, **94**, 8–25.
Kleywegt, G. J. and Jones, T. A. (1996) Phi/psi-chology: Ramachandran revisited. *Structure*, **4**, 1395–1400.
Marron, J. S. and Wand, M. P. (1992) Exact mean integrated square error. *Ann. Statist.*, **20**, 712–736.
Odlyzko, A. M. (1987) On the distribution of spacings between zeros of the Zeta function. *Math. Computn*, **48**, 273–308.
Olson, C. R., Gettner, S. N., Ventura, V., Carta, R. and Kass, R. E. (2000) Neuronal activity in macaque supplementary eye field during planning of saccades in response to pattern and spatial cues. *J. Neurphys.*, **84**, 1369–1384.
Parzen, E. (1962) On estimation of a probability density function and mode. *Ann. Math. Statist.*, **33**, 1065–1076.
Periwal, V. (1997) Reparametrization invariant statistical inference and gravity. *Phys. Rev. Lett.*, **78**, 4671–4674.
Politis, D. N. (2003) Adaptive bandwidth choice. *Nonparam. Statist.*, **15**, 517–533.
Ripley, B. D. and Sutherland, I. (1990) Finding spiral structures in images of galaxies. *Phil. Trans. R. Soc. Lond.*, **332**, 477–485.
Ruppert, D. and Kline, B. H. (1994) Bias reduction in kernel density estimation by smoothed empirical transformations. *Ann. Statist.*, **22**, 185–210.
Schmidt, D. M. (2000) Continuous probability distributions from finite data. *Phys. Rev.* E, **61**, 1052–1055.
Scott, D. (1979) On optimal and data-based histograms. *Biometrika*, **66**, 605–610.
Silverman, B. W. (1986) *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.
Ushakov, B. W. (1999) *Selected Topics in Characteristic Functions*. Utrecht: VSP.
Wand, M. P. and Jones, M. C. (1995) *Kernel Smoothing*. London: Chapman and Hall.
Watson, G. S. and Leadbetter, M. R. (1963) On the estimation of the probability density. *Ann. Math. Statist.*, **34**, 480–491.
Wiener, N. (1949) *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. New York: Wiley.

nature neuroscience

A reservoir of time constants for memory traces in cortical neurons

Alberto Bernacchia, Hyojung Seo, Daeyeol Lee & Xiao-Jing Wang

According to reinforcement learning theory of decision making, reward expectation is computed by integrating past rewards with a fixed timescale. In contrast, we found that a wide range of time constants is available across cortical neurons recorded from monkeys performing a competitive game task. By recognizing that reward modulates neural activity multiplicatively, we found that one or two time constants of reward memory can be extracted for each neuron in prefrontal, cingulate and parietal cortex. These timescales ranged from hundreds of milliseconds to tens of seconds, according to a power law distribution, which is consistent across areas and reproduced by a 'reservoir' neural network model. These neuronal memory timescales were weakly, but significantly, correlated with those of monkey's decisions. Our findings suggest a flexible memory system in which neural subpopulations with distinct sets of long or short memory timescales may be selectively deployed according to the task demands.

In economic behavior, choices that have a higher reward expectation are favored and adaptive decision making depends on our ability to learn reward expectation through past rewards associated with our actions. The neural mechanisms underlying this process have been the subject of growing interest, as they could provide important insights on how learning occurs in the brain and how humans and other animals make economic decisions. Neural correlates of reward valuation have been observed in different studies[1–3] and interpreted in the framework of reinforcement learning theory[4,5]. In the reinforcement learning model, reward expectation is computed by weighting the previous rewards through a temporal filter, which quantifies the memory trace of rewards. The optimal duration of the filter (memory) depends on the predictability of the environment. If the payoffs for the same option change often and unpredictably, then rewards should be filtered on short timescales to track the fast changes in a volatile environment; in contrast, if past rewards reliably predict future ones, then they should be filtered on long timescales to exploit a stable environment[6,7]. The neural mechanism underlying switching between long and short time constants for computing reward expectation remains poorly understood.

On which timescale does the brain filter rewards? To date, a few studies have estimated the time constant of this filter from behavior and assessed how past rewards affect choice selection[8–12], but the neural mechanisms responsible for such timescales are still unknown. To address this issue, we analyzed the activity of cortical neurons in monkeys performing a competitive game task. Using a method based on the idea that reward memory modulates neural activity multiplicatively, we found that memory time constants can be extracted from the activity of single neurons. We found that a different timescale for reward memory can be associated with each recorded neuron and that there is a wide range of timescales across neurons, obeying a power law distribution. The same distribution is found across three different cortical areas: anterior cingulate cortex (ACCd), dorsolateral prefrontal cortex (DLPFC) and lateral intraparietal cortex (LIP). Hence, each area is endowed with a reservoir of time constants for reward memory, which are distributed heterogeneously across neurons.

We found that the time constants estimated from pairs of simultaneously recorded neurons are uncorrelated, implying that our results cannot be explained by a single time constant for all neurons that changes slowly over time. On the other hand, our analysis of an animal's behavior suggests that the timescale over which reward events affect decisions changes across experimental sessions, possibly reflecting the animal's attempt to increase its payoff by exploring different strategies. The time constants for reward memory at the behavioral and neuronal levels were weakly correlated across experimental sessions. Finally, we found that a randomly connected circuit model, akin to a reservoir network[13–15], can reproduce the observed distribution of timescales, provided that the network operates at the critical point (or edge of chaos)[16–18]. Taken together, these findings suggest a distributed, flexible neural system for reward valuation and memory.

RESULTS

Multiplicative memory traces in cortical neurons

We analyzed single-neuron activity recorded from three cortical areas, ACCd[19] (154 neurons), DLPFC[20] (322 neurons) and LIP[21] (205 neurons) of six monkeys performing a matching pennies task[11,22] (**Fig. 1a**). In each trial, the monkey chose one of two targets by shifting its gaze and the computer made its choice by simulating a rational opponent; the animal received reward if its choice matched that of the computer. We computed firing rates of each neuron by counting the spikes in 12 time intervals of 250 ms (**Fig. 1a**), which are referred to as epochs. This includes six epochs (1.5 s) before saccade initiation (pre-fixation, fore-period and delay) and six epochs (1.5 s) after saccade completion (choice fixation, feedback and post-feedback). Consistent with previous studies[23–25], we found that the activity of neurons varied substantially in different trial epochs (99% of neurons,

Department of Neurobiology and Kavli Institute of Neuroscience, Yale University School of Medicine, New Haven, Connecticut, USA. Correspondence should be addressed to X.-J.W. (xjwang@yale.edu).

Received 18 October 2010; accepted 11 January 2011; published online 13 February 2011; doi:10.1038/nn.2752

