Received: 28 October 2019 | Revised: 19 May 2020 | Accepted: 1 June 2020

DOI: 10.1111/ejn.14860

RESEARCH REPORT

EJN European Journal of Neuroscience FENS WILEY

Neural networks of different species, brain areas and states can be characterized by the probability polling state

Zhi-Qin John Xu[1] | **Xiaowei Gu**[2,3] | **Chengyu Li**[2] | **David Cai**[1,4] | **Douglas Zhou**[1] | **David W. McLaughlin**[4]

[1]School of Mathematical Sciences, MOE-LSC and Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai, China

[2]CAS Center for Excellence in Brain Science and Intelligence Technology, Institute of Neuroscience, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China

[3]University of Chinese Academy of Sciences, Beijing, China

[4]Courant Institute of Mathematical Sciences and Center for Neural Science, New York University, New York, NY, USA

Correspondence
Douglas Zhou, School of Mathematics & Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai 200240, China.
Email:zdz@sjtu.edu.cn

and
David W. McLaughlin, Courant Institute, New York University, New York, New York 10012, USA.
Email: dwm1@nyu.edu

Funding information
National Key R&D Program of China, Grant/Award Number: 2019YFA0709503; National Science Foundation in China, Grant/Award Number: 11671259 and 11722107; Student Innovation Center at Shanghai Jiao Tong University; SJTU-UM Collaborative Research Program

Abstract

Cortical networks are complex systems of a great many interconnected neurons that operate from collective dynamical states. To understand how cortical neural networks function, it is important to identify their common dynamical operating states from the probabilistic viewpoint. Probabilistic characteristics of these operating states often underlie network functions. Here, using multi-electrode data from three separate experiments, we identify and characterize a cortical operating state (the "probability polling" or "p-polling" state), common across mouse and monkey with different behaviors. If the interaction among neurons is weak, the p-polling state provides a quantitative understanding of how the high dimensional probability distribution of firing patterns can be obtained by the low-order maximum entropy formulation, effectively utilizing a low dimensional stimulus-coding structure. These results show evidence for generality of the p-polling state and in certain situations its advantage of providing a mathematical validation for the low-order maximum entropy principle as a coding strategy.

KEYWORDS

coding, cortical state, maximum entropy, network, neuroscience

Abbreviations: MEP, maximum entropy principle; NFP, neural firing pattern; p-polling, probability polling.

Edited by John Foxe

The peer review history for this article is available at https://publons.com/publon/10.1111/ejn.14860

1 | INTRODUCTION

Recent technological advances are providing massive and intricate datasets of brain activity, from the scale of a single neuron to that of the whole brain (Nguyen et al., 2016). In neuroscience, a well-known framework to formalize what it means to understand the brain is David Marr's three-level understanding of a complex system (Marr, 1982): (a) What is the goal of the computation? (b) What representations of the input and output and what algorithms are employed to carry out this computation? (c) How can the representations and algorithms be realized physically or bio-physically. In this work, we identify a commonly occurring operating state, and take steps toward understanding how neural networks, in this operating state, represent information and realize computation, that is, Marr's second level. For the microprocessor, the fundamental computational rule in different-level computations is similar, for example, summation or subtraction through arithmetic logic units. Neural networks share similar network motifs (Alon, 2007), and individual neurons in neural networks can be characterized by similar dynamics (Gerstner & Kistler, 2002), e.g., the Hodgkin-Huxley (HH) type dynamics (Hodgkin & Huxley, 1952). Therefore, it is natural and important to explore whether, in brain networks, there exists a cortical dynamical state, common across different species, brain areas, and behavioral tasks, which can be characterized quantitatively and which can underlie basic neural computations and general population coding schemes of neural networks.

Both theoretical and experimental evidence indicate that populations of neurons perform computations probabilistically. An important example is the "Bayesian coding hypothesis": that the brain represents sensory information through probability distributions (Knill & Pouget, 2004). Probabilistic dependencies among neurons in the population—where neurons often cooperate to perform functions—underlie how information is encoded by probability distributions of a neural network's states. For example, Pearson correlation is a widely used quantity to characterize probabilistic dependency. Theoretical studies have shown that different Pearson correlation structures in neural dynamics can result in very different coding schemes (Abbott & Dayan, 1999; Averbeck, Latham, & Pouget, 2006; Moreno-Bote et al., 2014). In this work, we characterize the dependence between neurons probabilistically. Consider any two neurons in a network, denoted by i and j, respectively. Their dependence is characterized by the *difference* of the probability of neuron i firing under two different conditions—one with all other neurons silent, and the other with only neuron j firing spikes. Using a *probability increment* to characterize the dependence between neurons, we have quantitatively identified a stochastic operating state of the network, [referred to herein as the *"probability polling" (p-polling) state*]. This state defines a basic computational rule of neural networks on the probabilistic level. By analyzing in vivo multi-electrode data from three separate published experiments (Coen-Cagli, Kohn, & Schwartz, 2015; Kohn & Coen-Cagli, 2015; Kohn & Smith, 2016; Liu et al., 2014; see Section 2), we show that the p-polling state is a commonly occurring cortical operating state for mouse and monkey with performing different functions.

In the probabilistic view point, the statistical distribution of neural firing patterns (NFPs) encodes the information of external stimuli. When the cortical network is in the p-polling state and the probability increment δ is small, we have theoretically established (Xu, Bi, Zhou, & Cai, 2017) that the distribution of NFPs can be represented by a *low-order maximum entropy principle* (MEP) distribution. Here, for in vivo p-polling cortical operating points, *second-order* MEP distributions of NFPs, with constraints of mean firing for each neuron and second-order correlations of firing between each pair of neurons, accurately represent the observed distributions of NFPs, while first-order MEP distributions are not accurate. This good performance of second-order MEP distributions is consistent with other findings (Schneidman, Berry, Segev, & Bialek, 2006; Shlens et al., 2006; Tang et al., 2008; Xu et al., 2017). In addition to characterizing distributions, using the second-order MEP analysis of 78 neurons on CA1 in mice as they run along a virtual linear track (Meshulam, Gauthier, Brody, Tank, & Bialek, 2017), one study found that place cells and non-place cells are coupled together, that is, they are not distinct sub-networks, but a single network that can encode more than just place information (Meshulam et al., 2017). In the p-polling state, one can show that interaction strengths in the MEP model form a hierarchy in the power of probability increment. When the neural network is in the p-polling state and the probability increment δ is small, the neural coding scheme can be well described by the MEP framework. However, when the probability increment δ is not small, the system might still be in the p-polling state. In such a case, the MEP framework cannot well characterize the neural activity. Note that when the system is in the p-polling state, many constraints are imposed to the high-dimensional probability distribution of neural firing patterns (due to the linear superposition rule among conditional probability increments as discussed below), that is, the p-polling state may greatly reduce the degree of freedoms in NFPs and constrain the corresponding probability distribution to a low-dimensional manifold. However, the computation of this manifold is beyond the MEP framework and requires future investigation by taking advantage of certain prior knowledge or further dynamical properties.

