Journal of Neuroscience Methods 307 (2018) 175–187

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

A separable two-dimensional random field model of binary response data from multi-day behavioral experiments

Noa Malem-Shinitski[a,b], Yingzhuo Zhang[c,*], Daniel T. Gray[d,e], Sara N. Burke[f], Anne C. Smith[d], Carol A. Barnes[d,e,g], Demba Ba[c]

[a] Bernstein Center for Computational Neuroscience, Berlin, Germany
[b] Department of Artificial Intelligence, Technische Universität Berlin, Berlin, Germany
[c] John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA
[d] Evelyn F. McKnight Brain Institute, University of Arizona, Tucson, AZ, USA
[e] Division of Neural Systems, Memory & Aging, University of Arizona, Tucson, AZ, USA
[f] Evelyn F. McKnight Brain Institute, Department of Neuroscience, University of Florida, Gainesville, FL, USA
[g] Department of Psychology, Neurology and Neuroscience, University of Arizona, Tucson, AZ, USA

HIGHLIGHTS

- Behavioral experiments enable the quantification of changes in the brain due to aging.
- We propose a method to analyze binary data from multi-object, multi-day, experiments.
- We apply it to data from young and aged monkeys in an object-reward association task.
- Unlike existing methods, it captures differences in group and individual performance.
- The method can discriminate the age groups with high accuracy solely using behavior.

ARTICLE INFO

Article history:
Received 24 October 2017
Received in revised form 13 March 2018
Accepted 12 April 2018
Available online 19 April 2018

Keywords:
Behavior
Learning
Aging
Separable random field model
Change-point test
Reversal learning task
EM algorithm
Markov Chain Monte Carlo
Laplace prior
Bayesian inference
Gibbs sampling
MAP estimation

ABSTRACT

Background: The study of learning in populations of subjects can provide insights into the changes that occur in the brain with aging, drug intervention, and psychiatric disease.
New method: We introduce a separable two-dimensional (2D) random field (RF) model for analyzing binary response data acquired during the learning of object-reward associations across multiple days. The method can quantify the variability of performance within a day and across days, and can capture abrupt changes in learning.
Results: We apply the method to data from young and aged macaque monkeys performing a reversal-learning task. The method provides an estimate of performance within a day for each age group, and a learning rate across days for each monkey. We find that, as a group, the older monkeys require more trials to learn the object discriminations than do the young monkeys, and that the cognitive flexibility of the younger group is higher. We also use the model estimates of performance as features for clustering the monkeys into two groups. The clustering results in two groups that, for the most part, coincide with those formed by the age groups. Simulation studies suggest that clustering captures inter-individual differences in performance levels.
Comparison with existing method(s): In comparison with generalized linear models, this method is better able to capture the inherent two-dimensional nature of the data and find between group differences.
Conclusions: Applied to binary response data from groups of individuals performing multi-day behavioral experiments, the model discriminates between-group differences and identifies subgroups.
© 2018 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Understanding how learning differs between groups of subjects is important in many areas of neuroscience research. Obtaining an objective measure of performance is challenging because observa-

* Corresponding author.
E-mail address: yingzhuozhang@g.harvard.edu (Y. Zhang).

https://doi.org/10.1016/j.jneumeth.2018.04.006
0165-0270/© 2018 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

tions are often not continuous (e.g., binary) and learning is both dynamic and varies greatly across individuals.

A popular method to quantify performance changes in binary data is to use the change point test (Siegel and Castellan, 1988) or a moving average (Kim and Frank, 2009; Stefani et al., 2003; Wise and Murray, 1999). These methods are easy to implement but can be problematic in that they rely on a predefined null hypothesis (Siegel and Castellan, 1988) or arbitrary choice of window length (Kim and Frank, 2009; Stefani et al., 2003; Wise and Murray, 1999), respectively. State-space and hidden Markov approaches (Smith et al., 2004, 2005, 2007; Eldar et al., 2011) have been found to be a practical alternative since they make use of likelihood methods to fit the data, allowing accurate assessment of individual and group dynamics. These models are useful for analysis of behavioral data as they are able to track the dynamics of learning (and forgetting), and provide confidence bounds on estimates allowing for across trial inference.

We propose a two-dimensional model of binary response data from multi-day behavioral experiments termed the SMuRF (separably Markov random field) model, first introduced in Zhang et al. (2018). In the present study, we examine the binary response data from two age groups of animals learning multiple object-reward associations across days. Conventional one-dimensional methods for analyzing these data (Smith et al., 2004, 2005, 2007; Eldar et al., 2011) ignore their two-dimensional nature by aggregating the responses across object pairs or days, thus failing to capture subtle changes in learning dynamics. Two-dimensional state-space models have been applied to the analysis of neural spike rasters (Czanner et al., 2008; Rad and Paninski, 2010), with the goal of simultaneously capturing the within and cross trial variability. Their appeal, in the context of binary response data from behavioral experiments (e.g., repeated across multiple days), is the ability to simultaneously capture the within-day (across object pairs) and cross-day variability of learning. While classical two-dimensional state-space models (Czanner et al., 2008; Rad and Paninski, 2010) obviate the need for aggregation, they lack the interpretability of one-dimensional state-space models. In contrast, the SMuRF model (Zhang et al., 2018) is a two-dimensional model that inherits the interpretability of one-dimensional approaches. As we demonstrate formally in Zhang et al. (2018), the SMuRF model is *not* a two-dimensional state-space model in the classical sense. Applied to data from behavioral experiments, the SMuRF model expresses the probability of observing the binary outcomes as a function of two latent processes: a within-day learning process that reflects an animal's item-dependent performance level during a day, and a cross-day learning process that captures the dynamics of proficiency across days. The two processes play the same role as the latent processes in one-dimensional methods (Smith et al., 2004, 2005, 2007; Eldar et al., 2011), thus inheriting the interpretability of these methods. Unlike in one-dimensional methods, we combine the two processes, thus capturing the two-dimensional nature of the behavioral raster data as in Czanner et al. (2008) and Rad and Paninski (2010). In this context, the contribution of this paper is twofold. First, we substitute a Laplace random walk for the Gaussian random walk in Zhang et al. (2018), to allow for accurate and automatic modeling of sudden changes or jumps in behavioral data (e.g., due to reversal learning). We term the resulting model the Laplace SMuRF model, and that from Zhang et al. (2018), the Gaussian SMuRF model. The Laplace SMuRF model leads to algorithms for parameter estimation and inference (e.g., MAP inference) different from those derived in Zhang et al. (2018) for the Gaussian SMuRF model. Second, learning within a day is tracked over groups in addition to the individual learning across days. Using this formulation, we are able to obtain an accurate assessment of learning using a single model. We demonstrate that this model has superior predictive power, as measured by the WAIC (Vehtari et al., 2017), compared to one that aggregates the behavioral data within day, and is therefore better suited to perform inferences about the presence or absence of age differences.

We apply the Laplace SMuRF model to data acquired from young (mean 10.4 years, range 9.8–11.2) and aged (mean 23.2 years, range 20.2–27.7) female macaque monkeys performing a reversal learning task which was designed to identify alterations in behavior after reinforcement contingencies change (Gray et al., 2016). Using a linear mixed-effects model (LME), the analysis from Gray et al. (2016) indicated that older animals were impaired at affective shifting. In addition, by dividing each day into three epochs, it was observed that, for both age groups, within-day learning/performance was generally better in the middle of the session than at the beginning or the end, an effect which may be termed a "middle-list" effect. No between age difference was observed using this technique. Using the Laplace SMuRF model, we find that, as a group, the older monkeys find the object pairings harder than do the young monkeys, and that the ability to change behavior ("cognitive flexibility") of the younger group is higher. We are also able to use the results of our analysis to create clusters that, for the most part, coincide with those formed by the age groups. Simulation studies suggest that clustering based on the model's results captures inter-individual differences in performance levels, which allows us to identify exceptional behavior.

The remainder of our treatment begins in Section 2, where we describe the experimental behavioral paradigm that motivated the Laplace SMuRF model. We introduce the Laplace SMuRF model of learning, as well as algorithms for parameter estimation and inference, in Section 3. The interested reader can find the derivations relevant to Section 3 in the Appendix. In Section 4, we utilize the model from Section 3 to assess between-group differences in learning, and to predict group membership from behavior. We conclude in Section 5.

2. Experimental behavioral paradigm

We analyze data from 14 female macaque monkeys – 6 young (mean 10.4 years, range 9.8–11.2) and 8 old (mean 23.2 years, range 20.2–27.7) animals – performing a reversal learning task (Burke et al., 2014; Gray et al., 2016). This experiment was designed to identify differences in affective shifting between the two groups. Behavioral experiments followed guidelines established by the National Institutes of Health and were approved by the Institutional Animal Care and Use Committee at the University of Arizona.

In the first phase of the task, each animal was presented each day with the same 40 object pairs in the same order. By a process of trial and error the animal determined which object in the pair was associated with a food reward. The first phase of the task ended when a monkey reached a 90% performance criterion over 5 consecutive days (that is, 180/200 item pairs correct). After reaching criterion, the second phase began. In this part of the experiment, the reward for each object pair was switched to the previously unrewarded object and each animal was expected to learn these novel associations. The second phase ended when the same 90% criterion over 5 days was reached. Fig. 1 presents a schematic depiction of the object reversal paradigm.

Existing analysis methods: Conventional analyses of learning in this type of multi-day task collapse performance data within a day (i.e. ignore information regarding object pairs) (Shamy et al., 2010; Burke et al., 2014; Rapp and Amaral, 1989) and make use of analysis of variance (ANOVA) to identify between group learning differences. Because it was hypothesized in this experiment that pairs presented at the start and end of the 40 object pairings might be recalled more easily – primacy and recency effects – the original analysis in Gray et al. (2016) of data from these experiments

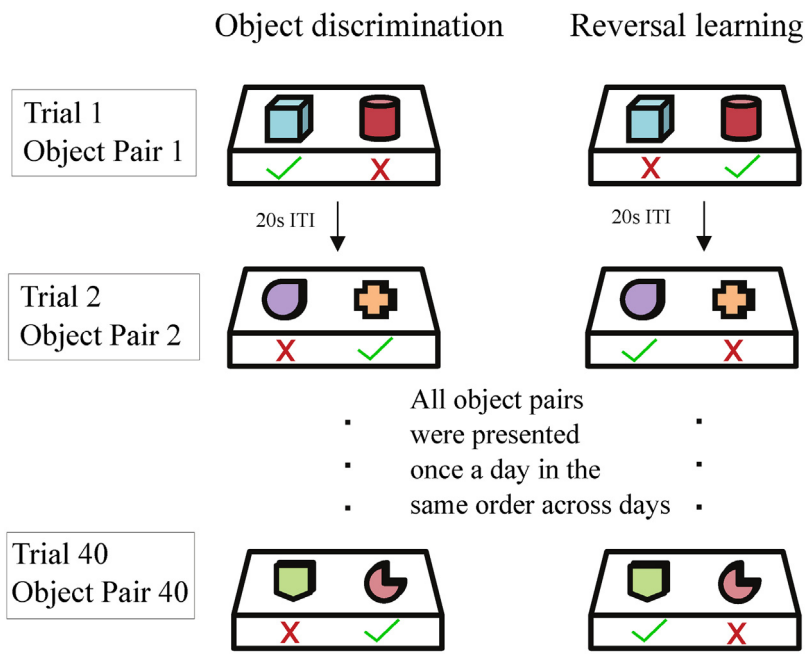

Fig. 1. Depiction of the object reversal learning paradigm. The object discrimination part of the learning task (left column) requires monkeys to learn 40 novel object pairs presented sequentially in the same order every session until they reach a 90% performance criterion. After reaching criterion, the object-reward associations are switched (reversal learning, right column), and the monkeys re-learn the new associations to the same criterion.