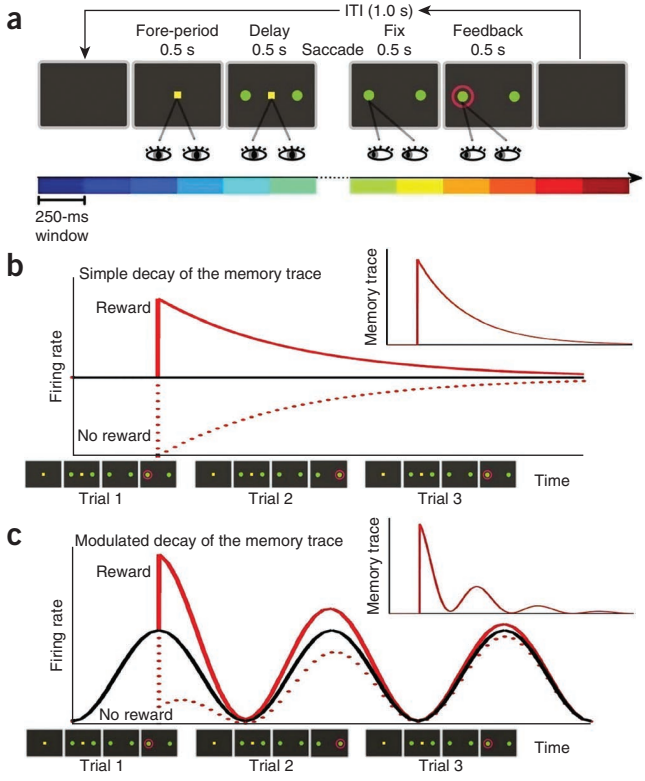

Figure 1 Behavioral task and schematic illustration of memory traces. (**a**) In the matching pennies task, the monkey was required to fixate a central spot during the fore-period (500 ms) and delay period (500 ms) while the two choice targets (green disks) were displayed. The central spot then disappeared and the monkey made a saccadic eye movement to one of the two choice targets and maintained its gaze on the chosen target for 500 ms (choice fixation). A red ring appearing around the correct target revealed the computer's choice, and if it matched the animal's choice (as illustrated), reward was delivered 500 ms later. ITI, inter-trial interval. Colored bars at the bottom show the 12 250-ms intervals (epochs) used to compute the firing rates in the analysis. (**b**,**c**) Two hypothetical neurons. The neuron in **b** has a constant average firing rate (black line), whereas the firing rate of neuron in **c** depends on the trial epoch, repeating in each of the three consecutive trials. Red lines show the change in activity as a result of the outcome in the first trial (continuous line indicates reward, dashed line indicates no reward). The inset shows the memory trace of the reward, given by the difference between the red and black lines. The memory trace of the neuron in **b** shows a simple decay, whereas that of the neuron in **c** is multiplicatively modulated by the epoch-dependent activity.

675 of 681, ANOVA, $P < 0.05$). The time course of the activity in successive epochs differs substantially in different neurons.

We then examined the effect of reward on the activity of neurons. Neural activity in all three cortical areas carries the information of past reward events[19–22]. We characterized the memory trace of each neuron using a new approach (**Fig. 1b**,**c**). Consider the time course of neural activity in different epochs, averaged across trials and hence across reward/no reward conditions. The difference in activity from the average time course, triggered by a reward/no reward event, was defined as the memory trace of the reward, which is positive for one of the outcomes and negative for the other. Thus, if the average activity of a neuron is zero in a given epoch, then the change by either outcome must be zero, and the memory trace in that epoch is therefore also zero. Starting with this intuition, we hypothesized that the memory trace in a given epoch is proportional to the average firing rate in that epoch. In that case the memory trace is modulated (multiplied) by the average firing rate.

We define the epoch code as the firing rate averaged across all trials, as a function of the different epochs, denoted by $g(k)$ ($k = 1,\ldots,12$ epochs, in temporal order). In one neuron recorded in ACCd, firing rate decreased after the saccade to a chosen target and increased after the feedback period (**Fig. 2a**). To separate the contributions of epoch and reward memory to neural activity, we modeled the firing rate measured in trial n and epoch k, denoted by $\mathrm{FR}(n,k)$, as the sum of the epoch code $g(k)$ and a filter $f(n',k)$ convolved with the animal's reward history in previous trials (last five trials; in each trial, Rew = +1 indicates reward; Rew = −1 indicates no reward).

$$\mathrm{FR}(n,k) = g(k) + \sum_{n'=0:5} f(n',k)\cdot \mathrm{Rew}(n-n') \quad (1)$$

The filter f describes how the reward in a given trial affects neural activity in the subsequent trials, assuming that the effects of rewards in successive trials are additive. For example, $f(3,4)$ describes the effect of a reward after 3 trials during epoch 4. The filter f corresponds to our definition of memory trace (**Fig. 1b**,**c**); it reflects the deviation from the epoch-dependent time course $g(k)$ resulting from a reward event. Because Rew(n) is nearly a random sequence[11] of +1 and −1, averaging the firing rates over all trials recovers the epoch code $g(k)$ (sum over n of $\mathrm{FR}(n,k)$). We estimated the memory trace $f(n',k)$ by applying multiple linear regression to the data according to equation (1). In an example negative neuron (that is, reward decreases the activity of this neuron in subsequent trials), the memory trace does not decay monotonically, but its strength is modulated throughout the trial consistent with the epoch code (**Fig. 2b**). According to the multiplicative

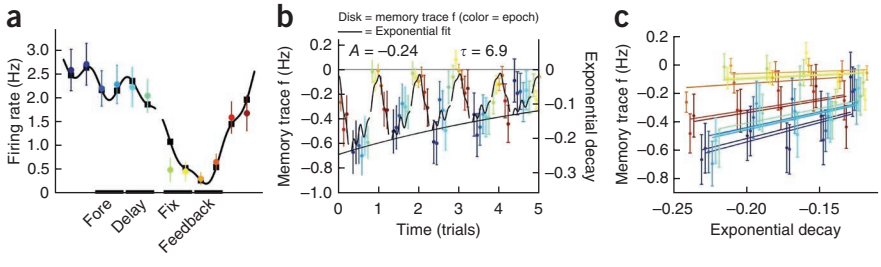

Figure 2 An example neuron in ACCd showing multiplicative modulation of memory traces by the epoch code. The colors in all panels denote trial epochs, following the format of **Figure 1a**. (**a**) The epoch code for an example neuron; that is, the firing rate computed in 12 250-ms epochs in a trial and averaged over all trials (black squares, interpolated by the black line, broken during the saccade). Colored disks correspond to the slopes fitted in **c** (error bars represent ±s.e.); their correlation with the epoch codes quantifies the multiplicative modulation and is referred to as the factorization index (0.97 in this example). (**b**) The memory trace f of past rewards in the same neuron, up to five trials in the past. Colored dots and error bars (±s.e.) show the results of the multiple linear regression model (equation (1)) and the black line is the exponential fit (equation (2), continuous line, exponential $\mathrm{ex}(t)$; broken line, modulated envelope $g\cdot\mathrm{ex}(t)$). The parameters for the fit are shown (A, amplitude; τ, timescale). (**c**) The memory trace f (from **b**), plotted as a function of the exponential function ex. The lines are least-squares fit, each line encompassing a particular epoch and all five trial lags. According to the factorization, the slopes should correspond to the epoch code, $f = g\cdot\mathrm{ex}$. The values of the slopes are plotted in **a** (colored squares) and compared with the epoch code $g(k)$.