2 | MATERIALS AND METHODS

2.1 | MEP analysis

We recapitulate the MEP analysis as follows and more details can be found in references (Schneidman et al., 2006;

Shlens et al., 2006; Tang et al., 2008). In a sampling time bin, a neuron has two states, that is, firing and silence. For a network of n neurons, these measurements will be represented by a binary vector $V(t)=(\sigma_1,\sigma_2,\ldots,\sigma_n)\in\{0,\ 1\}^n$, where $\sigma_i=\sigma_i(t)=0,\ 1$ represents the ith neuron being silent, or firing a spike within the time bin labeled by t, respectively. For a fixed sampling time bin, we partition the recording time by the bin size and record the state of neurons at every sampling time bin. Here, we use a time bin of 20 ms, which is widely used in experimental measurements (Schneidman et al., 2006; Tang et al., 2008). For any possible firing pattern V, we count the occurring frequency of V in all sampled bins. Then, the observed distribution of neural firing patterns, $P(V)$, is the frequency distribution of all possible states. To obtain correlations up to the mth order, it is required to calculate all possible $\langle\sigma_{i_1},\ldots,\sigma_{i_M}\rangle$, where $1\le i_1\le\ldots\le i_M\le n$, $1\le M\le m$, and $\langle\cdot\rangle_E$ is defined by $\langle g(l)\rangle_E=\sum_{l=1}^{N_L}g(l)/N_L$ for any function $g(l)$, and N_L is the total number of sampling time bins in the recording. The mth order MEP analysis is to find the desired probability distribution $P_m(V)$ for n neurons by maximizing the entropy $S\equiv-\sum_V P_m(V)\log P_m(V)$, subject to the constraints of correlations up to the mth order ($m\le n$). By using the method of Lagrange multipliers, the distribution is of the form

$$P_m(V)=\frac{1}{Z}\exp\left(\sum_{k=1}^{m}\sum_{i_1\le\ldots\le i_k}^{n}J_{i_1\ldots i_k}\sigma_{i_1}\ldots\sigma_{i_k}\right)\qquad(1)$$

where $J_{i_1\ldots i_k}$ is the kth order interaction ($1\le k\le m$), and the "partition function" Z is the normalization factor. Equation (1) is referred to as the mth order MEP distribution. For P_m with $m<n$, we use the same iterative scaling algorithm as in Tang et al. (2008) to numerically solve the above optimization problem to obtain the interaction parameters and the corresponding distribution (see Section 2.2 below for details).

Note that P_n, which will be referred to as the full-order MEP distribution, is identical to the observed distribution $P(V)$ (Amari, 2001; Xu, Zhou, & Cai, 2019). Then, by using $P(V)$, we can solve analytically for interactions of P_n. For illustration, considering a network of size $n=2$ and substituting 2^2 states of $V=(\sigma_1,\sigma_2)$ and $P(V)$ into Equation (1) with $m=n$, we have $P[V=(0,0)]=1/Z$, $P[V=(1,0)]=\exp(J_1)/Z$, $P[V=(0,1)]=\exp(J_2)/Z$, and $P[V=(1,1)]=\exp(J_1+J_2+J_{12})/Z$. The left side of above equations can be measured from experimental data. By taking the logarithm of both sides of above equations, we can have linear equations for $\log Z$, J_1, J_2, J_{12}, e.g., $\log P[V=(0,0)]+\log Z=0$ and $\log P[V=(1,0)]+\log Z=J_1$. Four unknowns, i.e., $\log Z, J_1, J_2, J_{12}$, in the above linear equations can then be found by solving this linear system. Similarly, for any n, substituting 2^n states of $V=(\sigma_1,\sigma_2,\ldots,\sigma_n)$ and the observed probability distribution $P(V)$ into the nth order MEP analysis [Equation (1) with $m=n$], then taking the logarithm of both sides of Equation (1), we obtain a system of 2^n linear equations with total number of 2^n unknowns (2^n-1 interactions and one normalization factor),

$$\sum_{k=1}^{m}\sum_{i_1\le\ldots\le i_k}^{n}J_{i_1\ldots i_k}\sigma_{i_1}\ldots\sigma_{i_k}=\log P(V)+\log Z,\qquad(2)$$

By solving the system of linear equations [Equation (2)], we can obtain the 2^n-1 interactions J's for the nth order MEP analysis. For example, $n=3$, we can obtain

$$J_{12}=\log\frac{P(\sigma_1=1|\sigma_2=1,\sigma_3=0)}{P(\sigma_1=0|\sigma_2=1,\sigma_3=0)}-\log\frac{P(\sigma_1=1|\sigma_2=0,\sigma_3=0)}{P(\sigma_1=0|\sigma_2=0,\sigma_3=0)}$$

Note that J_{12} can be interpreted as the increment of the first-order interaction $J_1=\log(P_{100}/P_{000})$ induced by the second neuron's firing, where $P_{\sigma_1\sigma_2\sigma_3}$ represents the network state with σ_i being the state of the ith neuron. Namely, we first switch the state of the second neuron in J_1 from silence to firing, to produce $\log(P_{110}/P_{010})$, which is denoted as J^1. Then, the second-order interaction is $J_{12}=J^1-J_1$. This recursive structure can be extended to the case of higher-orders, *that is*, the $(k+1)$th order interaction $J_{123\ldots(k+1)}$ can be obtained as follows (Xu et al., 2017): First, we switch the state of the $(k+1)$th neuron in $J_{123\ldots k}$ from silence to firing to obtain a new term J^1. Then, we subtract $J_{123\ldots k}$ from the new term to obtain $J_{123\ldots(k+1)}$, that is, $J_{123\ldots(k+1)}=J^1-J_{123\ldots k}$. In the p-polling state, the recursive relation implies that the interaction strengths form a hierarchy in the power of probability increment (Xu et al., 2017). For illustration, in a homogeneous network, *that is*, the probability increment $\delta_i=\delta$ for $1\le i\le n$, where δ_i is defined in Equation (5); then, $J_{123\ldots k}\sim O(\delta^{k-1})$.

2.2 | The iterative scaling algorithm

We use the same numerical algorithm as in Tang et al. (2008) to estimate the interactions of the MEP distribution in Equation (1) when $m<n$. For illustration, we show the process for J_i for the second-order MEP analysis, $P_2(V)$. The value of J_i is adjusted by an iterative procedure: $J_i^{\text{new}}=J_i^{\text{old}}+\alpha\,\text{sign}(\langle\sigma_i\rangle_E)\log(\langle\sigma_i\rangle_E/\langle\sigma_i\rangle_2)$, where $\langle\sigma_i\rangle_2\equiv\sum_{l=1}^{2^n}\sigma_i(V_l)P_2(V_l)$, the constant α is used to maintain the stability of the iteration. We use $\alpha=.75$ as in Tang et al. (2008). Adjustments are performed for each ensemble until the difference of the expected and the predicted values (means and pairwise correlations) are smaller than the tolerance 10^{-11}.