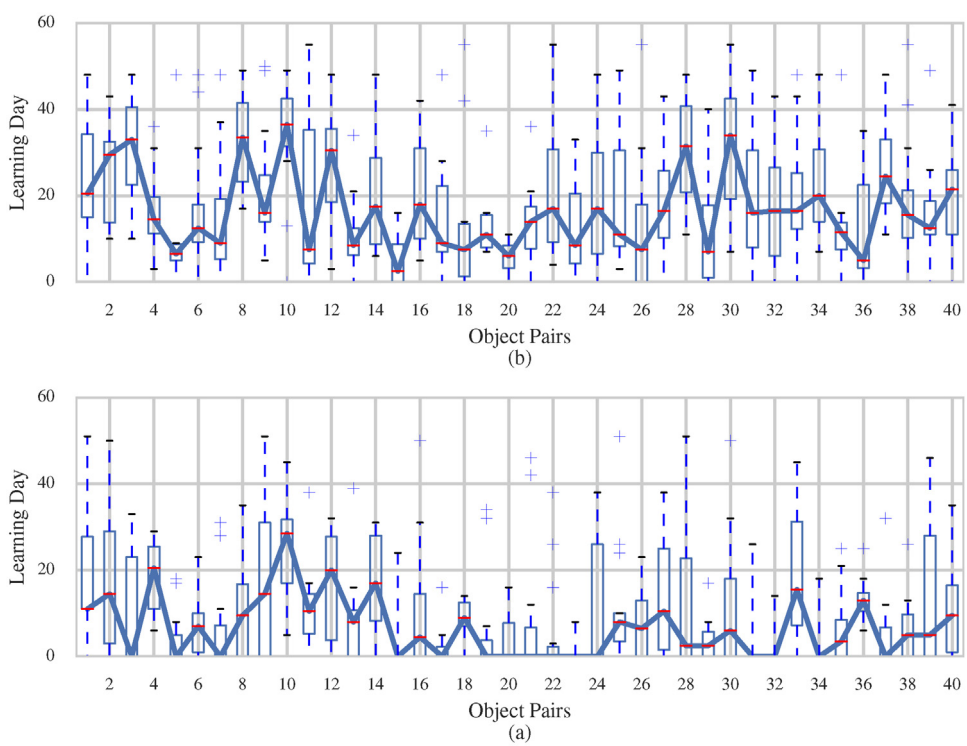

Fig. 2. Summary of estimated learning day for each object pair. Boxplots showing learning day estimated for all 40 object pairs in object discrimination (OD) (a) and reversal learning (RL) (b) parts of the task. Learning days are computed individually for each animal using a state-space approach. Boxes indicate the interquartile range (IQR) with whiskers extending to the most extreme data point that is no more than 1.5 IQR from the edge of the box. Blue line joins the median values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

examined within-day effects by using a hybrid state-space linear mixed model. A three-step process was employed. First, for each animal and object pair, the across-day performance was summarized by computing one learning day for the object discrimination task and one learning day for the reversal learning task, using one-dimensional state-space models (Smith et al., 2004). Fig. 2(a) shows a plot of boxplots of the learning days during the object discrimination task as a function of object pair. Fig. 2(b) shows the equivalent boxplots for the reversal learning task. For each object-pair, as many state-space models were fit as there were animals. As can be observed (and at odds with the literature on typical primacy and recency effects), there is a general trend towards quicker learning of the object-pairs in the middle of the 40-trial daily task. Second, as a consequence of the observation, these learning days were orga-

nized for each animal into three epochs: the first 14 object-pairs (epoch 1), the second 12 object-pairs (epoch 2) and the final 14 object pairs (epoch 3). Third, a linear mixed effects model (LME) was applied with age, task and epoch as factors. Results suggested that during both tasks, the second and third group of objects were learned significantly faster than were the first third of the objects. Also, the middle third of objects were learned significantly faster than the last third of objects. There were no significant age interactions with task or epoch. We hypothesize that this is due to the fact that the model, by partitioning every day of the experiment into three epochs, did not contain fine enough granularity to detect age differences at the level of object-pairs. In contrast (Section 4), we will see that the Laplace SMuRF model which deals with the within-day dynamics in greater detail is able to tease out between-group and inter-individual differences that are predictive of the group membership of the animals. In particular, we will show that a model that takes into account the within-day dynamics has better predictive power compared to one that does not, as measured by the Watanabe-Akaike Information Criterion (WAIC, Vehtari et al., 2017). The WAIC is a measure of predictive performance for Bayesian models that is an analogue to the Akaike Information Criterion (AIC, Akaike, 1998). The WAIC and AIC can each be thought of as an approximation to some form of cross-validation (Stone, 1977).

3. Methods

3.1. Laplace SMuRF model

The goal is to model the probability, $\pi^m_{k,r}$, of the event that a given monkey m picks the correct object in a given pair r, on a given day k. A precise assessment of this trial (object pair) and day-dependent probability would provide a complete characterization of the dynamics of learning of a monkey during the reversal learning task. We propose a *group-dependent* separably-Markov random field (SMuRF) model (Zhang et al., 2018) for these data. For simplicity, we describe the SMuRF model for a single group, with the understanding that the dynamics of learning within each group (young or old) follow separate SMuRF models. The Laplace SMuRF model expresses the dynamics of learning of a monkey as

$$\begin{cases} d_k = d_{k-1} + \delta_k, \quad \delta_k \sim \mathcal{N}(0, \sigma^2_\delta) \\ x^m_r = x^m_{r-1} + \epsilon^m_r, \quad \epsilon^m_r \sim \mathcal{L}\text{aplace}(\lambda) \\ \pi^m_{k,r} = \dfrac{e^{x^m_r + d_k}}{1 + e^{x^m_r + d_k}} \\ \Delta N^m_{k,r} | d_k, x^m_r \sim \text{Bernoulli}(\pi^m_{k,r}). \end{cases} \tag{1}$$

In Eq. (1), $(d_k)_{k=1}^K$ is a group-level latent random walk process that represents the group performance on object pairs presented within a day (within-day learning), and $(\delta_k)_{k=1}^K$ are independent and identically distributed (i.i.d.) Gaussian random variables with mean zero and variance σ^2_δ. $(x^m_r)_{r=1}^{R_m}$ represents monkey m's latent state of learning across days (cross-day learning). Across days the noise variables, $(\epsilon^m_r)_{r=1}^{R_m}$, are i.i.d. Laplace random variables (Eltoft et al., 2006) with parameter λ. Compared to a Gaussian prior, the Laplace prior can capture abrupt changes in cross-day learning due to the reversal of object-reward associations. The probability $\pi^m_{k,r}$ of the event that a given monkey m picks the correct object in a given pair k on a given day r is related to these latent processes via a logit link function which constrains the probabilities between 0 and 1. Responses are denoted by $(\Delta N^m_{k,r})_{k=1,r=1}^{K,R_m}$ which is 1 if monkey m produces a correct response when presented with object pair k on day r, and 0 otherwise. Across days, each animal is allowed its own personal learning curve that results in a latent state sequence $(x^m_r)_{r=1}^{R_m}$. There are a total of $K = 40$ object pairs, and the total number of days R_m of the experiment depends on the animal (Gray et al., 2016).

3.2. Parameter estimation in the SMuRF model by maximum likelihood

Let $\mathbf{x}_m = (x^m_1, \ldots, x^m_{R_m})^\mathrm{T}$ be a vector that denotes the learning state of the mth monkey, $\mathbf{x} = (\mathbf{x}^\mathrm{T}_1, \ldots, \mathbf{x}^\mathrm{T}_M)$ be a vector of learning states for all monkeys, and $\mathbf{d} = (d_1, \ldots, d_K)^\mathrm{T}$ be a vector that represents within-day group performance. We denote the response data from the mth monkey by $\boldsymbol{\Delta}\mathbf{N}^m = \left\{\Delta N^m_{k,r}\right\}_{k=1,r=1}^{K,R_m}$, $m = 1, \ldots, M$.

The goal is to maximize, with respect to the unknown parameters, $\theta = (\sigma^2_\delta, \lambda)$, the log-likelihood $L\left(\theta | (\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M\right)$ of the SMuRF model

$$\begin{aligned} L\left(\theta | (\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M\right) &= \log p\left((\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M; \theta\right) \\ &= \log \int_{\mathbf{x},\mathbf{d}} p\left((\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{x}, \mathbf{d}; \theta\right) d\mathbf{x} d\mathbf{d}. \end{aligned} \tag{2}$$

This is a challenging problem because of the high-dimensional integral that must be evaluated in Eq. (2). We propose to maximize this likelihood by Expectation-Maximization (EM). Given a candidate solution, EM (Dempster et al., 1977) builds an approximation of the likelihood in its E-step. Maximizing this approximation in the M-step is guaranteed to yield a new candidate solution that does not decrease the value of the likelihood function. By iterating this process, EM generates a sequence of iterates that converge to a local optimum of the likelihood function. Let $\theta^{(\ell)}$ be the candidate solution in the ℓth iteration of EM. In the context of the Laplace SMuRF model, the key challenge of EM is to compute $\mathcal{Q}(\theta | \theta^{(\ell)})$ defined as

$$\mathcal{Q}(\theta | \theta^{(\ell)}) = \mathbb{E}_{\mathbf{x},\mathbf{d}}\left[\log p((\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{x}, \mathbf{d}; \theta) | (\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \theta^{(\ell)}\right], \tag{3}$$

the expected value of the *complete-data likelihood* with respect to the *joint* posterior distribution of the missing data $(\mathbf{x}, \mathbf{d})$ conditioned on the observed data $(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M$ and the candidate solution $\theta^{(\ell)}$. This expectation is not tractable, i.e., it cannot be computed in closed-form. We propose to approximate the required expectations using Markov-Chain Monte-Carlo (MCMC) samples from $p(\mathbf{x}, \mathbf{d} | (\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M; \theta^{(\ell)})$. In particular, we will use Gibbs sampling (Casella and George, 1992), a Monte-Carlo technique, to generate samples from a distribution by sampling from its so-called full conditionals (conditional distribution of one variable given all others), thus generating a Markov chain that, under regularity conditions, can be shown to converge to a sample from the desired distribution.