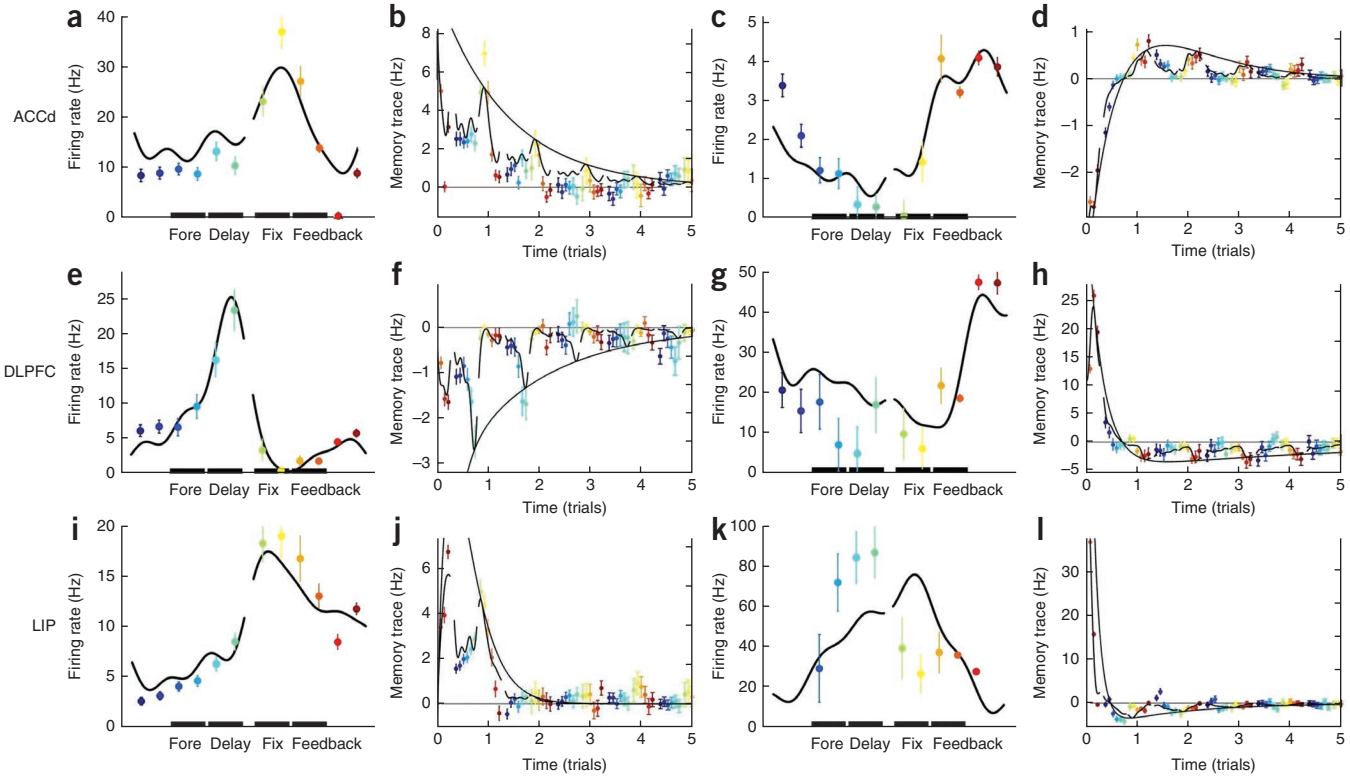

Figure 3 Firing rates and memory traces for six neurons, two for each of the three recorded areas. For each of the six neurons, epoch codes (first and third column) and memory traces (second and fourth column) are shown, presented as in **Figure 2a**,**b**. The second column shows monotonic decay of the memory trace and the fourth column shows biphasic memory traces (double exponential). Different neurons had different firing rates, both in magnitude and time course, and different types of memory decay, but they were all consistent with an exponential (single or double) decay of the memory modulated by the epoch code. The factorization indexes for those neurons are 0.98 (**a**,**b**), 0.91 (**c**,**d**), 0.98 (**e**,**f**), 0.84 (**g**,**h**), 0.97 (**i**,**j**) and 0.61 (**k**,**l**).

model (**Fig. 1b**,**c**), we assumed that the memory trace f is factorized into the epoch code $g(k)$ and an exponential function $\mathrm{ex}(t)$.

$$\mathrm{FR(n,k)} = \mathrm{g}(k) + \mathrm{g(k)} \cdot \sum_{n'=0:5} \mathrm{ex}(t) \times \mathrm{Rew}(n-n') \qquad (2)$$

The filter f considered in equation (1) is now replaced by the product of two factors $g(k)\cdot\mathrm{ex}(t)$, where $\mathrm{ex}(t) = Ae^{\frac{-t}{\tau}}$ is an exponential decay function and t is the time elapsed since the outcome (Online Methods). By applying this model to the example neuron (**Fig. 2a**,**b**), we obtained a timescale of memory decay τ = 6.9 trials and an amplitude A = −0.24. According to the factorization ($f = g\cdot\mathrm{ex}$), the constant of proportionality between the memory trace f and the exponential function ex, estimated in different epochs (**Fig. 2c**), should reproduce the epoch code $g(k)$. The epoch codes for the neuron closely followed these predictions, indicating that the factorization is nearly exact (**Fig. 2a**). The factorization index of a neuron, defined as the correlation coefficient between the epoch code and the proportionality constants (slopes), was 0.97 for this neuron.

The modulated decay of the memory trace was observed in the majority of the recorded neurons in all three cortical areas. In some cases, the sum of two exponential functions, $\mathrm{ex}(t) = A_1 e^{\frac{-t}{\tau_1}} + A_2 e^{\frac{-t}{\tau_2}}$, fitted the data better than a single exponential, in which case the memory trace often exhibited a biphasic characteristic (with A_1 and A_2 of the opposite sign; **Fig. 3**). Using the Bayesian Information Criterion, we found that the best fit was a single exponential for 269 neurons and double exponentials for 268 neurons, whereas the remaining 144 neurons were fitted best by a model with $\mathrm{ex}(t) = 0$. The latter is interpreted as no memory and the corresponding neurons were excluded from further analysis. We tested the validity of the fitting procedure by randomly reshuffling the order of trials in each session and we consistently found that 96% of neurons (656 of 681) showed no memory after reshuffling.

We examined the average firing rates and memory traces of all recorded neurons (examples shown in **Fig. 3**). Although the activity of most neurons is consistent with an exponential decay of the memory trace (79%, 537 of 681, single and double exponentials), a fraction of them did not show a modulation of the memory by the epoch code. This is quantified by the factorization index, which is significantly positive for approximately half of the neurons showing a memory effect (46%, 249 of 537, $P < 0.05$, t test). We found a small, but significant, difference in the fraction of neurons with memory across different areas (87% in ACCd, 75% in DLPFC and 78% in LIP, χ^2 test, $P = 0.01$).

We next investigated how the timescales of memory traces were distributed across neurons in different cortical areas. We determined that the distribution of timescales in all cortical areas could be fit with a power law with an exponent of −2 (**Fig. 4**). The power law implies that timescales are distributed in a wide range of values. In fact, for a power law distribution $\sim\tau^{-2}$, the variance increases with the sample size and, in principle, arbitrarily large timescales would be observed with a proportionally large increment in the number of recorded neurons. Note that the power law tail applies for timescales equal to or larger than one trial, which are those timescales that might be involved in memory (see below). About 20% of all recorded neurons (133 of 681)

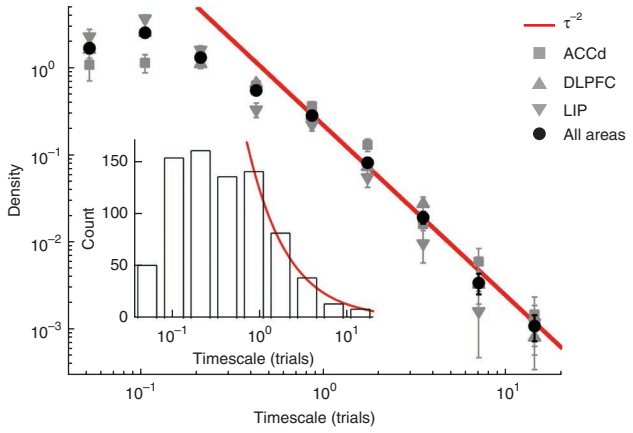

Figure 4 Distribution of the timescales characterizing the reward memory traces across neurons. Black disks show the density for the neurons in all three cortical areas in the corresponding bin; that is, the count of timescales divided by the bin length (error bars represent ±s.e.). The inset shows the count of the timescales in the same bins, in a linear scale (a total of 805 timescales). Grey markers show the density separately for each of the three different cortical areas (square, ACCd, 197 timescales; upward triangle, DLPFC, 362; downward triangle, LIP, 246). The red line (red curve in the inset) shows a power law fit (exponent = −2).

had timescale larger than one trial (29% in ACCd, 19% in DLPFC and 13% in LIP; χ^2 test, $P = 0.0005$; see **Supplementary Fig. 1c–e**). Because the timescales from one- (τ) and two-exponential functions (τ_1, τ_2) were distributed similarly (**Supplementary Fig. 1a,b**), we pooled all timescales (a total of 805 timescales from 269 single exponential and 268 double exponential; that is, 269 τ, 268 τ_1 and 268 τ_2). ACCd contributed 197 timescales from 71 single exponential and 63 double exponential functions (71 τ, 63 τ_1 and 63 τ_2), whereas 20 neurons had no memory. A total of 362 timescales were obtained from DLPFC with 124 single and 119 double exponential functions (124 τ, 119 τ_1 and 119 τ_2) and 79 DLPFC neurons had no memory. LIP neurons contributed 246 timescales from 74 single and 86 double exponential functions (74 τ, 86 τ_1 and 86 τ_2) and 45 LIP neurons showed no memory.