2.3 | Overview of experimental settings

Multi-electrode recording data from three separate experiments used in this work are summarized as follows. We select neurons with firing rates larger than 6 Hz to perform our

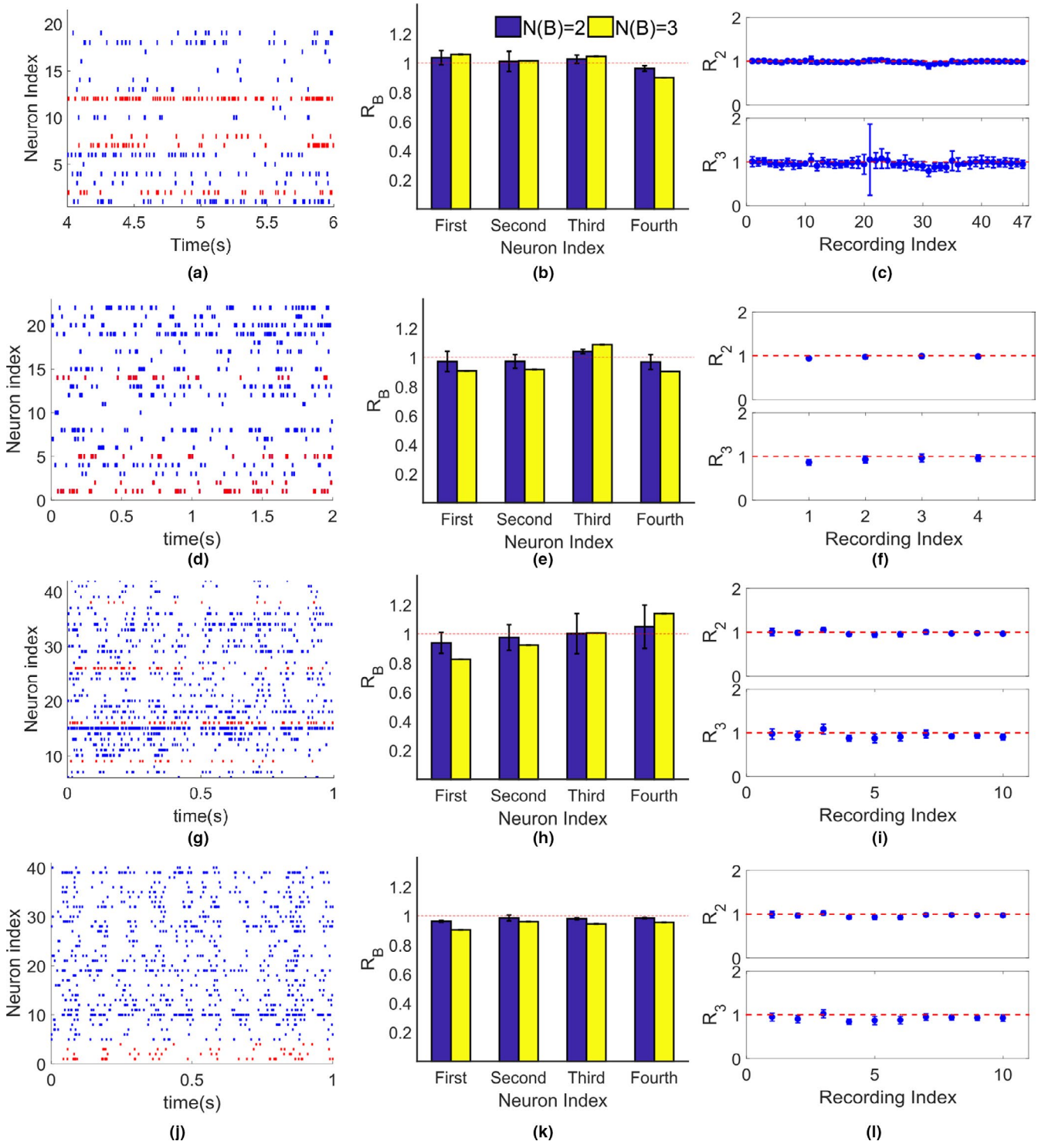

FIGURE 1 Linearity indexes of four-neuron sub-networks. The analysis of four datasets (see the main text) is displayed. (a, d, g, j): Raster plot with every short bar indicates that the neuron with certain index fires at certain time. For figures (a, d, g, j), we select four neurons in each of them which are denoted by red color. The linearity indexes R_B are computed from these selected four neurons correspondingly and results are illustrated in figures (b, e, h, k). Note that, according to the definition of R_B, subsets can have different sizes, we here choose $N(B) = 2$ and $N(B) = 3$ to compute R_B for each neuron and display them in blue bar and yellow bar, respectively. (c, f, i, l): Each dot indicates a total mean (see Section 2 for details) of one recording for $N(B) = 2$ (upper panel) and $N(B) = 3$ (lower panel), respectively. The red dashed line is $y = 1$ and the abscissa indicates the recording index for each dataset (see the main text) [Colour figure can be viewed at wileyonlinelibrary.com]

data analysis. To show that the NFP distributions are stable throughout the recording, for illustrations, we split the recording of the four selected neurons in Figure 1a,d,g,j into two parts. The NFP distributions obtained from each part are consistent with the distribution obtained from the total recording (Figure S1).

2.4 | Olfactory delayed non-match to sample task

The first experiment provides data for the first dataset. The data analyzed here were collected from a published study (Liu et al., 2014) and the details can be found therein. Head-fixed mice were trained to perform an olfactory delayed non-match to sample task. For each trial, an olfactory stimulus (ethyl acetate, EA, or 2-pentanone, 2P) was presented as the sample, followed by a delay period (5 s) and then a testing olfactory stimulus, either matched or non-matched to the sample. Water-restricted mice were rewarded with water if they licked within a response time window in the non-match but not match trials. The mouse performed tasks continuously with a 10 s rest between consecutive trials. The experiment is designed to separate the working memory period from the decision-making period. Here, we only focus on the data structure in this experiment, which was single-unit activity of medial prefrontal cortex (mPFC) measured with custom-made tetrodes. For each mouse in each recording, the spike trains of a group of neurons were simultaneously recorded from the mPFC of the mouse for 203 consecutive trials (3,756 s duration length in total). There are 47 recordings of 8 mice in total.

2.5 | Spontaneous firing in V1 in anesthetized macaque monkeys (pvc-11)

The second experiment provides data for the second dataset. The data were downloaded from the CRCNS website crcns.org (pvc-11; Kohn & Smith, 2016). The data used in this work is spontaneous V1 activity for anesthetized macaque monkeys. Sufentanil (4 – 18 microg kg^{-1} hr^{-1}) was used for anesthesia. The data were recorded by the Laboratory of Adam Kohn at the Albert Einstein College of Medicine and J. Anthony Movshon's laboratory at New York University. The spontaneous datasets that comprise spiking activity from ~70 to 100 neurons are recorded using "Utah" arrays placed in visual cortex of four adult Macaca fascicularis monkeys (a total number of six arrays). Recordings were obtained while a uniform gray screen was shown to the animals. Four recordings were obtained from V1 in four monkeys with the time duration of 1,234 s, 903 s, 1,874 s, 1,841 s, respectively (Kohn & Smith, 2016).

2.6 | V1 in anesthetized macaque monkeys under images (pvc-8)

The third experiment provides data for both the third and fourth datasets. The data were collected in the Laboratory of Adam Kohn at the Albert Einstein College of Medicine and downloaded from the CRCNS web site (Kohn & Coen-Cagli, 2015). These data consist of multi-electrode recordings from V1 in anesthetized macaque monkeys, while natural images and gratings were flashed on the screen. Natural images were presented at two sizes, 3–6.7 degrees and windowed to 1 degree, to quantify surround modulation. Stimuli were presented monocularly in a circular aperture surrounded by a gray field of average luminance. Recordings were performed using the "Utah" electrode array. Experimental procedures and stimuli are fully described in Coen-Cagli et al. (2015). Data from ten recordings are included. The data were recorded from three animals (recording 01 from animal 1; recording 02–07 from animal 2; recording 08–10 from animal 3).