Gibbs sampling is attractive in cases where sampling from the full conditionals is simple. However, sampling from the full conditionals $p(\mathbf{x} | \mathbf{d}, (\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M; \theta^{(\ell)})$ and $p(\mathbf{d} | \mathbf{x}, (\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M; \theta^{(\ell)})$ is not trivial, for reasons detailed in Zhang et al. (2018). We circumvent this difficulty by considering, instead of Eq. (1), the following equivalent *augmented* SMuRF model

$$\begin{cases} d_k = d_{k-1} + \delta_k, \quad \delta_k \sim \mathcal{N}(0, \sigma^2_\delta) \\ x^m_r = x^m_{r-1} + \nu^m_r \sqrt{z^m_r}, \quad \nu^m_r \sim \mathcal{N}(0, 1) \\ \pi^m_{k,r} = \dfrac{e^{x^m_r + d_k}}{1 + e^{x^m_r + d_k}} \\ \Delta N^m_{k,r} | d_k, x^m_r \sim \text{Bernoulli}(\pi^m_{k,r}) \\ w^m_{k,r} \sim \text{PG}(1, d_k + x^m_r) \\ z^m_r \sim \exp\left(\dfrac{\lambda^2}{2}\right) \end{cases} \tag{4}$$

where $(w_{k,r}^m)_{r=1,k=1}^{R_m,K}$ are i.i.d. Polya-Gamma (PG) (Polson et al., 2013) random variables, $m = 1, \ldots, M$, and $(z_r^m)_{r=1}^{R_m}$ are i.i.d. exponential random variables with parameter $\frac{\lambda^2}{2}$. To simplify notation, let $\mathbf{z}_m = (z_1^m, \ldots, z_{R_m}^m)^{\mathrm{T}}$, $\mathbf{z} = (\mathbf{z}_1^{\mathrm{T}}, \cdots, \mathbf{z}_M^{\mathrm{T}})^{\mathrm{T}}$, $\mathbf{W}^m \in \mathbb{R}^{R_m \times K}$, and $\mathbf{W}_{k,r}^m = (w_{k,r}^m)_{r=1}^{R_m}$ for $m = 1, \ldots, M$. The models in Eqs. (1) and (4) are equivalent in the sense that the joint distribution of $(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{x}, \mathbf{d}$ from the former is equivalent to that from the latter *marginalized* over $(\mathbf{W}^m)_{m=1}^M$ and $\mathbf{z}$. In what follows, it will be useful to let $y_r^m = 1/z_r^m$, $\mathbf{y}_m = (y_r^m, \ldots, y_{R_m}^m)^{\mathrm{T}}$ and $\mathbf{y} = (\mathbf{y}_1^{\mathrm{T}}, \ldots, \mathbf{y}_M^{\mathrm{T}})^{\mathrm{T}}$

Block Gibbs sampling from augmented SMuRF model: The augmented SMuRF model of Eq. (4) yields the following block Gibbs sampler, the derivation of which is in the Appendix. This Gibbs sampler forms the basis of the E-step of our Monte-Carlo EM algorithm for maximizing the likelihood for the SMuRF model of the learning dynamics. The full-conditionals (distributions from which one must obtain samples) for the Gibbs sampler are as follows

- $p(\mathbf{x}|(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{d}, \mathbf{z}, (\mathbf{W}^m)_{m=1}^M; \theta^{(\ell)}) = \prod_{m=1}^M p(\mathbf{x}_m|\boldsymbol{\Delta}\mathbf{N}^m, \mathbf{d}, \mathbf{z}_m, \mathbf{W}^m; \theta^{(\ell)})$. In other words, conditioned on group performance $\mathbf{d}$ and all of the data $(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M$, the learning states $\mathbf{x}_1, \ldots, \mathbf{x}_M$ of the M monkeys in the group are independent. We show in the Appendix that, for each $m = 1, \ldots, M$, $p(\mathbf{x}_m|\boldsymbol{\Delta}\mathbf{N}^m, \mathbf{d}, \mathbf{z}_m, \mathbf{W}^m; \theta^{(\ell)})$ is the posterior distribution of the state sequence in a linear-Gaussian state-space model. This leads to a very efficient forward-filtering backward-sampling algorithm.
- $p(\mathbf{d}|(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{x}, \mathbf{z}, (\mathbf{W}^m)_{m=1}^M; \theta^{(\ell)})$. This represents the full-conditional of within-day group performance, which obeys properties similar to the previous full conditional (by symmetry).
- $p(\mathbf{y}|(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{x}, \mathbf{d}, (\mathbf{W}^m)_{m=1}^M; \theta^{(\ell)}) = \prod_{m=1}^M \prod_{r=1}^{R_m} p(y_r^m|x_r^m, x_{r-1}^m; \theta^{(\ell)})$. A result from Eltoft et al. (2006) tells us that $p(y_r^m|x_r^m, x_{r-1}^m; \theta^{(\ell)}) \sim \mathrm{IG}\left(y_r^m; \frac{\lambda^{(\ell)}}{|x_r^m - x_{r-1}^m|, \lambda^{(\ell)2}}\right)$, where $IG(\cdot,)$ denotes the inverse Gaussian distribution. The full conditional for $\mathbf{z}$ is the product of independent inverse Gaussians, where the independence is across animals and days of the task.
- $p(\mathbf{W}|(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{x}, \mathbf{d}, \mathbf{z}; \theta^{(\ell)}) = \prod_{m=1}^M \prod_{r=1}^{R_m} \prod_{k=1}^K p(w_{k,r}^m|x_r^m, d_k)$. From Polson et al. (2013), $p(w_{k,r}^m|x_r^m, d_k) \sim \mathrm{PG}(1, x_r^m + d_k)$ that is, the full conditional for $\mathbf{W}$ is the product of independent Polya-Gammas, where the independent variable is across animals, days of the task, and object pairs within a day.

Derivations of the form of the first two full conditionals are in the Appendix. The forms of the last two full conditionals follow from the form of the SMuRF model, and results from Eltoft et al. (2006) and Polson et al. (2013) respectively.

In the M-step, we update the parameters σ_δ^2 and λ by maximizing $\mathcal{Q}(\theta|\theta^{(\ell)})$ with respect to θ. Following relatively simple algebra detailed in the Appendix, we obtain

$$\lambda^{(\ell+1)} = \left(\frac{2\sum_{m=1}^M R_m}{\sum_{m=1}^M \sum_{r=1}^{R_m} \mathbb{E}[z_r^m|x_r^m, x_{r-1}^m; \lambda^{(\ell)}]}\right)^{1/2} \tag{5}$$

$$\sigma_\delta^{2(\ell+1)} = \frac{1}{K}\sum_{k=1}^K \mathbb{E}\left[(d_k - d_{k-1})^2|(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{x}, \mathbf{z}, (\mathbf{W}^m)_{m=1}^M; \theta^{(\ell)}\right]. \tag{6}$$

where we approximate the necessary expectations using the Monte-Carlo samples from the E-step.

Computational complexity of block Gibbs sampling: Sampling from the full-conditional of learning states $\mathbf{x}_1, \ldots, \mathbf{x}_M$ of the M monkeys is the most computationally expensive step of the block Gibbs sampler. For each Gibbs sampling iteration, the complexity of this step is $\mathcal{O}(M \cdot R)$, where R is the number of days the slowest animal (young or old) takes to perform the experiment. That is, the complexity is linear in the maximum number of days R it takes an animal to complete the experiment. This is because the cost of forward-filter backward sampler is the same as the cost of a Kalman smoother which, for a given animal m, is linear in the number of days R_m it takes the animal to complete the experiment. Since the full-conditional factorizes over animals, we sample independently for each animal and therefore the cost per Gibbs iteration is also linear in the number of animals M in the group. To obtain the cost of the EM algorithm, we multiply the cost per Gibbs iteration by the number of Gibbs sample per EM iteration, and by the number of EM iterations.

We note that, compared to the block Gibbs sampler for the Gaussian SMuRF (Zhang et al., 2018), the Laplace SMuRF requires an additional step – third bullet point above – making it slightly more expensive computational than the Gaussian SMuRF.

3.3. Assessing the dynamics of learning from binary response data

Bayesian estimation of the Laplace SMuRF model enables us to infer detailed changes in learning dynamics, in particular to extract the within-day and cross-day components of performance dynamics that accompany the learning of a task by an animal. This is because, following estimation, inference in the Laplace SMuRF model yields the *joint* posterior distribution of the within and cross-day components of learning conditioned on the observed data. We can use this posterior distribution, in turn, to assess between group differences in performance within a day, as well as inter-individual differences in learning across days. If the model fits the data better than competing ones, then it forms a better basis for performing inference (e.g. assessing the effect of age on performance). We let $p\left(\mathbf{x}, \mathbf{d}|(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M; \hat{\theta}_{ML}\right)$ denote the posterior distribution of $\mathbf{x}$ and $\mathbf{d}$, given the binary response data$(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M$ and the maximum likelihood estimate $\hat{\theta}_{ML}$ of θ. It is understood that we use Gibbs samples $(\mathbf{x}_i, \mathbf{d}_i)_{i=1}^n$ from $p\left((\mathbf{x}, \mathbf{d}|\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M; \hat{\theta}_{ML}\right)$ to obtain an empirical estimate of the distribution. In order to compare the within-day group performance and the cross-day individual performances we estimate the mode of the posterior distribution using a maximum a posteriori (MAP) estimate. For the derivation of the MAP estimator, we refer the reader to the Appendix.

Model comparison: The Watanabe-Akaike Criterion (WAIC) (Vehtari et al., 2017) is a measure of model fit that balances the goodness-of-fit of a model with the number of parameters used. It can be computed from (Vehtari et al., 2017).

$$\mathrm{WAIC} = -2(\widehat{\mathrm{lpd}} - \hat{p}_{\mathrm{waic}}) \tag{7}$$

where $\widehat{\mathrm{lpd}}$ is the computed log pointwise predictive density

$$\widehat{\mathrm{lpd}} = \sum_{m=1}^M \sum_{k=1}^K \sum_{r=1}^{R_m} \log\left(\frac{1}{n}\sum_{i=1}^n p\left(\Delta N_{k,r}^m|x_r^{i,m}, d_k^i\right)\right) \tag{8}$$

and $\hat{p}_{\mathrm{waic}}$ is a measure of model complexity

$$\hat{p}_{\mathrm{waic}} = \sum_{m=1}^M \sum_{k=1}^K \sum_{r=1}^{R_m} V_{i=1}^n\left(\log p\left(\Delta N_{k,r}^m|x_r^{i,m}, d_k^i\right)\right) \tag{9}$$

where $V_{i=1}^n(a_{k,r}^{i,m}) = \frac{1}{n-1}\sum_{i=1}^n (a_{k,r}^{i,m} - \bar{a}_{k,r}^{i,m})^2$ is the sample variance and $a_{k,r}^{i,m} = \log p\left(\Delta N_{k,r}^m|x_r^{i,m}, d_k{}^i\right)$. Because the models for young and old are independent we can compute the WAIC (Eq. (7)) for each age group (for a given age group, M is understood as the number of animals in that group) separately and add them to obtain the

cumulative WAIC. A model with a lower WAIC can be assumed to enable better inference about the data.

Within-day learning differences: We can use our models to compute, for any of the object pairs, the posterior probability that young monkeys would perform better on that pairing than old monkeys. This probability is

$$\mathbb{P}[d_k^{\text{young}} > d_k^{\text{old}} | (\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M; \hat{\theta}_{ML}] \approx \frac{1}{n}\mathbb{I}_{\{d_k^{i,\text{young}} > d_k^{i,\text{old}}\}}, \tag{10}$$

where we approximate it by its frequency of occurrence in the posterior samples from the model.