Comparison with behavior

Are the neural memory timescales relevant for learning and decision making? The matching pennies task that we used does not necessarily require the memory of past rewards and the optimal strategy for the monkey is to choose randomly and unpredictably. Although the overall performance of monkeys was nearly optimal, their trial-by-trial decisions, locally in time, were influenced by previous rewards and actions[11,19–22]. We analyzed the behavior of monkeys in different experimental sessions by fitting their decisions with a standard reinforcement learning model[5] (reinforcement learning, Online Methods). The learning rate parameter (α) of the reinforcement learning model quantifies the behavioral timescale of the memory trace ($\alpha \sim 1/\tau$). The resulting likelihood was significantly larger than the likelihood for reshuffled trials and the model fit with behavioral data was significant in 78% of the sessions (196 of 250, $P < 0.05$). We found that the timescales of behavioral memory varied across sessions, possibly suggesting that monkeys adopted different strategies in successive sessions. For the 196 sessions fitted by the reinforcement learning model, the distribution of behavioral timescales followed a power law distribution (**Fig. 5a**) and the exponent was consistent with that measured in the neural distribution. Hence, the distributions of behavioral and neuronal timescales qualitatively matched with each other. This result suggests that there might be a relationship between the memory trace observed at the neural level and that observed at the behavioral level. We tested this hypothesis by comparing the neural timescale for reward memory observed during a given recording session with the behavioral timescale fit in that session (when both are available) and we found a small, but significant, correlation across sessions ($R = 0.12$, $P = 0.003$; **Fig. 5b**), suggesting that the activity of single neurons is related, albeit weakly, to the behavioral strategy of the animals.

Do the reward memory timescales also change in a single session? We determined whether the timescales are stable in a single recording session by dividing each session into two separate blocks (halves) of trials and we re-estimated both the neural and behavioral timescales separately in the two blocks. Both the behavioral and neural memory timescales were fairly stable in a single session (**Fig. 6**).

The neural and behavioral timescales might fluctuate together across sessions, but their small correlation indicates that there is only a weak coupling. Indeed, we found that at any moment, the timescales of reward memory varied across cortical neurons. In each recording session, only few neurons were simultaneously recorded (about two on average). When we estimated memory timescales for pairs of simultaneously recorded neurons, the correlation between their time constants was not significantly different from zero (312 pairs of timescales, $R = 0.07$, $P = 0.2$). This result suggests that the broad distribution of memory time constants observed in the data reflects a variability of timescales across different neurons, rather than resulting from a memory timescale fixed for all neurons that collectively changes across sessions.

Taken together, our results support the conclusion that a diverse collection of neural memory timescales, a reservoir, is available across cortical neurons at any given time. The animal's behavior may be determined by a readout system that is able to sample, at different times,

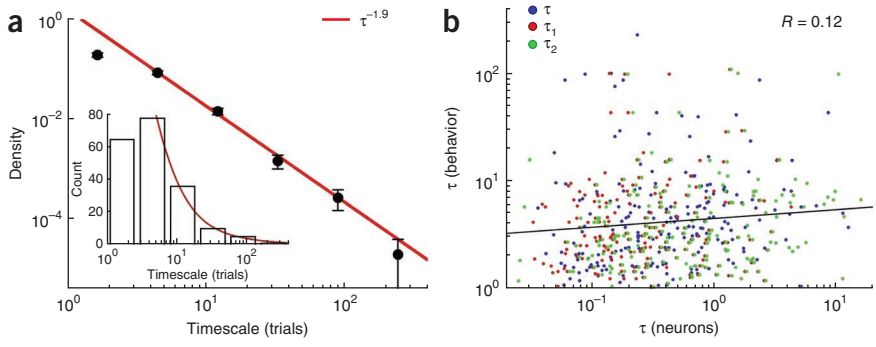

Figure 5 Distribution of behavioral timescales and their relationship with the neural memory timescales. (**a**) Time constant τ estimated from the learning rate α ($\tau \sim 1/\alpha$) of a reinforcement learning model fit to the monkey's behavioral data. Black disks show the density in the corresponding bin; that is, the count of timescales divided by the bin length (error bars represent ±s.e.). The inset shows the count of the timescales in the same bins, in linear scale (a total of 196 timescales). The red line (red curve in the inset) shows a power law fit (exponent = −1.9). (**b**) The scatterplot of behavioral versus neural memory timescales obtained from all sessions where both were available. Neural timescales from different types of fit (τ from single exponential and τ_1, τ_2 from double exponential) are shown in different colors. Behavioral and neural timescales show a small, but significant, correlation ($R = 0.12$, $P = 0.003$).

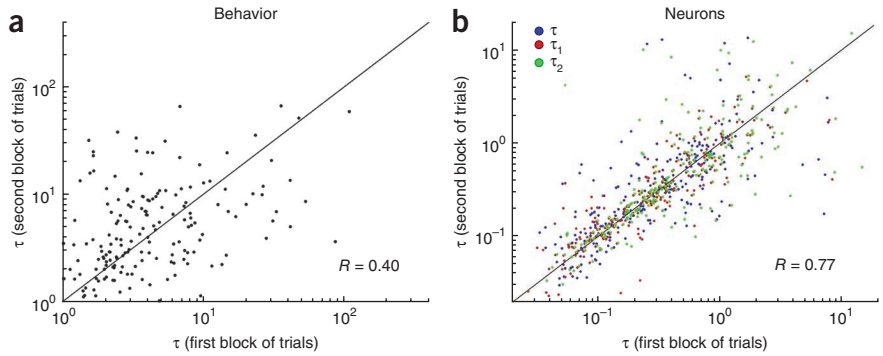

Figure 6 Stability of behavioral and neural memory timescales in an experimental session. (**a**,**b**) In both panels, the scatterplot of the timescales fitted in the second half of the trials is plotted against the timescales fitted in the first half of the trials in the same session. The correlation was significantly different from zero in both cases ($R = 0.4$ for behavioral timescales, $R = 0.77$ for neural timescales), suggesting that both types of timescales are fairly stable in a single session. Neural memory timescales from different types of fit (τ from single exponential and τ_1, τ_2 from double exponential) are shown in different colors.

from a variety of timescales present in the reservoir. The reservoir might not be static and it may change its distribution of timescales from day to day. During competitive games, the subjects might also take into account their recent choices to determine their future behavior. We therefore tested whether any memory trace of choice exists in the recorded neurons by applying the same analysis of equations (1) and (2) and substituting reward with choice. We found that multiplicative modulation and a power law distribution of memory timescales also hold for memory trace of past choices (**Supplementary Fig. 2**).