Each image stimulus (number of 956 in total) was presented in pseudo-random order for 100 ms, followed by a 200 ms uniform gray screen. Each stimulus was presented 20 times. We separate each recording into two parts, and treat the data recorded during uniform gray screen stimulus as one recording of the third dataset, and the data recorded during the image stimuli as one recording of the fourth dataset. The measured distributions of neural firing patterns of the third dataset are very different from those of the fourth dataset, as shown in Figure S2, justifying the network residing in two distinct operating points. The recording time of each recording is 3,824 s and 1,912 s for the third and the fourth dataset, respectively. Except for the first recording, since there are many neurons firing larger than 6 Hz, we randomly select 1,000 four-neuron sub-networks in each recording for the MEP analysis.

2.7 | Remarks on analysis details

We use the same bin size of 20 ms as used in experimental studies (Schneidman et al., 2006; Tang et al., 2008) to analyze the NFP data; however, the conclusions are still valid for reasonable choices of other bin sizes, for example, 10 and 40 ms (data not shown). The time scale of the auto-correlation for the spike trains of datasets used in this work is smaller than 20 ms. For example, the auto-correlation of a neuron in Figure 1a decays to 0.01 when the time lag is more than 10 ms.

The existence of the p-polling state is neither dependent on the threshold choice of firing rate (*e.g.*, 6 Hz) that is chosen to insure sufficient number of spikes for estimating distributions, nor dependent on the MEP analysis that is used for estimating the probability distribution of NFPs.

2.8 | Synchronization index

We use the following synchronization index (SI) to characterize the synchronization of a network of n neurons. The

number of firing neurons in the *i*th bin is denoted by n_i and the total number of the bins which have at least one neuron firing is denoted by L. The SI is defined as =1nL∑i=1Lni,

$$\mathrm{SI} = \frac{1}{nL}\sum_{i=1}^{L} n_i, \tag{3}$$

which measures the mean fraction of active neurons in the bins which have at least one neuron firing. When SI = 1, the network is in a fully synchronized state.

2.9 | The p-polling state

To define the p-polling state, we start with the definition of probability increment which characterizes the probabilistic dependence between neurons. The state of a network of n neurons is denoted by a binary vector $V(t) = (\sigma_1, \sigma_2, \ldots, \sigma_n) \in \{0, 1\}^n$ where $\sigma_i = \sigma_i(t) = 0, 1$ represents the *i*th neuron being silent, or firing a spike within a time bin labeled by *t,* respectively. Without loss of generality, we randomly select a neuron from the network and label it as the "first neuron" and consider the conditional probability of this first neuron firing while other neurons are silent, *i.e.*,

$$p = P(\sigma_1 = 1 | \sigma_i = 0, \forall i \in S), \tag{4}$$

where $S = \{2, 3, \ldots, n\}$. Then, we consider probabilistically the influence of the other neurons on the firing of the first neuron—the *probability increment* of the first neuron firing induced by the *i*th neuron firing, that is,

$$\delta_i = P(\sigma_1 = 1 | \sigma_i = 1, \sigma_k = 0, \forall k \in S \backslash \{i\}) - p. \tag{5}$$

An operating state of the network is defined as p-polling state if, for any neuron selected as the "first neuron," and for any subset $B \subseteq S$ of firing neurons, the *probability increment* of the first neuron firing induced by neurons in B firing is equal to the linear sum of the probability increments induced by each individual neuron in the subset B firing, *that is*,

$$P(\sigma_1 = 1 | \sigma_i = 1, \forall i \in B; \sigma_k = 0, \forall k \in S \backslash B) - p = \sum_{i \in B} \delta_i.$$

The p-polling state indicates a linear computational rule of the firing dependence among neurons. Thus, we introduce a *linearity index* R_B to estimate how close a neural operating state is to a p-polling state. This linearity index R_B is defined by

$$R_B = \frac{P(\sigma_1 = 1 | \sigma_i = 1, \forall i \in B; \sigma_k = 0, \forall k \in S \backslash B)}{p + \sum_{i \in B} \delta_i}. \tag{6}$$

A probability increment δ is used to characterize probabilistically the dependence of a pair of neurons. The p-polling state is characterized by a linear superposition among δ's of different pairs of neurons and does not rely on the strength of δ of any pair of neurons. The p-polling state is different from the weakly correlated state as follows. The correlation is a linear measure and is restricted to the second-order statistics of signals, while the probability characterization is a model-free nonlinear measure and is related not only to low-order statistics (e.g., mean and covariance), but also high-order statistics. The weak correlation focuses on the magnitude of second-order linear statistics, while the p-polling state focuses on the linear superposition relation among conditional probability increments. They characterize the dynamics from different perspectives. Our previous works establish quantitative relationships underlying the probability distributions, moments, and effective interactions in the MEP analysis (Xu, Crodelle, Zhou, & Cai, 2019). On the one hand, if the strengths of all the δ's among neurons are small in the p-polling state, the correlations among neurons are indeed small. However, if the strengths of all the δ's among neurons are not all small but still satisfy the linear superposition, then the p-polling state is not the weakly correlated state (Xu et al., 2017; Xu, Crodelle, et al., 2019). On the other hand, if the network is in the weakly correlated state, it only constrains the magnitude of the second-order statistics. However, the p-polling state is defined as the linear superposition among multi-variable conditional probability increments. Therefore, the weakly correlated state is not necessarily the p-polling state. In Section 4, we use two idealized examples to illustrate the difference between the weakly correlated state and the p-polling state.

Considering that δ_i is often small in asynchronous networks, it is usually difficult to accurately measure δ_i for a full large network. Alternatively, one may consider *measured sub-networks* S_m of m neurons—those m neurons in the cortical network whose response properties can be measured by a multi-electrode experiment. If R_B is unity for any subset $B \subseteq S_m$ and for any neuron selected as the first neuron, the neural sub-network's operating point is likely to be the p-polling state. Note that for the trivial case where $\delta_i = 0$ for any i, a network consists of neurons firing independently with one another. However, the p-polling state as discussed below does not belong to such a special case.

2.10 | Total mean of a recording

For a recording of a network, the linearity index can be computed for sub-networks with different combinations of neurons and for different subsets within a sub-network. For large networks in practice, one is forced to validate the p-polling state in sub-networks since the verification of large network size encounters the issue of under sampling. To illustrate whether a network is in a p-polling state, we define

a *total mean* of linearity index for a recording. An example of a *total mean* of a recording is computed as follows: Consider M four-neuron sub-networks which are selected from the recording. Define the subsets with size $N(B)$ for the jth neuron as all the possible combinations of $N(B)$ neurons (the jth neuron is excluded) in the considered network. For each neuron in a four-neuron sub-network, there are three subsets with size $N(B) = 2$ selected from the other three neurons in the four-neuron sub-network. For example, for neuron 2, subsets include {1, 3}, {1, 4}, {3, 4}. There are twelve subsets in total for this four-neuron sub-network (denoted by the index i). We compute the mean of all twelve R_B's, denoted by $\mu_{2,i}$ for the ith selected four-neuron sub-network. For the recording with M selected four-neuron sub-networks, the total mean is defined as the mean of $\{\mu_{2,i}\}_{i=1}^{M}$. The total mean for $N(B) = 3$ is defined similarly.