Cross-day learning differences: We define a monkey's cognitive flexibility measure as the difference between its learning rate before and after the reversal. Assuming reversal for monkey m occurs on day D_m, the monkey's cognitive flexibility is

$$\text{CF}^m = \left(\frac{1}{R_m - D_m}\sum_{r=D_m+1}^{R_m}(x_r^m - x_{r-1}^m)\right) - \left(\frac{1}{D_m}\sum_{s=1}^{D_m}(x_s^m - x_{s-1}^m)\right). \tag{11}$$

In practice, each Gibbs sample $(\mathbf{x}_i, \mathbf{d}_i)$ yields one number

$$\text{CF}^{i,m} = \left(\frac{1}{R_m - D_m}\sum_{r=D_m+1}^{R_m}\left(x_r^{i,m} - x_{r-1}^{i,m}\right)\right) - \left(\frac{1}{D_m}\sum_{s=1}^{D_m}\left(x_s^{i,m} - x_{s-1}^{i,m}\right)\right). \tag{12}$$

Repeating this calculation for all n Gibbs samples yields an empirical posterior distribution for the cognitive flexibility of the mth animal.

Armed with this definition of cognitive flexibility, we can compare animals by computing the probability that animal m is more cognitively flexible than animal $\tilde{m}$

$$\mathbb{P}[\text{CF}^m > \text{CF}^{\tilde{m}} | (\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M; \hat{\theta}_{ML}] \approx \frac{1}{n}\mathbb{I}_{\{\text{CF}^{i,m} > \text{CF}^{i,\tilde{m}}\}}. \tag{13}$$

4. Results

In what follows, we will consider two classes of models of the data from the reversal learning task. **Model Class 1** comprises two separate Laplace SMuRF models, one for young monkeys, and another for old monkeys. It assumes that animals within a group follow the same within-day learning profile. We consider two subclasses, **Model Class 1A** and **Model Class 1B**. In subclass A, we assume there are no dynamics within a day i.e. the random walk for d_k defined in Eq. (1) is set to have a negligibly small variance. In subclass B, we allow the possibility of dynamics by estimating the random walk variance σ_δ^2 from the data. Comparing the predictive accuracy of **Model Class 1A**, which assumes no within-day dynamics, to **Model Class 1B**, which assumes the presence of within-day dynamics, would let us determine the importance of dynamics for the purpose of making inferences in this experiment (e.g. the effect of age on performance). In **Model Class 2**, we assume that each animal is in a group of its own. That is, **Model Class 2** comprises 14 separate Laplace SMuRF models, one for each animal.

4.1. Model comparison to assess importance of within-day dynamics

We use the WAIC to assess the predictive accuracy of **Model Class 1A** compared to **1B** (see Table 1). The WAIC values reported in the first and second column confirm that incorporating within-day dynamics leads to lower WAIC, i.e. a model of the multi-day behavioral experiment that is better predictive of animal performance across the experiment. For the remainder of the paper, references to **Model Class 1** should be understood as the dynamic version (**Model Class 1B**).

Table 1
Comparison of the predictive accuracy of the three models of the multi-day behavioral experiment based on the WAIC. Lower values of the WAIC signify better predictive accuracy. The WAIC values reported in the first (**Model Class 1A**) and second column (**Model Class 1B**) confirm that incorporating within-day dynamics (second column) leads to a model of the multi-day behavioral experiment that is better predictive of animal performance across the experiment. The third column compares the predictive accuracy of the dynamic Laplace SMuRF model (**Model Class 1B**) to one that employs the Student-*t* distribution instead, termed Student-*t* SMuRF. The Student-*t* distribution is a popular choice to capture abrupt changes in data. The WAIC values indicate that the Laplace SMuRF model is a better fit to the data compared to the Student-*t* SMuRF.

	Laplace SMuRF (static)	Laplace SMuRF (dynamic)	Student-*t* SMuRF (dynamic)
WAIC	36,929	**35,988**	37,959

4.2. Student-t alternative to Laplace prior

The Student-*t* distribution is a popular choice to capture abrupt changes in data. The third column in Table 1 compares the predictive accuracy of the dynamic Laplace SMuRF model to one that employs the Student-*t* distribution instead, termed Student-*t* SMuRF. The WAIC values indicate that the Laplace SMuRF model (**Model Class 1B**) is a better fit to the data compared to the Student-*t* SMuRF. We hypothesize that this is because the Student-*t* distribution does not capture the abruptness in change in performance due to the object-pair reversals as accurately as the Laplace distribution.

4.3. Comparing learning between groups

Two objectives of the study in Gray et al. (2016) were to identify learning effects within a daily session and establish age differences in ability to learn tasks known to rely on different brain networks. Using **Model Class 1**, we can perform these comparisons easily with our paradigm because it provides a summary for learning within the day for old and young animals, and a learning curve across days for each animal.

As described above, we can use the Gibbs samples yielded at convergence of the Monte Carlo EM algorithm to compare the groups' performance on the different object pairs (Eq. (10)). We consider the overall item-pair performance at one point in the daily session to be greater than the performance at another point if this probability is greater than 0.95. Learning across days can be analyzed using the probability distribution of each animal's learning curve.

Within-day performance for young and old animals is shown in Fig. 3(a). Both groups show similar trends. Performance was low for two distinct periods near the start each day (item pairs 0–3 and 7–11), reached a peak near item pair 20, and tailed down at the end of the session (item pairs 26 onwards).

In general the older animals perform equivalently or worse than the younger animals. Comparison between the underlying distributions is shown in Fig. 3(b). The curve shows the item-by-item probability that the young animals performance is better than the old animals. When the curve lies above 0.95 we can be fairly confident that the young animals performance is better than the old animals. Thus, from pairs 6–16 and pairs 24–39 (with the exception of pair 37) we observe that, according to the model, with 95% certainty, the younger animals perform better than older.

Fig. 4(a) and Fig. 4(b) show the individual learning curves for the young and old monkeys, respectively, aligned by the day of reversal. In general the older monkeys took longer to learn both before and after reversal. Our model does not *a priori* know the reversal point in the data but in all cases accurately identifies the change point.

Fig. 5 shows raw data in the form of a raster and a summary of our analyses for one young monkey and one old monkey. The black dots in the raster plots indicate rewarded, correct behavior,

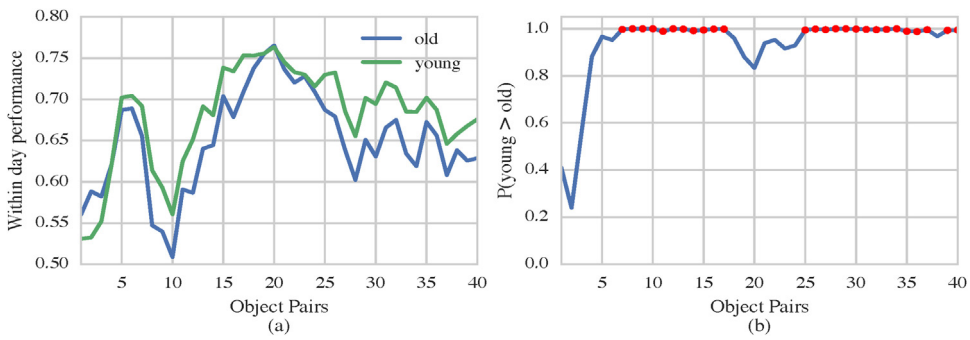

Fig. 3. Within-day performance comparison between young and old animals. (a) MAP estimates of the within-day performance of young (green) and old (blue) animals across all 40 object pairs. (b) Probability that the young animals' performance is greater than the old animals' (Eq. (10)). Values where the probability is greater than 0.95 are indicated with red dots.

whereas a space indicates unrewarded, incorrect behavior. The red dashed line highlights the day when reversal of reward occurs.

The bottom panels depict the within-day performance curve for both the young and the old group. The left panel of Fig. 5(a) and (b) are the MAP estimates of the learning curves for the corresponding young and old monkey, respectively. This figure demonstrates that the Laplace SMuRF model can provide a succinct summary of the dynamics of learning in this behavioral experiment across multiple days, and with multiple observations within a day.

4.4. Predicting age-group membership from behavior

In this section, we assume the subjects' ages are not known and investigate whether their behavior naturally falls into two or more groups using clustering. Three metrics (described below) are considered: a single cognitive flexibility measure and two simple daily performance measures (i.e., averaged response data across days and model-based within-day performance).

Clustering based on cognitive flexibility: A single monkey's cognitive flexibility measure is defined as the difference between its learning rate before and after the reversal. The learning rate in the pre and post-reversal periods is calculated based on the Gibbs sample estimate rate differences (Eq. (11)). The resulting difference provides an empirical cognitive flexibility distribution for each animal. The measure was chosen for this experiment as each animal performed the task over a different number of days. We used **Model Class 1** to compute the empirical cognitive flexibility distribution of each animal (Eq. (11)).

To visualize comparison of this measure with that of other monkeys, we plot a cognitive flexibility heatmap (Fig. 6(a)). Each block in the heatmap reflects the probability that the monkey on the y-axis is more cognitively flexible than the monkey on the x-axis (Eq. (12)). These values are computed based on comparing samples from each animal's cognitive flexibility posterior distribution. A monkey is defined here to be more flexible than another monkey if, with high probability, its cognitive flexibility is higher than that of the other monkey. We cluster the individual monkeys into groups using the hierarchical clustering algorithm from Müllner (2018). Clusters are shown in Fig. 6, with gray exterior lines. Clustering results indicate two main clusters: one with five old animals (o_1, o_3, o_4, o_5, o_7) and a second with all young animals and three old animals (o_2, o_6, o_8). This result is consistent with other aging studies that indicate that there is considerable individual variation in effects of aging on learning and memory (Roberts et al., 2014; Spiegel et al., 2013; Gray et al., 2016).

Clustering based on simple measures: The clustering method described above is based on a probabilistic definition of cognitive flexibility derived from the Laplace SMuRF model. We compare this method with two possible alternate methods based on mean daily measures. The first is based on the average of the response data across days, resulting in one 40-dimensional vector for each animal. The second is based on **Model Class 2**, which yields empirical posterior averages of the within-day performance $(d_k^m)_{k=1}^{40}$ for each animal, $m = 1, \ldots, 14$. This method also results in one 40-dimensional vector for each animal. It produces a smoother estimate of daily performance than the average of the response data across days.

Fig. 6(b) and Fig. 6(c) show the corresponding surfaces representing these values. For the first metric based on raw data (Fig. 6(b)), the main two groups are a mixture of old and young animals (Group 1: y_4, y_5, o_3, o_4, o_5, Group 2: $y_1, y_2, y_3, y_6, o_1, o_2, o_8$). For the second metric based on the latent process (Fig. 6(c)) the main two groups are divided into a single young animal (y_3) and the rest of the animals. These latter two methods are harder to interpret than one based on cognitive flexibility and do not provide predicted group memberships as consistent with the true groups.

4.5. Simulation study: predicting performance-group membership from behavior

In the previous analysis, we were able to use cognitive flexibility (Eq. (12)) to cluster the animals into groups that are, for the most part, consistent with the young and old age groups. In agreement with their learning curves (Fig. 4), monkeys o_2, o_6, o_8 were assigned to the cluster containing all young animals, suggesting that they are high-performing old animals. This is a common observation in aging studies. Ideally, we would want to cluster animals based on performance, that is, to use the binary response data to assign every animal to one of two groups: a low-performing group, and a high-performing one.