Neural network model for memory traces

What neural mechanism(s) accounts for the statistical properties of reward memory described above? To address this question, we constructed a simple neural network model that reproduces the observed neural memory traces (**Fig. 7** and **Supplementary Fig. 3**). Model neurons integrate the reward signals by receiving a current impulse whenever a reward is obtained. Because neurons are recurrently connected and form loops, their activities reverberate and are able to maintain the memory of reward events. However, those memories decay and are slowly forgotten according to a time course that depends on the pattern of synaptic connections among neuron pairs. Specifically, the activity of neurons evolve according to $\frac{d\mathbf{v}}{dt} = J \cdot v(t) + h \cdot \mathrm{Rew}(t)$, where **v** is a vector of M components, each component is the activity of a different neuron in the reservoir (M = 1,000 neurons in simulations), J is the synaptic connectivity matrix of their interactions and **h** is a vector representing the relative strength of the reward input Rew(t) to each neuron. For our purposes, the specific form of the input signals is not important; the results depend only on the synaptic matrix J. We assumed that the connection weights (the entries of the matrix J) were randomly distributed and we looked for candidate probability distributions such that the network model reproduces the distributions of timescales and amplitudes observed in the neural data from behaving monkeys (see **Supplementary Text**). Amplitudes determine the extent of the immediate response of neurons to reward, with respect to the average activity. Time constants had a power-law distribution (**Fig. 4**) and the distribution of amplitudes was exponential (**Fig. 8a**, where we used A for one exponential and $A_1 + A_2$ for two exponentials).

First, we found that the connection weights must be broadly distributed among neuron pairs and that this endows the network with a wide variety of timescales. Intuitively, the stronger the connection, the longer the reverberation of the input and hence the timescale of the memory trace. However, if connections are also heterogeneous, then weaker connections and smaller timescales will also contribute to the memory traces. If the width of the distribution of connection weights reaches a certain threshold, a power-law distribution of timescales is observed (**Fig. 7b**), which is characterized by a high probability for both small and large timescales. This is a distinct type of network state 'at a critical point' (or edge of chaos in nonlinear systems), which have been proposed to be desirable for many kinds of computations[16–18]. In our model, the criticality corresponds to the situation where the system is on the verge of losing stability. When the width of the connection distribution exceeds the critical level, the linear system is unstable and the model would need to be extended to include nonlinearities such as saturation of neural activity. For the sake of simplicity, we limited ourselves to the linear model, which is sufficient for the purpose of reproducing the observed power-law distribution of timescales under specific conditions.

A second desirable property of the network is that its dynamics are robust with respect to small changes of the connection strengths. If the coding of the memory changes markedly as a result of small changes in the connection strengths (for example, synaptic noise), it would be difficult for a downstream system to interpret that code. A known property of the connection matrix J that ensures that kind of robustness is normality, which guarantees that there is an orthogonal set of eigenvectors[26] (but see refs. 27–29 for non-normal neural network models). If J is normal, we found that the amplitudes of the memory traces followed an exponential distribution (**Fig. 8b**), consistent with the experimental observations (**Fig. 8a**). To the best of our knowledge, our results provide the first complete statistical description of

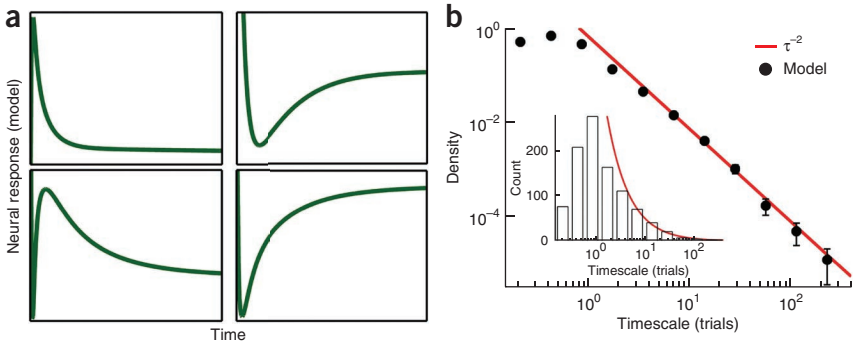

Figure 7 Neural responses (memory traces) in the model and distribution of timescales of the memory traces in model neurons. (**a**) The memory traces of four model neurons. (**b**) The black disks show the density of timescales in the corresponding bin; that is, the count of timescales divided by the bin length (error bars represent ±s.e.). The inset shows the count of the timescales in the same bins, in linear scales (a total of 1,000 timescales). The red line (red curve in the inset) shows a power law fit (exponent = –2).

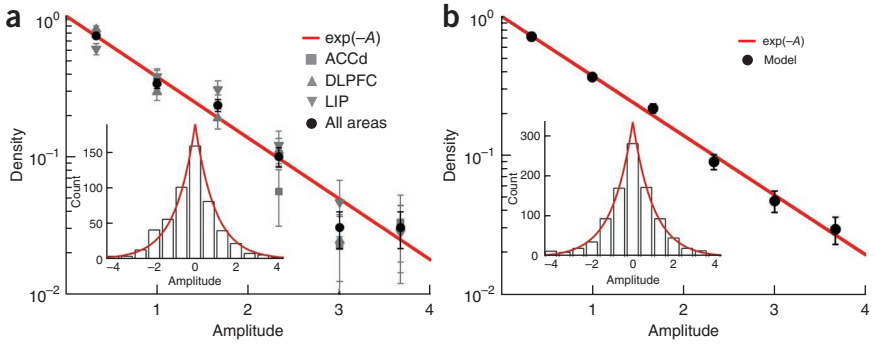

Figure 8 Distribution of amplitudes of the memory traces in the neural data and model. (**a**,**b**) In both panels, black disks show the density in the corresponding bin; that is, the count of timescales divided by the bin length (error bars represent ±s.e.). The inset shows the count of the amplitudes in the same bins, in a linear scale (537 amplitudes in the data, 1,000 in the model). Amplitudes are plotted as absolute values, as the distribution was approximately symmetric (symmetry is shown in the inset). Grey markers show the density separately for the three different recorded areas (squares, ACCd, 134 amplitudes; upward triangles, DLPFC, 243; downward triangles, LIP, 160). The red line (red curve in the inset) shows an exponential fit ($e^{-|A|}$).

a network connection matrix based on *in vivo* neuronal recordings of behaving animals (see also refs. 30–32).

DISCUSSION

The power law of timescales suggests that the duration of reward memory trace is highly diverse across cortical neurons. The same diversity is observed across three cortical areas, suggesting that the computation of reward memory is a distributed process. This finding is consistent with an increasing appreciation that neural encoding of cognitive variables is highly heterogeneous and distributed[33,34]. Prefrontal cortex is important for dynamic decision processes encoding and updating values[1–4]. Although anterior cingulate cortex has been implicated in monitoring conflict between incompatible response processes[35] or detecting performance errors[36], recent studies have placed more emphasis on its role in representing both positive and negative values[19,37]. Parietal cortex has also been implicated in decision making on the basis of the value representation and the accumulation of sensory evidence[38,39].

Our work provides a comprehensive description of memory traces in terms of a specific distribution of timescales across a population of neurons and introduces a framework that could potentially be applicable to different brain areas and different types of memory. The concept of multiplicative modulation of memory traces can be used to deduce the neural memory timescales in various tasks and to test the idea that a different set of time constants is selected to adapt to a specific environment[6,7]. Although the global optimal strategy for the matching pennies task is to choose randomly and therefore does not require memory, the animals made their decisions largely on the basis of their reward history[11,19–22]. Perhaps in the persistent search for an appropriate strategy, they sampled different timescales across experimental sessions. We found that those behavioral timescales followed a similar distribution and were weakly, but significantly, correlated with the timescales observed at the neural level. This suggests the possibility that the behavior might be driven by a mechanism that appropriately samples from a range of timescales in a neural network, which has yet to be elucidated. Alternatively, this weak correlation might be caused by factors that are currently not understood. Note that the observed range is different for the neuronal versus behavioral time constants. Also, we have not attempted to fit the behavioral data by a reinforcement learning model endowed with multiple time constants. Future work is needed to further assess the correlation between neural memory traces and behavior. Regardless, our results suggest that reward memory with multiple time constants might be used to compute the value functions in reinforcement learning theory in more than one timescale. Similarly, the double exponential decay of memory may correspond to a reward prediction error signal; if the short timescale (τ_1) is small enough (about one trial or smaller), then the corresponding exponential filter will respond primarily to the reward in the present trial, whereas the long timescale (τ_2) may provide a value signal by weighting the rewards in the past few trials. When the two exponentials have opposite signs, they roughly subtract the value from the actual reward signal, therefore providing a reward prediction error. It has been noted that a biphasic filtering in dopamine neurons might provide a reward prediction error[40].