3 | RESULTS

3.1 | Examinations in physiological experiments

We examine the cortical operating points in three physiological in vivo experiments (Coen-Cagli et al., 2015; Kohn & Coen-Cagli, 2015; Kohn & Smith, 2016; Liu et al., 2014; see Section 2 for details).

In the first experiment, data were recorded from the medial prefrontal cortex when mice were performing an olfactory delayed non-match to sample task (Liu et al., 2014; dataset one). In the second experiment, data were recorded from the primary visual cortex (V1) when four anesthetized adult macaca fascicularis monkeys viewed a uniform gray screen (Kohn & Smith, 2016; dataset two). In the third experiment, data were recorded from the primary visual cortex when anesthetized macaque monkeys received each image stimulus presented in pseudo-random order for 100 ms (data of all image stimuli are collected for dataset three), followed by a 200 ms uniform gray screen (data of all gray screen stimuli are collected for dataset four; Coen-Cagli et al., 2015; Kohn & Coen-Cagli, 2015). The four datasets are analyzed in four rows in Figure 1, respectively. In addition to two distinct animal models (mouse and monkey), these experimental settings include several representative (a) brain states (anesthetized spontaneous state, performing tasks of working memory, and visual information processing); (b) brain areas (mPFC and V1); (c) brain functions (working memory and visual information processing). Thus, these experimental settings represent some degree of generality for investing the commonality of the p-polling state.

The neural populations recorded in above experiments are in asynchronous state. For illustration, we show the raster plot of all neurons in one selected recording for each dataset, as shown in Figure 1a,d,g,j). The synchronization indexes are 0.16, 0.22, 0.22, and 0.26 for Figure 1a,d,g,j, respectively. We *randomly* select four neurons (marked by red color in Figure 1a,d,g,j) as examples to show that the linearity indexes are close to unity. We examine R_B for the subsets with size $N(B) = 2$ and $N(B) = 3$, as shown in Figure 1b,e,h,k. For each neuron, we compute the mean value and standard deviation of R_B with size $N(B) = 2$ and size $N(B) = 3$. The mean values of R_B of all subsets are very close to unity, which indicates that these four-neuron sub-networks in each recording are operating in a state very close to the p-polling state. To be more convincing, we would next examine the p-polling state in massive randomly selected sub-networks.

For a large network, it often requires an unrealistically long recording in order to compute the linearity index. An alternative way is to examine the linearity index for sufficient number of randomly selected sub-networks. We then examine R_B for four-neuron sub-networks with firing rates larger than 6 Hz in all recordings. Note that 6 Hz is an empirical threshold to insure sufficient number of spikes to estimate the NFP distribution; it is not a threshold for the existence of the p-polling state. For illustration, we present *total means* (see Section 2 for details) of *each recording* respectively for $N(B) = 2$ and $N(B) = 3$. Each total mean is close to one indicated by one dot in Figure 1c,f,i,l, which are indexed by their recording indexes (abscissa). Most mean values of R_B consisting in the total means for the subsets with size $N(B) = 2$ and $N(B) = 3$ are also very close to unity. For illustration, for the recording in Figure 1a, we present the mean R_B of all randomly selected four-neuron sub-networks for $N(B) = 2, 3$ in Figure S3. The linearity indexes are also close to unity for sub-networks of six and ten neurons (see an example in Figure S4), that is, the p-polling state is also verified for sub-networks of six and ten neurons.

We have also probed into the p-polling state in different behavior states of one animal. For the first dataset, we have examined the operating state when the mouse is in the delay period and in the resting period, respectively. In both periods, neural networks are also close to be in the p-polling state, as shown in the analysis of one recording of Mouse Three in Figure S5. We also shuffle data in four datasets (see an example in Figure S5) to confirm that the state of neuron-independence is a trivial case of the p-polling state. For the trivial case, all the probability increments are zero, and the conditional probability reduces to the unconditional probability. The experiments we study in this work have probability increments that are not zero, which can be seen in Figure 2c,f,i,l discussed in the next section. Their linearity indexes close to unity shows that the networks are operating in a p-polling state. Note that, the linearity indexes can deviate significantly from unity as the