In the context of an experiment, we cannot assess the ability of a model such as the Laplace SMuRF model to produce clusters consistent with ground-truth performance level as the only information available is age. In this section, we design a simulated reversal learning experiment comprising two groups of monkeys differing in their intrinsic performance levels, as opposed to age. We begin by assessing the extent to which cognitive flexibility computed based on **Model Class 2** can produce clusters consistent with the low and high performing groups. Then, we assess how well the learning state can be used for clustering in case all monkeys have the same number of trials. We cannot do this for the experiment from Section 2 because the 90% stopping criterion leads to learning state vectors of different length for different animals. We use the sensitivity and positive predictive value (PPV) curves to quantify the performance of each clustering method as we vary the extent to which the groups can be differentiated.

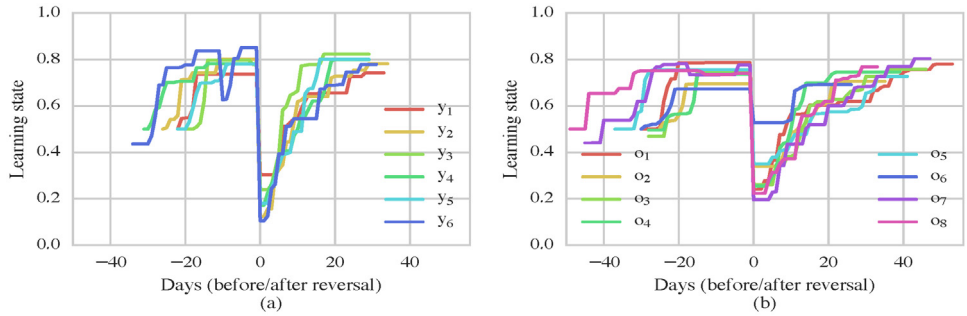

Fig. 4. MAP estimates of cross-day performance for all 14 animals. (a) Young animals ($y_1, \ldots, y_6$). (b) Old animals ($o_1, \ldots, o_8$). Note that older animals tend to take longer to learn both before and after reversal and that the model clearly captures the sharp change in behavior after reversal.

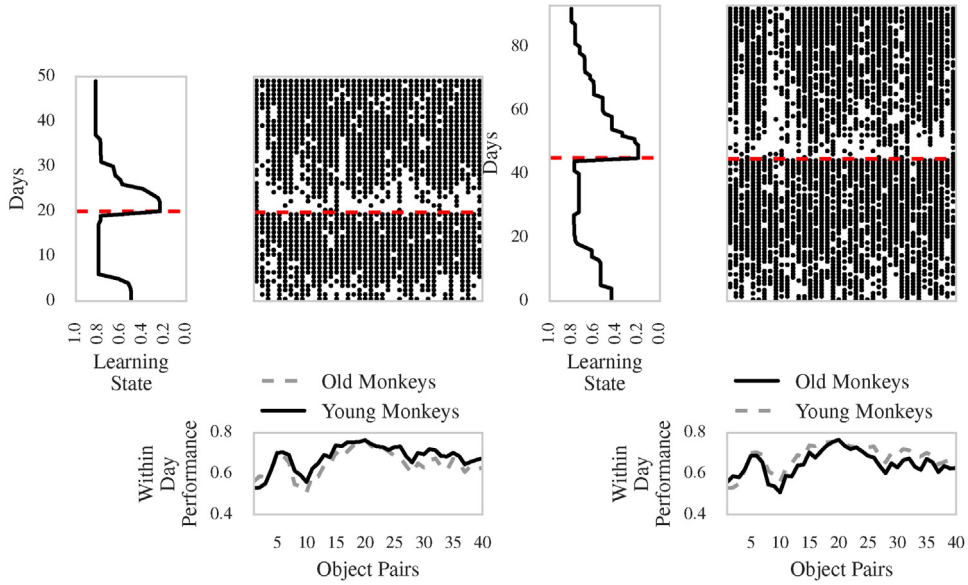

Fig. 5. Summary of analyses of representative young and old monkeys. (a) Young animal. (b) Old animal. The raster plots indicate performance within day (x-axis) and across days (y-axis). A black dot indicates rewarded, correct behavior, and a space indicates unrewarded, incorrect behavior. The red dashed line highlights the day when reversal of reward occurs. In each panel, the plot on the left represents the MAP estimate of the learning state for the animal. Both bottom panels show the within-day performance curves from Fig. 3(a).

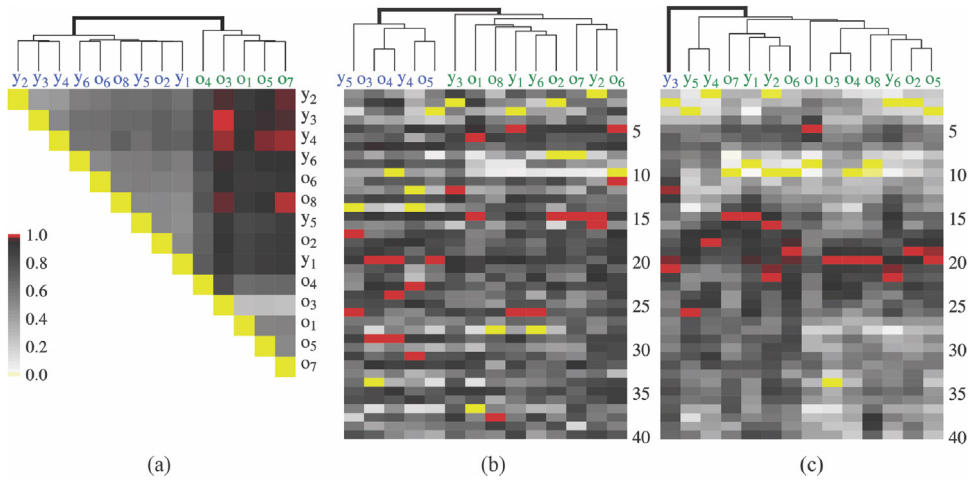

Fig. 6. Clustering of animal responses based on raw data without knowledge of age. (a) Heatmap of relative cognitive flexibility for each animal (Eq. (13)). Shading indicates the probability level that one animal's flexibility is higher than a second animal's. Values greater than 0.95 are indicated in red. Individual animals are labeled by their actual chronological age (old are $o_1, \ldots, o_8$ and young are $y_1, \ldots, y_6$) even though we do not use this information. (b) Heatmap indicating average response across days (y-axis) for each animal. (c) Empirical posterior average of the within-day performance for each animal, obtained by fitting a separate Laplace SMuRF to each animal. This leads to estimates of daily performance that are smoother than in panel (b). In each panel, the gray lines above the heatmap indicate the hierarchical group clusters and the bold line indicates the clustering into two groups according to which the labels are colored.

Simulation paradigm for binary response data from a reversal-learning experiment: We propose a framework for simulating binary response data in a reversal learning experiment. In the first phase of the task, we assume each simulated animal is presented each day with the same 40 object pairs ($K = 40$) in the same order. The first phase of the task ends after 30 days, followed by a second phase (the reversal phase), which lasts 40 days, for a total of $R = 70$ days for all simulated animals.

Model of binary response data from simulated experiment: The simulated data are binary outcomes encoding the occurrence of the event that a simulated animal picks the correct object in a given pair, on a given day. We express the probability of this event as a function of a within-day learning process that reflects the dependence of an animal's performance on object pairs, and a cross-day learning process that captures the dynamics of proficiency across days.

We consider two groups of animals: a high-performing (H) one and low-performing (L) one. We assume that all animals, irrespective of their group membership, follow the same within-day learning process. That is, the only difference between the two groups is the evolution of the cross-day learning process.

Motivated by the preceding analysis (Fig. 3), we assume that the evolution of the within-day learning process follows the parabola

$$\text{expit}(d_k) = 0.6 - 0.1\left[\frac{(k-20)}{20}\right]^2, \quad k = 1, \ldots, 40. \tag{14}$$

The cross-day learning process for the high-performing group follows the piecewise logistic functions

$$\begin{cases} \text{expit}(x_r^H) = 0.5 + \dfrac{0.4}{1+\mathrm{e}^{-0.3(r-15)}}, & r = 1, \ldots, 30 \\ \text{expit}(x_r^H) = 0.1 + \dfrac{0.8}{1+\mathrm{e}^{-0.3(r-50)}}, & r = 31, \ldots, 70. \end{cases} \tag{15}$$

The cross-day learning process for the low-performing group follows the piecewise logistic function

$$\begin{cases} x_r^L = x_r^H, \quad r = 1, \ldots, 30 \\ \text{expit}(x_r^L) = 0.1 + \dfrac{p_R^L - 0.1}{1+\mathrm{e}^{-0.3(r-50)}}, \quad r = 31, \ldots, 70. \end{cases} \tag{16}$$

The goal of the simulation is to determine the extent to which we can distinguish the two groups as we vary the similarity in their cross-day learning processes. Therefore, we consider 9 different values for the final level of the low performing group probability, p_R^L, ranging from 0.1 to 0.9. Fig. 7 shows plots of our models for the within and cross-day performance curves.

Simulating from the model: We simulate binary response data from the reversal-learning task by generating Bernoulli random variables for each object pair k and day r as follows:

$$\begin{cases} \pi_{k,r} = \dfrac{e^{x_r + d_k}}{1 + e^{x_r + d_k}} \\ \Delta N_{k,r} | d_k, x_r \sim \text{Bernoulli}(\pi_{k,r}). \end{cases} \tag{17}$$

We refer to d_k and x_r as the "generating curves" of the simulation. Note that x_r is either given by Eq. (15) or Eq. (16), depending on whether we are simulating data from a high or low performing animal. In the latter case, we vary p_R^L as described previously.

Simulation results: We simulated binary response data from 8 animals in each group. Fig. 8 shows the result of applying **Model Class 2** to the simulated data. The estimates of within-day performance (Fig. 8(a)) for the two groups overlap and are not significantly different. This is consistent with the fact that the generating within-day performance curve is the same for the two groups. Fig. 8(b) demonstrates the accurate reconstruction of the generating learning state curve for two example low and high performing animals.

We quantify our ability to cluster the monkeys into two groups using the cognitive flexibility measures from Section 3 (Eqs. (11)–(13)) by varying how close the cross-day learning processes of the two groups are to each other. We use $p^{H-L} = \max(p_r^H) - \max(p_r^L)$ as a scalar measure of closeness of the cross-day learning curves from the two groups. For each of the 9 possible values of p^{H-L}, we performed clustering as in Section 4 and computed the positive predictive value (PPV) and sensitivity (Fig. 9). In Fig. 9(a), for instance, we see as expected that, when there is no group difference ($p^{H-L} = 0$), the PPV and sensitivity are 0.5. We are completely able to distinguish the groups when $p^{H-L} \geq 0.5$ and perform quite well already when $p^{H-L} \geq 0.1$. The values in Fig. 9 are computed from the average of 10 simulations. Very little between-simulation variation was observed so we do not show confidence bounds.