Besides the memory for reward, the activity of primate cortical neurons reflects other types of short-term memory. The time course of memory-related activity varies across different neurons and different task protocols, including persistent, ramping and multi-phasic activity[41–43]. Memory traces in the neural signals are mixed with other task-dependent factors[44,45] and it has been debated as to whether other processes involved in goal-directed behavior could be inter-mixed with a memory trace, such as spatial attention[46], motor planning[47], anticipation of future events[48] or timing[49]. The epoch code in the present task might include many of those processes and we found that memory signals could be dissociated from those factors by assuming a multiplicative computation. The hypothesis of a multiplicative effect of memory on neural activity could be tested by looking more closely at the multi-phasic time course of memory-related activity observed in other experiments. The computational advantage of the multiplicative effect of memory needs to be further investigated. For example, it may serve the appropriate recall of memories at different epochs (see **Supplementary Text**), as observed in a recent study[50].

Reservoir-type networks have been the subject of active research in computational neuroscience and machine learning[13–15], but experimental support that such networks are adopted by the brain has been lacking. Those models predict that the memory of input signals is stored in a large, recurrent and heterogeneous network (reservoir) in a distributed manner and that a desired output is obtained by a trainable combination of the response signals in the reservoir. The heterogeneous encoding of the input allows the flexible learning of different output functions. In our context, that may correspond to a flexible change in strategy resulting from the variety of timescales for reward memory present in the reservoir. We present direct experimental evidence, at the level of single neurons, for a high-dimensional reservoir network of reward memory traces in prefrontal, cingulate and parietal areas of the primate cortex. This empirical finding is reproduced by a simple computational model, which suggests that reward filtering in the cortex involves a dynamic reservoir network operating at the critical point, leading to a power-law distribution of time constants. The output of the network, supposedly driving the animal's behavior, is not explicitly modeled in our equations. Further studies are necessary to elucidate how the motor areas read out the memory of reward and choices and how the two are combined to subserve adaptive choice behavior.

Power-law distributions are unusual, as they imply a high probability for both large and small time constants. A diversity of time constants also means a broad range of learning rates, as the two are inversely

related to each other. This is noteworthy, as a shift from an exploitive to an exploratory strategy as the environment becomes uncertain is often assessed by an increase in the learning rate[10]. Our work suggests that a broad range of learning rates are available in the system, a subset of which (fast or slow) might be selectively utilized according to which strategy is behaviorally desirable. Ultimately, this framework could lead to a new model for predicting how reward expectation is computed and how reward memory affects decision making.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/natureneuroscience/.

Note: Supplementary information is available on the Nature Neuroscience website.

ACKNOWLEDGMENTS

We thank J. Mazer and M.W. Jung for comments on an earlier version of the manuscript, and R. Chaudhuri, M. Harre and J. Murray for discussions. This work was supported by the US National Institutes of Health grant R01 MH062349 and the Swartz Foundation (A.B. and X.-J.W.), and by US National Institutes of Health grants R01 MH073246 (X.-J.W. and D.L.) and DA029330 (D.L.).

AUTHOR CONTRIBUTIONS

All of the authors participated in the research design and the preparation of the manuscript. H.S. collected the data, A.B. and H.S. analyzed data, and A.B. and X.-J.W. performed modeling.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/natureneuroscience/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Kable, J.W. & Glimcher, P.W. The neurobiology of decision: consensus and controversy. *Neuron* **63**, 733–745 (2009).
2. Rushworth, M.F. & Behrens, T.E. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat. Neurosci.* **11**, 389–397 (2008).
3. Wang, X.-J. Decision making in recurrent neural circuits. *Neuron* **60**, 215–234 (2008).
4. Soltani, A., Lee, D. & Wang, X.-J. Neural mechanism for stochastic behavior during a competitive game. *Neural Netw.* **19**, 1075–1090 (2006).
5. Sutton, R.S. & Barto,, A.G. *Reinforcement Learning, an Introduction* (MIT Press, Cambridge, Massachusetts, 1998).
6. Behrens, T.E., Woolrich, M.W., Walton, M.E. & Rushworth, M.F. Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).
7. Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B. & Dolan, R.J. Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879 (2006).
8. Lau, B. & Glimcher, P.W. Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* **84**, 555–579 (2005).
9. Corrado, G.S., Sugrue, L.P., Seung, H.S. & Newsome, W.T. Linear-nonlinear-Poisson models of primate choice dynamics. *J. Exp. Anal. Behav.* **84**, 581–617 (2005).
10. Kennerley, S.W., Walton, M.E., Behrens, T.E., Buckley, M.J. & Rushworth, M.F. Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* **9**, 940–947 (2006).
11. Lee, D., Conroy, M.L., McGreevy, B.P. & Barraclough, D.J. Reinforcement learning and decision making in monkeys during a competitive game. *Brain Res. Cogn. Brain Res.* **22**, 45–58 (2004).
12. Kim, S., Hwang, J., Seo, H. & Lee, D. Valuation of uncertain and delayed rewards in primate prefrontal cortex. *Neural Netw.* **22**, 294–304 (2009).
13. Maass, W., Natschläger, T. & Markram, H. Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Comput.* **14**, 2531–2560 (2002).
14. Jaeger, H., Lukosevicius, M., Popovici, D. & Siewert, U. Optimization and applications of echo state networks with leaky-integrator neurons. *Neural Netw.* **20**, 335–352 (2007).
15. Verstraeten, D., Schrauwen, B., D'Haene, M. & Stroobandt, D. An experimental unification of reservoir computing methods. *Neural Netw.* **20**, 391–403 (2007).
16. Sussillo, D. & Abbott, L.F. Generating coherent patterns of activity from chaotic neural networks. *Neuron* **63**, 544–557 (2009).
17. Bertschinger, N. & Natschlager, T. Real-time computation at the edge of chaos in recurrent neural networks. *Neural Comput.* **16**, 1413–1436 (2004).
18. Langton, C.G. Computation at the edge of chaos: phase transitions and emergent computations. *Physica D* **42**, 12–37 (1990).
19. Seo, H. & Lee, D. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* **27**, 8366–8377 (2007).
20. Seo, H., Barraclough, D.J. & Lee, D. Dynamic signals related to choices and outcomes in the dorsolateral prefrontal cortex. *Cereb. Cortex* **17**, i110–i117 (2007).
21. Seo, H., Barraclough, D.J. & Lee, D. Lateral intraparietal cortex and reinforcement learning during a mixed-strategy game. *J. Neurosci.* **29**, 7278–7289 (2009).
22. Barraclough, D.J., Conroy, M.L. & Lee, D. Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* **7**, 404–410 (2004).
23. Lapish, C.C., Durstewitz, D., Chandler, L.J. & Seamans, J.K. Successful choice behavior is associated with distinct and coherent network states in anterior cingulate cortex. *Proc. Natl. Acad. Sci. USA* **105**, 11963–11968 (2008).
24. Sigala, N., Kusonoki, M., Nimmo-Smith, I., Gaffan, D. & Duncan, J. Hierarchical coding for sequential task events in the monkey prefrontal cortex. *Proc. Natl. Acad. Sci. USA* **105**, 11969–11974 (2008).
25. Jin, D.Z., Fujii, N. & Graybiel, A.N. Neural representation of time in cortico-basal ganglia circuits. *Proc. Natl. Acad. Sci. USA* **106**, 19156–19161 (2009).
26. Trefethen, L.N. & Embree, M. *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators* (Princeton University Press, Princeton, New Jersey, 2005).
27. Murphy, B.K. & Miller, K.D. Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron* **61**, 635–648 (2009).
28. Ganguli, S., Huh, D. & Sompolinsky, H. Memory traces in dynamical systems. *Proc. Natl. Acad. Sci. USA* **105**, 18970–18975 (2008).
29. Goldman, M.S. Memory without feedback in a neural network. *Neuron* **61**, 621–634 (2009).
30. Schneidman, E., Berry, M.J., Segev, R. & Bialek, W. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* **440**, 1007–1012 (2006).
31. Brunel, N., Hakim, V., Isope, P., Nadal, J.-P. & Barbour, B. Optimal information storage and the distribution of synaptic weights: perceptron versus purkinje cell. *Neuron* **43**, 745–757 (2004).
32. Ganguli, S. *et al.* One-dimensional dynamics of attention and decision making in LIP. *Neuron* **58**, 15–25 (2008).
33. Duncan, J. An adaptive coding model of neural function in prefrontal cortex. *Nat. Rev. Neurosci.* **2**, 820–829 (2001).
34. Rigotti, M., Rubin, D.B.D., Wang, X.-J. & Fusi, S. Internal representation of task rules by recurrent dynamics: the importance of the diversity of neural responses. *Front. Comput. Neurosci.* **4**, 24 (2010).
35. Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S. & Cohen, J.D. Conflict monitoring and cognitive control. *Psychol. Rev.* **108**, 624–652 (2001).
36. Holroyd, C.B. & Coles, M.G.H. The neural basis of human error processing: reinforcement learning, dopamine and error-related negativity. *Psychol. Rev.* **109**, 679–709 (2002).
37. Wallis, J.D. & Kennerley, S.W. Heterogeneous reward signals in prefrontal cortex. *Curr. Opin. Neurobiol.* **20**, 191–198 (2010).
38. Platt, M.L. & Glimcher, P.W. Neural correlates of decision variables in parietal cortex. *Nature* **400**, 233–238 (1999).
39. Roitman, J.D. & Shadlen, M.N. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* **22**, 9475–9489 (2002).
40. Bayer, H.M. & Glimcher, P.W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129–141 (2005).
41. Rainer, G. & Miller, E.K. Time course of object-related neural activity in the primate prefrontal cortex during a short-term memory task. *Eur. J. Neurosci.* **15**, 1244–1254 (2002).
42. Machens, C.K., Romo, R. & Brody, C.D. Functional, but not anatomical, separation of "what" and "when" in prefrontal cortex. *J. Neurosci.* **30**, 350–360 (2010).
43. Shafi, M. *et al.* Variability in neuronal activity in primate cortex during working memory tasks. *Neuroscience* **146**, 1082–1108 (2007).
44. Curtis, C.E. & Lee, D. Beyond working memory: the role of persistent activity in decision making. *Trends Cogn. Sci.* **14**, 216–222 (2010).
45. Passingham, D. & Sakai, K. The prefrontal cortex and working memory: physiology and brain imaging. *Curr. Opin. Neurobiol.* **14**, 163–168 (2004).
46. Lebedev, M.A., Messinger, A., Kralik, J.D. & Wise, S.P. Representation of attended versus remembered locations in prefrontal cortex. *PLoS Biol.* **2**, e365 (2004).
47. Funahashi, S., Chafee, M.V. & Goldman-Rakic, P.S. Prefrontal neuronal activity in rhesus monkeys performing a delayed anti-saccade task. *Nature* **365**, 753–756 (1993).
48. Rainer, G., Rao, S.G. & Miller, E.K. Prospective coding for objects in primate prefrontal cortex. *J. Neurosci.* **19**, 5493–5505 (1999).
49. Brody, C.D., Hernandez, A., Zainos, A. & Romo, R. Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex. *Cereb. Cortex* **13**, 1196–1207 (2003).
50. Bromberg-Martin, E.S., Matsumoto, M., Nakahara, H. & Hikosaka, O. Multiple timescales of memory in lateral habenula and dopamine neurons. *Neuron* **67**, 499–510 (2010).