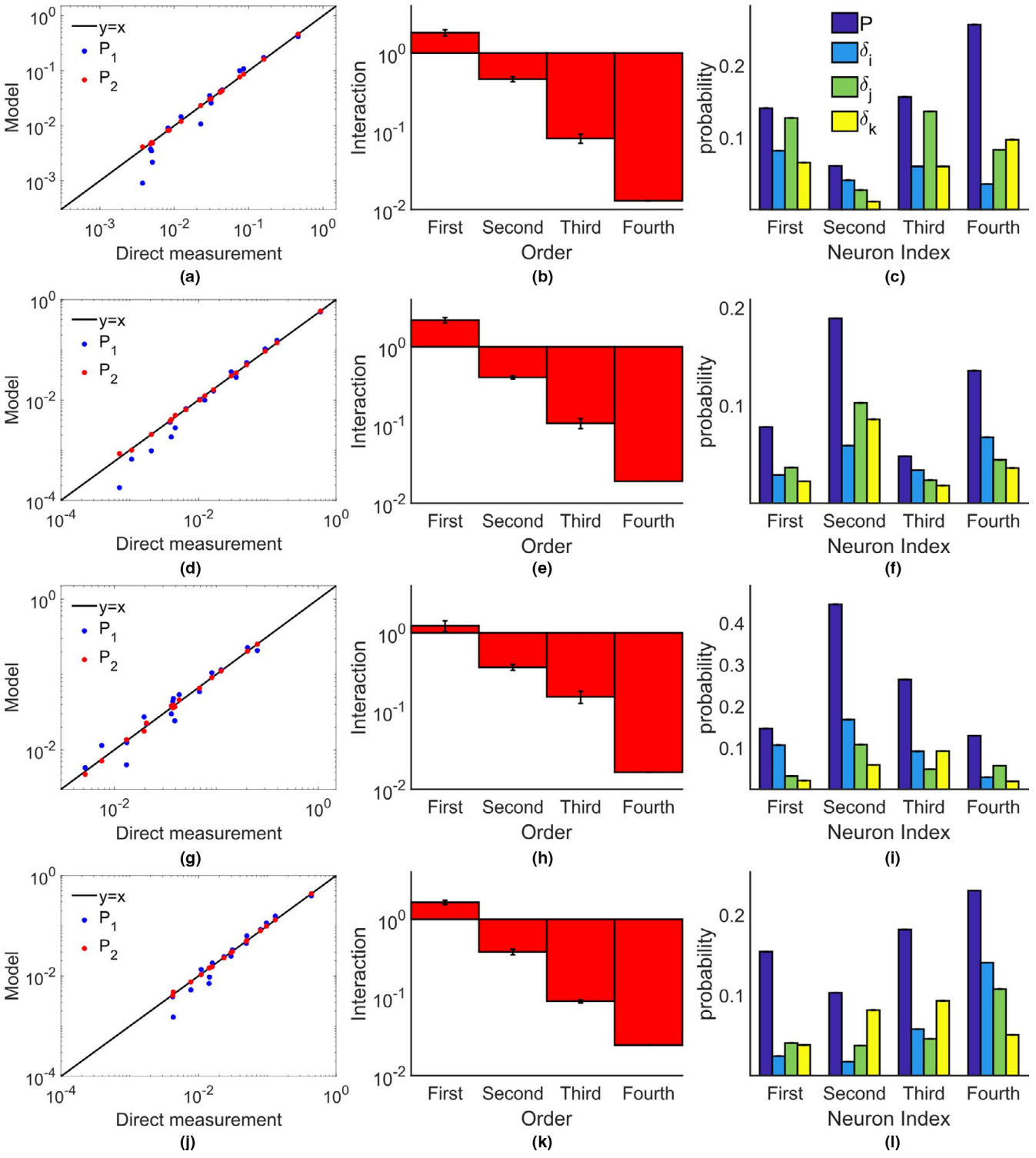

FIGURE 2 MEP analysis. Selected sub-networks in four datasets are the same as sub-networks marked by red color in Figure 1a,d,g,j, respectively. (a, d, g, j): The probability of occurrence of each firing pattern (denoted by one dot) predicted from the MEP analysis, P_2 (red) and P_1 (blue), is plotted against the measured probability. (b, e, h, k): Each bar is the mean of absolute values of interaction strengths of the order indicated by the abscissa. The standard deviation is also indicated by the error bar around the mean. (c, f, i, l): For each neuron, the blue bar is the probability of the neuron firing conditioned on other three neurons being silent (see Equation 4). Each bar with non-blue color is the probability increment of the neuron, whose index is indicated by the abscissa, firing induced by one of other neurons (see Equation 5) [Colour figure can be viewed at wileyonlinelibrary.com]

neural system displays more synchrony as discussed in our previous works (Xu et al., 2017), which indicates the breakdown of the p-polling state. Taken together, these datasets provide evidence that the p-polling state is a common cortical dynamical state—across different species and different behaviors.

3.2 | The second-order MEP analysis

The p-polling state, a common cortical dynamical state examined above, provides an efficient way to study the neural coding embedded in statistical distributions of NFPs. For example, statistical distributions of NFPs are found to be used to perform awake replay of remote experiences in rat hippocampus (Karlsson & Frank, 2009). However, the degree of freedom in distributions of NFPs, 2^n, grows exponentially with the neuron number n. Next, we would show for the four in vivo datasets, that these networks operating in the p-polling state with small magnitude of probability increments simplify the description of the NFP distribution by low-order MEP analysis with constraints of low-order correlations (see Section 2).

First, we show the high accuracy of the second-order MEP analysis for the NFP of the four experimental datasets. As shown in Figure 2a,d,g,j, for the four-neuron sub-networks (those marked by red color in Figure 1a,d,g,j), the second-order MEP distributions, P_2 (red), are in excellent agreement with the observed distributions; however, the first-order MEP distributions, P_1 (blue), have significant deviations from the observed distributions. The results of the MEP analysis for all analyzed four-neuron sub-networks selected from Figure 1a,d,g,j are similar (shown in Figure S6). Note that these results on the accuracy of low-order MEP analysis are consistent with previous experimental results on neural network systems (Schneidman et al., 2006; Shlens et al., 2006; Tang et al., 2008).

Since the NFP distribution of n neurons is the same as the distribution of its corresponding full-order (nth-order) MEP distribution (Xu, Crodelle, et al., 2019), which is constrained by all-order correlations. The good performance of the second-order MEP analysis leads us to the question of whether high-order interactions in the full-order MEP analysis are small compared with lower-order ones. Thus, in these experiments, we then examine the interaction strengths of different orders in the full-order MEP analysis. The interaction strengths of different orders can be derived from the observed distribution (see Section 2). As shown in Figure 2b,e,h,k, for illustration, we only show results of those sub-networks marked by red color in Figure 1a,d,g,j. The bars are the means of absolute interaction strengths of different orders. It can be seen clearly that the means of higher-order absolute interaction strengths are much smaller than those of the first and second orders.

In our theoretical work (Xu et al., 2017), we proposed a mechanism to explain the situations where high-order interactions are small. There is a recursive relation in the full-order MEP analysis, namely, the $(k+1)$th order interaction $J_{123\ldots(k+1)}$ is the increment of the kth order interaction $J_{123\ldots k}$ induced by the $(k+1)$th neuron firing (see Section 2). This recursive relation can be analyzed theoretically in the p-polling state. For illustration, in a homogeneous network, *that is*, the probability increment satisfies $\delta_i = \delta$ for $2 \le i \le n$, the interaction strengths form a hierarchy in the power of probability increment δ, *that is*, $J_{123\ldots k} \sim O(k^{k-1})$. Importantly, in asynchronous neural networks, δ_i $(2 \le i \le n)$ is usually small when compared to p, the neuron's firing probability conditioned on the silence of all other neurons. This fact has been confirmed computationally for HH neural networks in a wide range of dynamical regimes (Xu et al., 2017), and it is also verified for the four in vivo cortical sub-networks shown in Figure 2c,f,i,l. Therefore, high-order interactions are much smaller compared with lower-order ones when the network is under the p-polling state. As discussed previously, if the system is in a p-polling state, only when the strengths of all the probability increments δ's are small, can the second-order MEP analysis be applied. Therefore, the MEP framework might be regarded as a special case of the coding strategy in the p-polling state.

4 | DISCUSSION

We discuss the relationship of p-polling state to other studies, and also some of its limitations.

4.1 | Balanced states in networks

In cortical regions (Haider, Duque, Hasenstaub, & McCormick, 2006), neurons are often found to operate in a "balanced state," in which the excitatory and inhibitory synaptic inputs into each neuron are dynamically balanced on average (Shadlen & Newsome, 1994; Van Vreeswijk & Sompolinsky, 1996). These balanced-state networks can display various firing patterns (Boerlin, Machens, & Denève, 2013; Denève & Machens, 2016; Moreno-Bote, 2014; Renart et al., 2010;), and often possess asynchronous states with low firing rates. As we have computed in neural network simulations (see an example in Figure S7), the linearity indexes of balanced states are near unity—showing that the p-polling state holds well in balanced states. Since the p-polling state might be more general operating points than balanced states, the theoretical studies of coding schemes in balanced-state networks can be extended to the p-polling state.

4.2 | Weak correlations

Weak correlation between neural spike trains is widely observed in experiment (Cohen & Kohn, 2011); although weak, these correlations can significantly impact the collective behavior (Schneidman et al., 2006). Pairwise correlations are not always detrimental for the neural coding since

decorrelation in such case does not imply an increase in information (Moreno-Bote et al., 2014). However, differential correlations, which are proportional to the product of the derivatives of the tuning curves, can limit substantially the signal capacity of neural populations (Moreno-Bote et al., 2014; Nogueira et al., 2019). Studies of coding schemes are often based upon the correlation structure of the network (Abbott & Dayan, 1999; Kanitscheider, Coen-Cagli, & Pouget, 2015; Quiroga & Panzeri, 2009) or inverse population covariability (Nogueira et al., 2019). Thus, a deeper and more quantitative understanding of the structure of correlations may help to discover new coding principles. The analysis framework of the p-polling state provides one reason that weak correlations among neurons are often observed (Cohen & Kohn, 2011). To illustrate this, we consider an example of a homogeneous network of n neurons (Xu et al., 2017), operating in a p-polling state. In this setting, correlations between neurons can be expressed in terms of p and δ; then, the correlation among neurons $\{i_1, i_2, \ldots, i_k\}$, *i.e.*, $C_{i_1,i_2,\ldots,i_k}$, forms a hierarchy in the power (Xu et al., 2017) of δ, *i.e.*, $C_{i_1,i_2,\ldots,i_k} = O(\delta^{k-1})$. Therefore, the p-polling state shows a quantitative understanding of why weak correlation is common.