The results from Fig. 9(a) suggest that the clustering of the monkeys using **Model Class 2** (Section 4, Fig. 6) reflects difference in performance among the two groups identified. The fact that these two groups do not overlap with the age groups underscores the utility of the Laplace SMuRF model, and our approach, for the identification of similar groups of monkeys solely based on performance, and irrespective of age.

Fig. 9(b) and Fig. 9(c) show the PPV and sensitivity plots obtained by clustering using the simple measures described in Section 4 (Fig. 6(b) and (c)). They further demonstrate the strength of clustering based on differences in cognitive flexibility (Eq. (13)) when the group differences are very small. Indeed, this method can identify the two groups even when $p^{H-L} \approx 0.1$, while the methods based on raw averages across object pairs (Fig. 9(b)) or using the estimates of the learning states from the model (Fig. 9(c)) cannot. This result is consistent with the fact that averaging across object pairs preserves information regarding the within day performance, which is the same for both groups, and thus reduces the ability to cluster. In contrast, our model separates the learning state from the within day performance.

5. Discussion

At present, there is no consensus on the best method to analyze population behavioral data that spans trials within a day and across many days. We have proposed a two-dimensional state-space approach for the analysis of data from two groups of animals performing multiple object pairings across days (Gray et al., 2016). The model, termed Laplace SMuRF model, expresses the probability of observing the binary outcomes as a function of two latent processes: a within-day learning process that reflects an animal's item-dependent performance level during a day, and a cross-day learning process that captures the dynamics of proficiency across days.

The standard for analysis of behavioral data is to make use of linear mixed effects (LME) models, which are preferred over t-tests as, being model based, they are amenable to probabilistic inference, i.e. to computing the posterior probability of events of interest under the model (e.g. the probability of there being an age difference given the data). Applied to behavioral data, LME models typically require the aggregation of the data over user-defined epochs, which can obfuscate the presence of group differences, as well as detailed changes in the learning dynamics. The Laplace SMuRF model presents several advantages over LME models. First, the Laplace SMuRF model is a generative model that obviates the need for aggregating the behavioral data into groups (either over days or within day Gray et al., 2016). Our analysis demonstrates that this model, which preserves object-pair information (**Model Class 1B**) (i.e. does not aggregate the behavioral data within day), provides a description of the experimental data that is more predictive, as measured by the WAIC (Vehtari et al., 2017), than one that

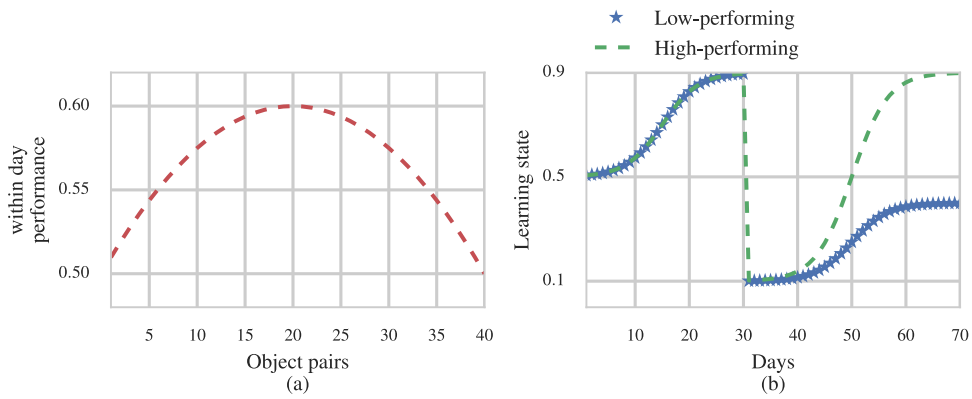

Fig. 7. Within and cross-day performance curves from simulation. (a) Within-day performance. (b) Cross-day performance.

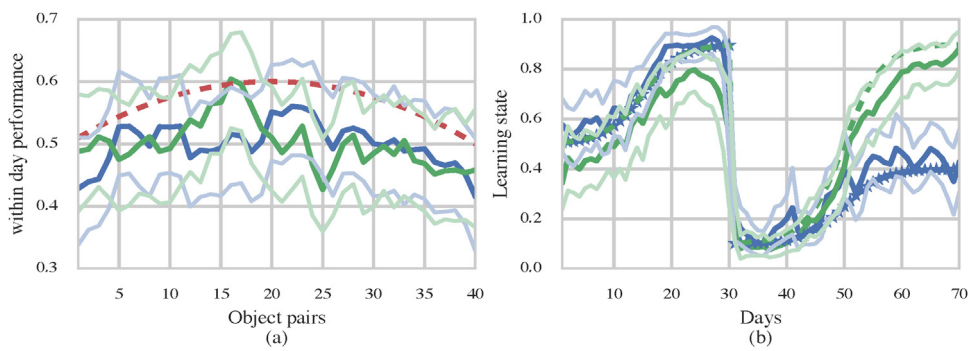

Fig. 8. Results from one example simulation. (a) The red dashed line shows true performance within days and green and blue lines show the estimated performance within days for simulated young and old animals. (b) The dashed lines indicate true performance across days. 95% confidence bounds are indicated with the lighter shaded lines.

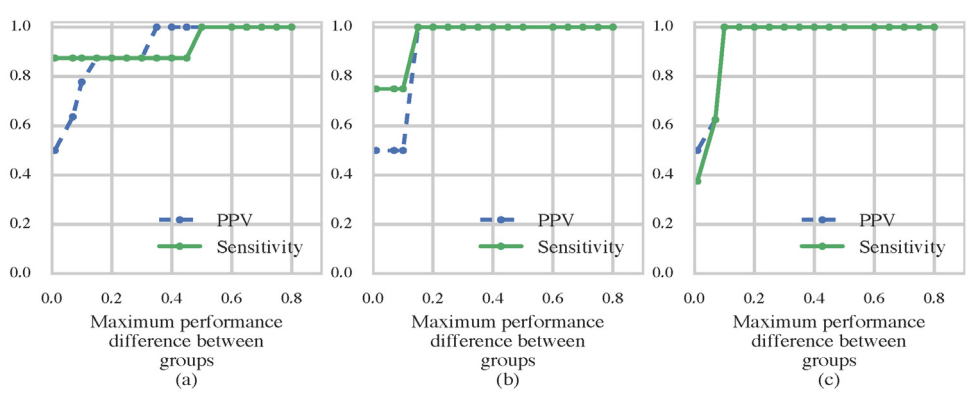

Fig. 9. Positive predictive value and sensitivity results for three clustering methods as a function of overall difference in cross-day performance (p^{H-L}). (a) Clustering based on cognitive flexibility. (b) Clustering based on raw data. (c) Clustering based on across day performance.

aggregates the behavioral data within day (**Model Class 1A**). Second, having superior predictive power, the Laplace SMuRF model is closer to the underlying true model that generated the data, and therefore better from the perspective of performing inferences. In hypothesis testing, the need to correct for multiple hypotheses stems from the absence of a model for the joint distribution of the data. By modeling this joint distribution and demonstrating the predictive power of the model, we put ourselves in a position to perform inferences, with the understanding that "all models are wrong, some are useful". Last, the Laplace SMuRF model provides a detailed characterization of the learning dynamics underlying behavioral data, leading to a more precise characterization of between-group differences. Indeed, we found that, across object pairs, young animals performed better than did older animals. This finding was not apparent from previous analyses (Gray et al., 2016) based on aggregating the data into "beginning", "middle" and "end" object pair epochs. It should be emphasized that the model does not assume *a priori* the presence of age differences. These are uncovered after fitting the model. When we define cognitive flexibility as the difference between the learning rate of an animal following reversal and that preceding reversal, we find that, as a group, younger monkeys are more cognitively flexible than older ones. The model is also able to tease out inter-individual differences in cognitive flexibility as demonstrated by our ability to cluster the animals post-hoc into

two groups that are consistent with the a priori-unknown performance level of the animals. We used simulation studies to validate these findings and conclude that the clustering is indeed reflective of the animals' performance. This suggests that the Laplace SMuRF model is a powerful tool for the classification of animals across the spectrum of performance in old age and provides a basis for evaluating individual differences.

Our analyses were facilitated by the derivation of powerful algorithms for estimation and inference in the Laplace SMuRF model. Compared to its Gaussian counterpart (Zhang et al., 2018) the Laplace prior explicitly models the abrupt change in performance at the onset of reversal learning. However, inference in the Laplace SMuRF is more expensive computationally. We derived a Monte Carlo EM algorithm for parameter estimation that uses an elegant Gibbs sampler. We also derived an efficient algorithm for MAP estimation that allows us to automatically capture abrupt changes (e.g., due to reversal learning).

Overall, the Laplace SMuRF model provides a simple, yet powerful, tool for the analysis of binary response data from behavioral experiments.

Acknowledgments

Demba Ba thanks the Alfred P. Sloan Foundation for its generous support. C.A. Barnes, D.T. Gray, S.N. Burke, and A.C. Smith thank the NIH (Grant R01 AG050548) and the McKnight Brain Research Foundation.

Appendix: Supporting derivations for the Methods section

Complete-data likelihood from augmented Laplace SMuRF model

We derive the expression for $\log p((\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{x}, \mathbf{d}; \theta)$, the log of the complete-data likelihood. The following derivations are parallel to those in the Appendix of Zhang et al. (2018). We sketch only the main steps here.

By definition of the augmented SMuRF model, the complete-data likelihood is

$$\begin{aligned} &p((\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{x}, \mathbf{d}, \mathbf{z}, (\mathbf{W}^m)_{m=1}^M; \theta) \\ &\quad = \prod_{m=1}^M p(\boldsymbol{\Delta}\mathbf{N}^m|\mathbf{x}_m, \mathbf{d}, \mathbf{z}_m, \mathbf{W}^m; \theta) p(\mathbf{x}_m|\mathbf{z}_m) p(\mathbf{z}_m) p(\mathbf{W}_m) p(\mathbf{d}). \end{aligned} \tag{18}$$

Therefore,

$$\begin{aligned} &\log p((\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{x}, \mathbf{d}, \mathbf{z}, (\mathbf{W}^m)_{m=1}^M; \theta) \\ &\quad = \sum_{m=1}^M \log p(\boldsymbol{\Delta}\mathbf{N}^m|\mathbf{x}_m, \mathbf{d}, \mathbf{z}_m, \mathbf{W}^m; \theta) \\ &\qquad + \log p(\mathbf{x}_m|\mathbf{z}_m) + \log p(\mathbf{z}_m) + \log p(\mathbf{W}_m) + \log p(\mathbf{d}) \\ &\quad = \sum_{m=1}^M KR_m \log(2) + R_m \log(\frac{\lambda^2}{2})) + \log p(\mathbf{x}_m|\mathbf{z}_m) + \log p(\mathbf{d}) \\ &\qquad + \sum_{m=1}^M \sum_{r=1}^{R_m} \sum_{k=1}^K \left\{ \Delta N_{k,r}^m (x_r^m + d_k) - \frac{x_r^m + d_k}{2} - \frac{(x_r^m + d_k)^2 w_{k,r}^m}{2} \right\} \\ &\qquad + \sum_{m=1}^M \sum_{r=1}^{R_m} \sum_{k=1}^K \left\{ \log \left(\sum_{i=1}^\infty (-1)^i \frac{(2i+1)}{\sqrt{2\pi {w_{k,r}^m}^2}} e^{-\frac{(2i+1)^2}{8 w_{k,r}^m}} \right) \right\} \\ &\qquad - \frac{\lambda^2}{2} \sum_{m=1}^M \sum_{r=1}^{R_m} z_r^m, \end{aligned} \tag{19}$$

where $\log p(\mathbf{x}_m|\mathbf{z}_m) = \sum_{r=1}^{R_m} - \frac{1}{2}\log(2\pi z_r^m) - \frac{(x_r^m - x_{r-1}^m)^2}{2{z_r^m}^2}$ and $\log p(\mathbf{d}) = \sum_{k=1}^K - \frac{1}{2}\log(2\pi\sigma_\delta^2) - \frac{(d_k - d_{k-1})^2}{2\sigma_\delta^2}$

Without loss of generality, we assume that $d_0 = 0$, and $x_0^m = 0$, $m = 1, \ldots, M$.