ONLINE METHODS

Animal preparation and electrophysiological recording. All of the data were collected using the same behavioral task and electrophysiological techniques. These techniques have been described previously[19–21]. We used six rhesus monkeys (five male and one female). The animal's head was fixed during the experiment and eye movements were monitored at a sampling rate of 225 Hz with a high-speed eye tracker (Thomas Recording). Animals performed an oculomotor free-choice task[22] (matching pennies; **Fig. 1a**). Trials began with the animal fixating a small yellow square (0.9° × 0.9°) displayed at the center of the computer screen for a 0.5-s fore-period. Two identical green disks were presented at 5° eccentricity in diametrically opposed locations along the horizontal meridian for a 0.5-s delay period. The extinction of the central target signaled the animal to shift its gaze toward one of the targets within 1 s. After the monkey maintained its fixation on the chosen peripheral target for 0.5 s, a red ring appeared around the target selected by the computer. The animal was rewarded only if it chose the same target as the computer, which simulated a rational decision maker in the matching pennies game trying to minimize the animal's expected payoff. Before each trial, the computer made a prediction for the animal's choice by computing the conditional probabilities for the animal to choose each target given its choices and rewards in the preceding four trials. The computer made a random choice if the probabilities were consistent with unbiased behaviors, otherwise it would bias its selection against the prediction. Single-unit activity was recorded using a five-channel multi-electrode recording system (Thomas Recording) from three cortical regions: the ACCd[19] (area 24c, two male monkeys, 8–12 kg), DLPFC[20,22] (anterior to the frontal eye field; four male and one female monkeys, 5–12 kg) and LIP[21] (two male and one female monkeys, 5–11 kg). All the neurons were recorded without pre-screening. The placement of the recording chamber was guided by magnetic resonance images and confirmed by metal pins inserted in known anatomical locations at the end of the experiment in some animals. In three animals, two recording chambers were used for simultaneous recording of DLPFC and LIP. All the experimental procedures were approved by the Institutional Animal Care and Use Committee at Yale University and conformed to the Public Health Services Policy on Humane Care and Use of Laboratory Animals and the Guide for the Care and Use of Laboratory Animals.

Multiple regression analysis of memory traces. To estimate the memory traces $f(n,k)$ from the observed neuronal firing rates and sequence of rewards, we computed the firing rates in each trial in 12 time intervals of 250 ms each (**Fig. 1a**). The following model was used to fit the firing rates. The firing rate of a neuron depends on the trial epoch k, following the epoch code $g(k)$; after the outcome is revealed (feedback period) in each trial, the firing rate is changed by an amount of $+f(n',k)$ for reward and $-f(n',k)$ for no reward, where n' is the number of trials elapsed since that outcome. The effects of outcomes in successive trials are additive. The firing rate $FR(n,k)$ is thus described by

$$FR(n,k) = g(k) + \sum_{n'=0:5} f(n',k)\mathrm{Rew}(n-n') + \mathrm{noise} \quad (3)$$

where the index k labels the epoch ($k = 1,\ldots,12$) and the indices n and n' label trials. The effect of reward extends up to five trials ($n' = 0,\ldots,5$), while the index n runs over all N trials available in each neuron recording (starting after the first five trials, $n = 6,\ldots, N$). To determine $f(n,k)$ and $g(k)$, we applied a multiple regression model by using the known $FR(n,k)$ and $\mathrm{Rew}(n)$ (+1/−1 for reward/no reward). Note that the epoch code $g(k)$ depends on the twelve different epochs within a trial, whereas the reward $\mathrm{Rew}(n)$ depends only on trial number. As a consequence, the regression can be applied separately for each epoch. For a fixed epoch k, the seven unknown variables $g(k)$, $f(0,k)$, $f(1,k)$, $f(2,k)$, $f(3,k)$, $f(4,k)$ and $f(5,k)$ can be determined by using the known values of $FR(n,k)$ and $\mathrm{Rew}(n)$ in $N-5$ trials ($n = 6,\ldots, N$). Using a parsimonious matrix notation and omitting the epoch label k, equation (3) can be rewritten as

$$\mathrm{FR} = \mathrm{Rew} \bullet f + \mathrm{noise} \quad (4)$$

where the vector of the known firing rates FR is equal to

$$\mathrm{FR} = [FR(6,k), FR(7,k), \ldots, FR(N,k)]^T \quad (5)$$

The seven unknown variables have been rewritten by a single vector $\mathbf{f}$

$$\mathbf{f} = [g(k), f(0,k), f(1,k), f(2,k), f(3,k), f(4,k), f(5,k)]^T \quad (6)$$