The p-polling state can be different from the weakly correlated state. For example, consider an idealized homogeneous network of three neurons, which might be realized within a large network, that is,

$$P_{100} = P_{010} = P_{001},$$

$$P_{110} = P_{011} = P_{101},$$

where $p_{\sigma_1\sigma_2\sigma_3}$ is the probability of the state $(\sigma_1, \sigma_2, \sigma_3)$, σ_i is the state of the ith neuron. Assume that p, the probability of each neuron firing conditioned on that other two neurons are silent, is 0.1, that is,

$$p = \frac{P_{100}}{P_{100} + P_{000}},$$

and assume that δ, the probability increment of one neuron, is 0.44, that is,

$$\delta = \frac{P_{110}}{P_{110} + P_{010}} - p.$$

Due to the normalization of all probabilities, we have

$$\sum p_{\sigma_1\sigma_2\sigma_3} = 1.$$

If the network is in the p-polling state, we further have

$$\frac{P_{111}}{P_{111} + P_{011}} - p = 2\delta.$$

The above equations can be analytically solved since the number of constraints is the same as that of unknown variables. The Pearson correlation coefficient in such case can be obtained as .786. Compared with often observed magnitude of the correlation coefficient (Cohen & Kohn, 2011) around .1, correlation of .786 in this idealized example is far from being weakly correlated. On the other hand, we assume $p = .3$, $\delta = .01$, and replace the last equation by

$$\frac{P_{111}}{P_{111} + P_{011}} - p = 25\delta,$$

which means the p-polling state does not hold. All the linearity indexes for subsets with size 2 is 1.72. In this case, the Pearson correlation coefficient can also be obtained as .10, that is, the network is weakly correlated but not in the p-polling state. These examples indicate that the p-polling state can be different from a weakly correlated state.

4.3 | Limitations

There are several limitations to this work. *First*, in physiological animal experiments, the length of recording time is always limited. This practical limitation on the length of the time series data places significant restrictions on the data analysis of the linear computational rule in the p-polling state of a large network and temporal correlations in population responses. For example, the mean linearity indexes for $N(B) = 2$ of sub-networks with 4, 6, and 10 neurons (Figures S3 and S4) are 1.00, 1.00, and 1.01, respectively, while their standard deviations are 0.04, 0.06, and 0.10, respectively. Note, the standard deviation increases with the sub-network size. For the large sub-network size, the estimation of the probability distribution function may suffer from the curse of dimensionality—making it difficult to estimate accurately the probability distribution function and thus lead to high fluctuations of the calculated linearity index. However, neural networks have many features, such as sparse connections and sparse firings, which make it also promising to develop methods to study the p-polling state in large networks. *Second*, it is still unclear how the p-polling state emerges from neural dynamics. The p-polling state is a common cortical operating point, but the cortex does not always operate in the p-polling state. For example, synchrony certainly prevents its emergence, as synchrony is far from p-polling state, as has been shown in computational HH model networks (Xu et al., 2017). Cortical networks normally operate in asynchronous states (Renart et al., 2010; Van Vreeswijk & Sompolinsky, 1996), such as the balance state, where the p-polling state holds well. For experimental data measured by calcium imaging from hippocampal neurons and

glial cells from E18 rat embryos in vitro (Lau & Bi, 2005), we have seen an indication that synchrony is far from the p-polling state; however, those data are too short to be conclusive and cortical networks normally operate in asynchronous states (Renart et al., 2010; Van Vreeswijk & Sompolinsky, 1996); It is expected that synchrony of the whole network may occur in abnormal situations such as epilepsy (Traub & Wong, 1982) and Alzheimer (Kuchibhotla, Lattarulo, Hyman, & Bacskai, 2009)—states which are far from p-polling.

There is a direct connection between the degree of synchrony and the probability increment δ, as δ is related to a commonly used "synchrony index" between a pair of neurons (*e.g.*, neuron 1 and 2) (Shlens et al., 2006); that is, $I_{\text{syn}} = \log_2[P(\sigma_1 = 1,\ \sigma_2 = 1)/(P(\sigma_1 = 1)P(\sigma_2 = 1))]$. By implementing p and δ_2 into this formula, $I_{\text{syn}} = \log_2[(p+\delta_2)/(p+P(\sigma_2 = 1)\delta_2)]$. When (a) $\delta_2 = 0,\ I_{\text{syn}} = 0$, *that is*, totally independent; (b) $\delta_2 > 0,\ I_{\text{syn}} > 0$, *that is*, tending to synchrony; (c) $\delta_2 < 0,\ I_{\text{syn}} < 0$, *that is*, tending to antisynchrony.

In any case, as the linearity index moves away from unity, it is necessary to further investigate the structure of the operating states which is beyond the p-pooling state.

5 | CONCLUSION

In this work, by calculating the linearity indexes directly from multi-electrode recording data measured in several distinct experiments on mouse and monkey, we have shown that the p-polling state is a rather common operating point of the cortical network, across different species and in different behavioral states. The mathematical characterization of the p-polling state is consistent with the theoretical assumption in the previous study on HH model simulations (Xu et al., 2017). As a step toward the Marr's second-level understanding, the commonality of the p-polling state indicates that although animals perform different tasks, in different behavioral states, they use similar representations and algorithms; that is, a linear computation rule of the dependence among neurons on the probability level defined by the p-polling state.

We then show that under the p-polling cortical operating state with small magnitude of probability increments, the second-order MEP distribution provides an efficient representation of the observed neural firing patterns. The NFP distribution is important for understanding neural systems. For example, the NFP distribution has been utilized both to study neural decoding (Karlsson & Frank, 2009) and the interaction between neurons of different types (Meshulam et al., 2017). Our theoretical work shows that in the p-polling state, a second-order MEP analysis performed with a short time recording produces a much more accurate probability distribution than the probability distribution directly constructed from the short time recording (Xu, Crodelle, et al., 2019). In addition, our theoretical work (Xu, Zhou, et al., 2019) indicates that sparsity of connections can lead to sparse high-order interactions in the representation of the distribution, and thus the accumulation of high-order interaction effects is small and the accuracy of low-order MEP distributions can be guaranteed in larger networks compared with that discussed in previous works (Ganmor, Segev, & Schneidman, 2011; Meshulam et al., 2017). Thus, for sparsely connected networks, the p-polling state provides a common operating point from which the NFP distribution can be sparsely represented, providing a natural sparse-coding network perspective to study information processing.

In summary, we have shown a p-polling state that is defined by an underlying linearity structure. When the second-order MEP framework cannot be applied, the linearity of the p-polling state, which constrains the probability distribution to a low-dimensional manifold, may provide a new perspective to the understanding of the coding scheme of neural network dynamics.