Derivation of full conditionals of learning states $(\boldsymbol{x}_m)_{m=1}^M$

From Eq. (19), we see that

$$\begin{aligned} &\log p\left(\mathbf{x}|(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{d}, \mathbf{z}, (\mathbf{W}^m)_{m=1}^M; \theta^{(\ell)}\right) \\ &\quad \propto \sum_{m=1}^M \sum_{r=1}^{R_m} \sum_{k=1}^K \left\{ \Delta N_{k,r}^m (x_r^m + d_k) - \frac{x_r^m + d_k}{2} - \frac{(x_r^m + d_k)^2 w_{k,r}^m}{2} \right\} \\ &\qquad + \log p(\mathbf{x}_m|\mathbf{z}_m) \\ &\quad = \sum_{m=1}^M \log p(\mathbf{x}_m|\boldsymbol{\Delta}\mathbf{N}^m, \mathbf{d}, \mathbf{z}_m, \mathbf{W}^m; \theta^{(\ell)}) \end{aligned} \tag{20}$$

Therefore, conditioned on $(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{d}, \mathbf{z}, (\mathbf{W}^m)_{m=1}^M$, $(\mathbf{x}_m)_{m=1}^M$ are independent. For sampling, we can restrict ourselves to a single m. By completing the square, we can simplify $\log p(\mathbf{x}_m|\boldsymbol{\Delta}\mathbf{N}^m, \mathbf{d}, \mathbf{z}_m, \mathbf{W}^m; \theta^{(\ell)})$ as follows

$$\begin{aligned} &\log p(\mathbf{x}_m|\boldsymbol{\Delta}\mathbf{N}^m, \mathbf{d}, \mathbf{z}_m, \mathbf{W}^m; \theta^{(\ell)}) \\ &\quad \propto \sum_{r=1}^{R_m} \sum_{k=1}^K \{ \Delta N_{k,r}^m (x_r^m + d_k) - \frac{x_r^m + d_k}{2} - \frac{(x_r^m + d_k)^2 w_{k,r}^m}{2} \} \\ &\qquad + \sum_{r=1}^{R_m} - \frac{1}{2}\log(2\pi z_r^m) - \frac{(x_r^m - x_{r-1}^m)^2}{2{z_r^m}^2} \\ &\quad \propto \sum_{r=1}^{R_m} \sum_{k=1}^K - \frac{1}{2\frac{1}{w_{k,r}^m}} \frac{(\Delta\tilde{N}_{k,r}^m - x_r^m + d_k)^2}{1/w_{k,r}^m} - \frac{(x_r^m - x_{r-1}^m)^2}{2{z_r^m}^2} \end{aligned} \tag{21}$$

It follows that sampling from $p(\mathbf{x}_m|\boldsymbol{\Delta}\mathbf{N}^m, \mathbf{d}, \mathbf{z}_m, \mathbf{W}^m; \theta^{(\ell)})$ is equivalent to sampling from the posterior of the state in the following linear-Gaussian state-space model

$$\begin{cases} x_r^m = x_{r-1}^m + \tilde{\epsilon}_r^m, \quad \tilde{\epsilon}_m^r \sim \mathcal{N}(0, z_r^m) \\ \Delta\tilde{N}_{k,r}^m = x_r^m + d_k + \tilde{v}_{k,r}, \quad \tilde{v}_{k,r} \sim \mathcal{N}(0, {w_{k,r}^m}^{-1}), \text{ i.i.d.}, \quad k = 1, \ldots, K \\ \Delta\tilde{N}_{k,r}^m = \Delta N_{k,r}^m - \frac{1}{2} \end{cases}$$

Derivation of full conditionals of group performance $(d_k)_{k=1}^K$

From Eq. (19), we see that

$$\begin{aligned} &\log p(\mathbf{d}|(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{x}, \mathbf{z}, (\mathbf{W}^m)_{m=1}^M; \theta^{(\ell)}) \\ &\quad \propto \sum_{m=1}^M \sum_{r=1}^{R_m} \sum_{k=1}^K \left\{ \Delta N_{k,r}^m (x_r^m + d_k) - \frac{x_r^m + d_k}{2} - \frac{(x_r^m + d_k)^2 w_{k,r}^m}{2} \right\} \\ &\qquad - \sum_{k=1}^K \frac{(d_k - d_{k-1})^2}{2\sigma_\delta^2} \\ &\quad \propto \sum_{k=1}^K \sum_{m=1}^M \sum_{r=1}^{R_m} - \frac{1}{2\frac{1}{w_{k,r}^m}} \frac{(\Delta\tilde{N}_{k,r}^m - x_r^m + d_k)^2}{1/w_{k,r}^m} - \frac{(d_k - d_{k-1})^2}{2\sigma_\delta^2} \end{aligned} \tag{22}$$

Therefore, sampling from $p(\mathbf{d}|(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{x}, \mathbf{z}, (\mathbf{W}^m)_{m=1}^M; \theta^{(\ell)})$ is equivalent to sampling from the posterior of the state in the following linear-Gaussian state-space model

$$\begin{cases} d_k = d_{k-1} + \delta_k, \quad \delta_k \sim \mathcal{N}(0, \sigma_\delta^2) \\ \Delta\tilde{N}_{k,r}^m = x_r^m + d_k + \tilde{v}_{k,r}, \quad \tilde{v}_{k,r} \sim \mathcal{N}(0, w_{k,r}^{m\ -1}), \quad \text{i.i.d.}, \quad m = 1, \ldots, M, \quad r = 1, \ldots, R_m \\ \Delta\tilde{N}_{k,r}^m = \Delta N_{k,r}^m - \frac{1}{2} \end{cases}$$

This form leads to an efficient forward-filtering backward-sampling algorithm (see Zhang et al., 2018).

M-step of EM for augmented SMuRF model

In the M-step, we obtain new values $\lambda^{(\ell+1)}$ and $\sigma_\delta^{2(\ell+1)}$ by maximizing $\mathcal{Q}(\theta|\theta^{(\ell)})$. We first derive the update equation for $\lambda^{(\ell+1)}$. From Eq. (19), taking terms in λ, and writing $\alpha = \frac{\lambda^2}{2}$,

$$\alpha^{(\ell+1)} = \arg\max_\alpha \sum_{m=1}^M R_m \log\alpha - \alpha \sum_{r=1}^{R_m} \mathbb{E}[z_r^m | x_r^m, x_{r-1}^m; \lambda^{(\ell)}]. \tag{23}$$

Taking the derivative of the objective with respect to α and setting equal to zero yields

$$\lambda^{(\ell+1)} = \left(\frac{2\sum_{m=1}^M R_m}{\sum_{m=1}^M \sum_{r=1}^{R_m} \mathbb{E}[z_r^m | x_r^m, x_{r-1}^m; \lambda^{(\ell)}]} \right)^{1/2}, \tag{24}$$

where we approximate the expectations in the denominator from the Monte-Carlo samples computed in the E-step.

Similarly,

$$\begin{aligned} \sigma_\delta^{2(\ell+1)} = & \arg\max_{\sigma_\delta^2} K \log(2\pi\sigma_\delta^2) \\ & - \frac{1}{2\sigma_\delta^2} \sum_{k=1}^K \mathbb{E}[(d_k - d_{k-1})^2 | (\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{x}, \mathbf{z}, (\mathbf{W}^m)_{m=1}^M; \theta^{(\ell)}], \end{aligned} \tag{25}$$

which yields

$$\sigma_\delta^{2(\ell+1)} = \frac{1}{K} \sum_{k=1}^K \mathbb{E}[(d_k - d_{k-1})^2 | (\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M, \mathbf{x}, \mathbf{z}, (\mathbf{W}^m)_{m=1}^M; \theta^{(\ell)}]. \tag{26}$$

Derivation of algorithm for computing MAP estimates of $\boldsymbol{d}$ *and* $(\boldsymbol{x}_m)_{m=1}^M$

The Monte-Carlo EM iterative algorithm for maximizing the likelihood of the SMuRF model results in estimates of the model parameters σ_δ^2 and λ. At convergence of the EM algorithm, samples from the joint posterior $p(\mathbf{x}, \mathbf{d}|(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M; \theta^{(\infty)})$ are also available. From these samples, we can approximate the posterior means of $\mathbf{x}$ and $\mathbf{d}$. The use of the Laplace distribution in the SMuRF model allows us to automatically capture abrupt changes, due to reversal learning, in the learning states of the animals. This, however, is a property of the mode of the posterior distribution and not necessarily its mean. Indeed, due to the potentially complex nature of the posterior distribution, its mean and its mode may not coincide. Computing the mode of the posterior distribution is very challenging, which motivates the iterative algorithm we derive below. In what follows, we denote the posterior by $p(\mathbf{x}, \mathbf{d}|(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M)$, leaving implicit its dependence on σ_δ^2 and λ.

Finding the maximum of the posterior distribution corresponds to solving the following optimization problem:

$$\begin{aligned} \max_{\mathbf{x},\mathbf{d}} \log(p(\mathbf{x}, \mathbf{d}|(\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M)) \propto & \sum_{m=1}^M \sum_{r=1}^{R_m} \sum_{k=1}^K \Delta N_{k,r}^m (x_r^m + d_k) - \log(1 + e^{x_r^m + d_k}) \\ & - \lambda \sum_{m=1}^M \sum_{r=1}^R |x_r^m - x_{r-1}^m| - \frac{1}{2\sigma^2} \sum_{k=1}^K (d_k - d_{k-1})^2 \end{aligned} \tag{27}$$

Invoking Eq. (19), one can see that

$$\begin{aligned} \log p\left(\mathbf{x}, \mathbf{d}, \mathbf{z}, (\mathbf{W}^m)_{m=1}^M | (\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M\right) \propto & \sum_{m=1}^M \sum_{r=1}^{R_m} \sum_{k=1}^K - \frac{(\Delta\tilde{N}_{k,r}^m - (x_r^m + d_k))^2}{2/w_{k,r}^m} \\ & - \sum_{k=1}^K \frac{(d_k - d_{k-1})^2}{2\sigma^2} - \sum_{m=1}^M \sum_{r=1}^R \frac{(x_r^m - x_{r-1}^m)^2}{2z_r^m}, \end{aligned} \tag{28}$$

with $\Delta\tilde{N}_{k,r}^m = \frac{\Delta N_{k,r}^m - 1/2}{w_{k,r}^m}$.