The matrix Rew is known, given by

$$\mathrm{Rew} = \begin{pmatrix} 1 & \mathrm{Rew}_6 & \mathrm{Rew}_5 & \mathrm{Rew}_4 & \mathrm{Rew}_3 & \mathrm{Rew}_2 & \mathrm{Rew}_1 \\ 1 & \mathrm{Rew}_7 & \mathrm{Rew}_6 & \mathrm{Rew}_5 & \mathrm{Rew}_4 & \mathrm{Rew}_3 & \mathrm{Rew}_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & \mathrm{Rew}_N & \mathrm{Rew}_{N-1} & \mathrm{Rew}_{N-2} & \mathrm{Rew}_{N-3} & \mathrm{Rew}_{N-4} & \mathrm{Rew}_{N-5} \end{pmatrix} \quad (7)$$

Because the sequence of rewards is nearly random and N is large, different columns of the matrix Rew are nearly orthogonal. This implies that the matrix product $(\mathrm{Rew}^T \cdot \mathrm{Rew})$ is well conditioned and that the solution f_{sol} minimizing the variance of the noise (or squared error) is robust and given by

$$f_{sol} = (\mathrm{Rew}^T \bullet \mathrm{Rew})^{-1} \bullet \mathrm{Rew}^T \bullet \mathrm{FR} \quad (8)$$

This expression is used to obtain the results. The confidence intervals for f_{sol} are derived from the residual errors according to the Matlab (Mathworks) function regress.

The matrix product $(\mathrm{Rew}^T \cdot \mathrm{Rew})$ is approximately proportional to the identity matrix. When $\mathrm{Rew}^T \cdot \mathrm{Rew} = I$, the filter is equal to the firing rate averaged over all trials, where the average is conditioned on the past rewards. This is equivalent to the cross-correlation between the input (rewards) and output (firing rates) and its application would correspond to a reverse correlation method, commonly used in the analysis of sensory neural coding. Here, however, we only showed results from the multiple regression analysis. For simplicity, we used an average over all trials as the definition of epoch code $g(k)$ in the main text, making use of the above approximation.

Exponential memory traces and model selection. The model considered here is similar to that of equation (3), but we assumed that memory traces are exponential function $\mathrm{ex}(t)$ rescaled by the epoch code $g(k)$.

$$FR(n,k) = g(k) + g(k)\sum_{n'=0.5} \mathrm{ex}(t)\mathrm{Rew}(n-n') + \mathrm{noise} \quad (9)$$

The filter f considered in equation (3) is replaced by $g(k) \cdot \mathrm{ex}(t)$. We considered two different exponential functions, a single exponential and the sum of two exponentials.

$$\mathrm{ex}_1(t) = Ae^{\frac{-t}{\tau}} \quad (10)$$

$$\mathrm{ex}_2(t) = A_1 e^{\frac{-t}{\tau_1}} + A_2 e^{\frac{-t}{\tau_2}} \quad (11)$$

where $\tau_1 < \tau_2$. The physical time t depends on all indices k, n and n' because the time elapsed between different epochs and between successive trials is variable, due to the variability in the time taken by the animal to start a trial and to make a saccade to one of the two targets. On the basis of the time stamps generated during the experiment, we computed the physical time $t = t(n,k,n')$ as the difference between the time corresponding to a given trial and epoch (n,k) and the time corresponding to the feedback epoch of n' trials in the past (up to five trials). Note that the memory trace f obtained by the multilinear regression is not computed in physical time. In that case, we assumed that the saccade reaction time of the animal in all trials is equal to 120 ms (average) and that the time elapsed between the initiation of two successive trials is 3.4 s (median).

The epoch code $g(k)$ was fixed by the firing rates averaged across trials, whereas the parameters of the exponential function (two parameters (A,τ) when using equation (10) and four parameters (A_1,τ_1,A_2,τ_2) when using equation (11)) were estimated using a nonlinear curve-fitting procedure, implemented by the Matlab function fminsearch, minimizing the variance of the noise (sum of squared errors) in equation (9). Fitting was repeated ten times for each neuron and each model in the search for a global minimum of the error. Any parameters resulting in unrealistic values were discarded, such as negative values of τ, τ_1 or τ_2, values of

τ larger than 20 trials, and the absolute value of A or $(A_1 + A_2)$ larger than 4. We determined the parameters for all neurons in both exponential models, single and double exponential and denoted the corresponding square errors by σ_1^2 and σ_2^2, respectively. We also computed the variance of firing rate, σ_0^2, as the square error for a zero filter model, that is, ex = 0 or FR = g + noise. Among the three models, the selection of the appropriate one for each neuron was determined according to the Bayesian information criterion (BIC)

$$\mathrm{BIC}_i = m\log(\sigma_i^2) + p_i \log(m) \tag{12}$$

where p_i denotes the number of parameters in the model, and $p_0 = 1$, $p_1 = 3$, $p_2 = 5$, for 0, 1 and 2 exponential fit, respectively (note that the variance σ_i^2 is also a parameter), and m is the number of data points ($m = 12(N - 5)$; 12 epochs and $N - 5$ trials for each neuron). The model with the minimum BIC was chosen for each neuron. As a control of the fitting procedure, we reshuffled the label n in the firing rates FR(n,k), assigning to each firing rate the value of a random trial, and we repeated the entire procedure.

Reinforcement learning fit of behavior. We applied a standard reinforcement learning model[5], separately for each recording session, to analyze how the animal's choice was influenced by the outcomes of its previous choices. For example, when right target R was chosen in trial t, the value function for R, denoted by $Q_R(t)$, was updated according to

$$Q_R(t+1) = Q_R(t) + \alpha[\mathrm{Rew}(t) - Q_R(t)] \tag{13}$$

where Rew(t) denotes the reward received by the animal in trial t, and the term inside square is commonly defined as the reward prediction error; that is, the discrepancy between the actual reward and the expected reward. A similar equation holds for the left value function $Q_L(t)$. The probability that the animal would choose the rightward target in trial t, $P_R(t)$, was determined by the SoftMax transformation

$$P_R(t) = \frac{\exp(\beta Q_R(t))}{\exp(\beta Q_L(t)) + \exp(\beta Q_R(t))} \tag{14}$$

where β, referred to as the inverse temperature, determines the randomness of the animal's choices. Model parameters (α,β) were estimated separately for each recording session by using a maximum likelihood procedure, where the likelihood is the product of probabilities in all trials (equation (14)), in each trial using R or L according to the actual monkey's choice. The parameter values maximizing the likelihood were found by using the Matlab function fminsearch. The significance of the estimation was assessed, for each session, by constructing 100 surrogate sessions, each one obtained by reshuffling of the order of trials. The distribution of 100 maximum likelihoods obtained by the estimation procedure was then compared with the maximum likelihood of the non-reshuffled case, which was considered to be significant if not smaller than the five largest reshuffled likelihoods.

Value functions and reward prediction error signals can be related to the exponential filters estimated for individual neurons. If a single value function (for a given stimulus/action) and a single reward (delivered at time zero) are considered, the solution of equation (13) can be approximated by an exponential response, $Q(t) = (1 - \frac{1}{\tau})^t \sim \exp(\frac{-t}{\tau})$, provided that τ is larger than one trial. When a sequence of rewards is delivered instead of a single one, the value is a superposition of the exponential responses for each reward.

doi:10.1038/nn.2752