It is possible to show that

$$\mathbb{E}_{\mathbf{z},(\mathbf{W}^m)_{m=1}^M}[\log p(\mathbf{x}, \mathbf{d}, \mathbf{z}, (\mathbf{W}^m)_{m=1}^M | (\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M), \mathbf{x}^{(n)}, \mathbf{d}^{(n)}]$$

is a majorizer (Hunter and Lange, 2004) of the objective function in Eq. (27). That is, if $(\mathbf{x}^{(n)}, \mathbf{d}^{(n)})$ is a candidate solution for the maximizer of Eq. (27), we can generate a new candidate solution $(\mathbf{x}^{(n)}, \mathbf{d}^{(n)})$ by solving

$$\max_{\mathbf{x},\mathbf{d}} \mathbb{E}_{\mathbf{z},(\mathbf{W}^m)_{m=1}^M}[\log p(\mathbf{x}, \mathbf{d}, \mathbf{z}, (\mathbf{W}^m)_{m=1}^M | (\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M), \mathbf{x}^{(n)}, \mathbf{d}^{(n)}]. \tag{29}$$

After some algebra,

$$\begin{aligned} \mathbb{E}_{\mathbf{z},(\mathbf{W}^m)_{m=1}^M} & [\log p(\mathbf{x}, \mathbf{d}, \mathbf{z}, (\mathbf{W}^m)_{m=1}^M | (\boldsymbol{\Delta}\mathbf{N}^m)_{m=1}^M), \mathbf{x}^{(n)}, \mathbf{d}^{(n)}] \\ & \propto \sum_{m=1}^M \sum_{r=1}^{R_m} \sum_{k=1}^K - \frac{(\mathbb{E}_{w_{k,r}^m}[\Delta\tilde{N}_{k,r}^m | x_r^{m(n)}, d_k^{(n)}] - (x_r^m + d_k))^2}{2} \mathbb{E}[w_{k,r}^m | x_r^{m(n)}, d_k^{(n)}] \\ & - \sum_{k=1}^K \frac{(d_k - d_{k-1})^2}{2\sigma^2} - \sum_{m=1}^M \sum_{r=1}^R \frac{(x_r^m - x_{r-1}^m)^2}{2} \mathbb{E}[1/z_r^m | x_r^{m(n)}, x_{r-1}^{m(n)}]. \end{aligned} \tag{30}$$

We first define

$$\mathbb{E}_{w_{k,r}^m}[\Delta\tilde{N}_{k,r}^m | x_r^{m(n)}, d_k^{(n)}] \stackrel{d}{=} \frac{\Delta N_{k,r}^m - \frac{1}{2}}{\mathbb{E}[w_{k,r}^m | x_r^{m(n)}, d_k^{(n)}]} \tag{31}$$

and note that (Polson et al., 2013; Eltoft et al., 2006)

$$\mathbb{E}[w_{k,r}^m | x_r^{m(n)}, d_k^{(n)}] = \frac{1}{2(x_r^{m(n)} + d_k^{(n)})} \tanh\left(\frac{(x_r^{m(n)} + d_k^{(n)})}{2}\right) \tag{32}$$

and

$$\mathbb{E}[1/z_r^m | x_r^{m(n)}, x_{r-1}^{m(n)}] = \frac{\lambda}{|x_r^{m(n)} - x_{r-1}^{m(n)}|} \tag{33}$$

We can maximize this function by alternative maximization. First, given $(\mathbf{x}^{(n)}, \mathbf{d}^{(n)})$, we fix $\mathbf{d}^{(n)}$, and find $\mathbf{x}^{(n+1)}$ by maximizing:

$$\max_{\mathbf{x}} \sum_{m=1}^{M}\sum_{r=1}^{R_m}\sum_{k=1}^{K} - \frac{\left(\mathbb{E}_{w_{k,r}^m}[\Delta\tilde{N}_{k,r}^m | x_r^{m(n)}, d_k^{(n)}] - (x_r^m + d_k^{(n)})\right)^2}{2}\mathbb{E}[w_{k,r}^m | x_r^{m(n)}, d_k^{(n)}] - \sum_{m=1}^{M}\sum_{r=1}^{R}\frac{(x_r^m - x_{r-1}^m)^2}{2}\mathbb{E}[1/z_r^m | x_r^{m(n)}, x_{r-1}^{m(n)}]. \quad (34)$$

Using an argument similar to the derivation of the full conditional of the learning states $(\mathbf{x}_m)_{m=1}^M$ in the Gibbs sampler, we can solve this problem in parallel (across animals) using a Kalman smoother for each animal.

Lastly, given $\mathbf{x}^{(n+1)}$, we solve for $\mathbf{d}^{(n+1)}$ by computing

$$\max_{\mathbf{d}} \sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{r=1}^{R_m} - \frac{\left(\mathbb{E}_{w_{k,r}^m}[\Delta\tilde{N}_{k,r}^m | x_r^{m(n+1)}, d_k^{(n)}] - (x_r^{m(n+1)} + d_k)\right)^2}{2}\mathbb{E}[w_{k,r}^m | x_r^{m(n+1)}, d_k^{(n)}] - \sum_{k=1}^{K}\frac{(d_k - d_{k-1})^2}{2\sigma^2}. \quad (35)$$

Using an argument similar to the derivation of the full conditional of group performance $\mathbf{d}$ in the Gibbs sampler, we can solve this problem using a Kalman smoother.

In this manner, we can generate a sequence $\{(\mathbf{x}^{(n)}, \mathbf{d}^{(n)})\}_{n=1}^{\infty}$ that is provably convergent to the maximizer of the posterior (Eq. (27)). We omit the proof here (see Hunter and Lange, 2004).

References

Akaike, H., 1998. Information theory and an extension of the maximum likelihood principle. In: Selected Papers of Hirotugu Akaike, Springer, pp. 199–213.

Burke, S.N., Thome, A., Plange, K., Engle, J.R., Trouard, T.P., Gothard, K.M., Barnes, C.A., 2014. Orbitofrontal cortex volume in area 11/13 predicts reward devaluation, but not reversal learning performance, in young and aged monkeys. J. Neurosci. 34 (30), 9905–9916, http://dx.doi.org/10.1523/JNEUROSCI.3918-13.2014, arXiv:http://www.jneurosci.org/content/34/30/9905.full.pdf, http://www.jneurosci.org/content/34/30/9905.

Casella, G., George, E., 1992. Explaining the Gibbs sampler. Am. Stat. 46 (3), 167–174.Czanner, G., Eden, U., Wirth, S., Yanike, M., Suzuki, W., Brown, E., 2008. An analysis of between-trial and within-trial neural spiking dynamics. J. Neurophysiol. 99 (5), 2672–2693.

Dempster, A., Laird, N., Rubin, D., 1977. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. B (Methodol.), 1–38.

Eldar, E., Morris, G., Niv, Y., 2011. The effects of motivation on response rate: a hidden semi-Markov model analysis of behavioral dynamics. J. Neurosci. Methods 201 (1), 251–261, http://dx.doi.org/10.1016/j.jneumeth.2011.06.028.

Eltoft, T., Kim, T., Lee, T., 2006. On the multivariate Laplace distribution. IEEE Signal Process. Lett. 13 (5), 300–303.

Gray, D., Smith, A., Burke, S., Gazzaley, A., Barnes, C., 2016. Attentional updating and monitoring and affective shifting are impacted independently by aging in macaque monkeys. Behav. Brain Res., http://dx.doi.org/10.1016/j.bbr.2016.06.056 http://www.sciencedirect.com/science/article/pii/S016643281630420X.

Hunter, D.R., Lange, K., 2004. A tutorial on MM algorithms. Am. Stat. 58 (1), 30–37.

Kim, S., Frank, L., 2009. Hippocampal lesions impair rapid learning of a continuous spatial alternation task. PLoS ONE 4 (5), e5494, http://dx.doi.org/10.1371/journal.pone.0005494.

D. Müllner, Modern hierarchical, agglomerative clustering algorithms, arXiv:1109.2378.

Polson, N., Scott, J., Windle, J., 2013. Bayesian inference for logistic models using pólya-gamma latent variables. J. Am. Stat. Assoc. 108 (504), 1339–1349.

Rad, K.R., Paninski, L., 2010. Efficient, adaptive estimation of two-dimensional firing rate surfaces via Gaussian process methods. Network Comput. Neural Syst. 21 (3–4), 142–168.

Rapp, P.R., Amaral, D.G., 1989. Evidence for task-dependent memory dysfunction in the aged monkey. J. Neurosci. 9 (10), 3568–3576.

Roberts, J.M., Ly, M., Murray, E., Yassa, M.A., 2014. Temporal discrimination deficits as a function of lag interference in older adults. Hippocampus 24 (10), 1189–1196, http://dx.doi.org/10.1002/hipo.22303.

Shamy, J.L., Habeck, C., Hof, P.R., Amaral, D.G., Fong, S.G., Buonocore, M.H., Stern, Y., Barnes, C.A., Rapp, P.R., 2010. Volumetric correlates of spatiotemporal working and recognition memory impairment in aged rhesus monkeys. Cereb. Cortex 21 (7), 1559–1573.

Siegel, S., Castellan, N.J., 1988. Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill, New York.

Smith, A., Frank, L., Wirth, S., Yanike, M., Hu, D., Kubota, Y., Graybiel, A., Suzuki, W., Brown, E., 2004. Dynamic analysis of learning in behavioral experiments. J. Neurosci. 24, 447–461.

Smith, A., Stefani, M., Moghaddam, B., Brown, E., 2005. Analysis and design of behavioral experiments to characterize population learning. J. Neurophysiol. 93, 1776–1792.

Smith, A., Wirth, S., Suzuki, W., Brown, E., 2007. Bayesian analysis of interleaved learning and response bias in behavioral experiments. J. Neurophysiol. 97, 2516–2524.

Spiegel, A.M., Koh, M.T., Vogt, N.M., Rapp, P.R., Gallagher, M., 2013. Hilar interneuron vulnerability distinguishes aged rats with memory impairment. J. Comp. Neurol. 521 (15), 3508–3523.

Stefani, M., Groth, K., Moghaddam, B., 2003. Glutamate receptors in the rat medial prefrontal cortex regulate set-shifting ability. Behav. Neurosci. 117 (4), 728–737, http://dx.doi.org/10.1037/0735-7044.117.4.728.

Stone, M., 1977. An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. J. R. Stat. Soc. Ser. B (Methodol.) 39 (1), 44–47.

Vehtari, A., Gelman, A., Gabry, J., 2017. Practical bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat. Comput. 27 (5), 1413–1432.

Wise, S., Murray, E., 1999. Role of the hippocampal system in conditional motor learning: mapping antecedents to action. Hippocampus 9 (2), 101–117, http://dx.doi.org/10.1002/(SICI)1098-1063(1999)9:2<101::AID-HIPO3>3.0.CO;2-L.

Zhang, Y., Malem-Shinitski, N., Allsop, S.A., Tye, K.M., Ba, D., 2018. Estimating a separably Markov random field from binary observations. Neural Comput. 30 (4), 1046–1079, http://dx.doi.org/10.1162/neco_a_01059.