ACKNOWLEDGEMENTS

The authors thank Zhongqi Tian for providing the code of the balanced state simulation. This work was supported by National Key R&D Program of China (2019YFA0709503) (D.Z., Z.X.), National Science Foundation in China with grants no. 11671259, 11722107, and the Student Innovation Center at Shanghai Jiao Tong University (D.Z.); by SJTU-UM Collaborative Research Program (D.C. and D.Z.)."

CONFLICT OF INTEREST

The authors declare no competing financial interests.

AUTHOR CONTRIBUTIONS

Z.X., D.C., D.Z., and D.M. involved in conceptualization and resources. Z.X. involved in methodology, software, formal analysis, investigation, data curation, and writing—original draft preparation. X. G. and C. L. involved in mice experiments. Z.X. and D.Z involved in validation and visualization. Z.X., D.Z., and D.M. involved in writing—review and editing. D.C., D.Z., and D.M. involved in supervision, project administration, and funding acquisition.

DATA AVAILABILITY STATEMENT

The data of the first scenario are openly available upon request. The data of other scenarios are openly available from crcns.org.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ejn.14860.

ORCID

Zhi-Qin John Xu https://orcid.org/0000-0003-0627-3520
Douglas Zhou https://orcid.org/0000-0001-6631-2961

REFERENCES

Abbott, L. F., & Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. *Neural Computation*, *11*(1), 91–101. https://doi.org/10.1162/089976699300016827

Alon, U. (2007). Network motifs: Theory and experimental approaches. *Nature Reviews Genetics*, *8*(6), 450. https://doi.org/10.1038/nrg2102

Amari, S. I. (2001). Information geometry on hierarchy of probability distributions. *IEEE Transactions on Information Theory*, *47*(5), 1701–1711. https://doi.org/10.1109/18.930911

Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews Neuroscience*, *7*(5), 358–366. http://www.nature.com/nrn/journal/v7/n5/full/nrn1888.html

Boerlin, M., Machens, C. K., & Denève, S. (2013). Predictive coding of dynamical variables in balanced spiking networks. *PLoS Computational Biology*, *9*(11), e1003258. https://doi.org/10.1371/journal.pcbi.1003258

Coen-Cagli, R., Kohn, A., & Schwartz, O. (2015). Flexible gating of contextual influences in natural vision. *Nature Neuroscience*, *18*(11), 1648–1655. https://doi.org/10.1038/nn.4128

Cohen, M. R., & Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nature Neuroscience*, *14*(7), 811–819. https://doi.org/10.1038/nn.2842

Denève, S., & Machens, C. K. (2016). Efficient codes and balanced networks. *Nature Neuroscience*, *19*(3), 375. https://doi.org/10.1038/nn.4243

Ganmor, E., Segev, R., & Schneidman, E. (2011). Sparse low-order interaction network underlies a highly correlated and learnable neural population code. *Proceedings of the National Academy of Sciences*, *108*(23), 9679–9684. https://doi.org/10.1073/pnas.1019641108

Gerstner, W., & Kistler, W. M. (2002). *Spiking neuron models: Single neurons, populations, plasticity*. Cambridge, UK: Cambridge University Press.

Haider, B., Duque, A., Hasenstaub, A. R., & McCormick, D. A. (2006). Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *Journal of Neuroscience*, *26*(17), 4535–4545. https://doi.org/10.1523/JNEUROSCI.5297-05.2006

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, *117*(4), 500–544. https://doi.org/10.1113/jphysiol.1952.sp004764

Kanitscheider, I., Coen-Cagli, R., & Pouget, A. (2015). Origin of information-limiting noise correlations. *Proceedings of the National Academy of Sciences*, *112*(50), E6973–E6982. https://doi.org/10.1073/pnas.1508738112

Karlsson, M. P., & Frank, L. M. (2009). Awake replay of remote experiences in the hippocampus. *Nature Neuroscience*, *12*(7), 913. https://doi.org/10.1038/nn.2344

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*(12), 712–719. https://doi.org/10.1016/j.tins.2004.10.007

Kohn, A., & Coen-Cagli, R. (2015). Multi-electrode recordings of anesthetized macaque V1 responses to static natural images and gratings. CRCNSorg. Retrieved from http://dxdoiorg/106080/K0SB43P8

Kohn, A., & Smith, M. A. (2016). Utah array extracellular recordings of spontaneous and visually evoked activity from anesthetized macaque primary visual cortex (V1). CRCNSorg. Retrieved from http://dxdoiorg/106080/K0NC5Z4X

Kuchibhotla, K. V., Lattarulo, C. R., Hyman, B. T., & Bacskai, B. J. (2009). Synchronous hyperactivity and intercellular calcium waves in astrocytes in Alzheimer mice. *Science*, *323*(5918), 1211–1215. https://doi.org/10.1126/science.1169096

Lau, P. M., & Bi, G. Q. (2005). Synaptic mechanisms of persistent reverberatory activity in neuronal networks. *Proceedings of the National Academy of Sciences*, *102*(29), 10333–10338. https://doi.org/10.1073/pnas.0500717102

Liu, D., Gu, X., Zhu, J., Zhang, X., Han, Z., Yan, W., … Li, Y.(2014). Medial prefrontal activity during delay period contributes to learning of a working memory task. *Science*, *346*(6208), 458–463. https://doi.org/10.1126/science.1256573

Marr, D. (1982). Vision: A computational approach. Freeman [aAC].

Meshulam, L., Gauthier, J. L., Brody, C. D., Tank, D. W., & Bialek, W. (2017). Collective behavior of place and non-place neurons in the hippocampal network. *Neuron*, *96*(5), 1178–1191. https://doi.org/10.1016/j.neuron.2017.10.027

Moreno-Bote, R. (2014). Poisson-like spiking in circuits with probabilistic synapses. *PLoS Computational Biology*, *10*(7), e1003522. https://doi.org/10.1371/journal.pcbi.1003522

Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., & Pouget, A. (2014). Information-limiting correlations. *Nature Neuroscience*, *17*(10), 1410. https://doi.org/10.1038/nn.3807

Nguyen, J. P., Shipley, F. B., Linder, A. N., Plummer, G. S., Liu, M., Setru, S. U., … Leifer, A. M. (2016). Whole-brain calcium imaging with cellular resolution in freely behaving *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences*, *113*(8), E1074–E1081.

Nogueira, R., Peltier, N. E., Anzai, A., DeAngelis, G. C., Martínez-Trujillo, J., & Moreno-Bote, R. (2019). The effects of population tuning and trial-by-trial variability on information encoding and behavior. *The Journal of Neuroscience*, *40*(5), 1066–1083. https://doi.org/10.1523/JNEUROSCI.0859-19.2019

Quiroga, R. Q., & Panzeri, S. (2009). Extracting information from neuronal populations: Information theory and decoding approaches. *Nature Reviews Neuroscience*, *10*(3), 173–185. https://doi.org/10.1038/nrn2578

Renart, A., dela Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., & Harris, K. D. (2010). The asynchronous state in cortical circuits. *Science*, *327*(5965), 587–590. https://doi.org/10.1126/science.1179850

Schneidman, E., Berry, M. J., Segev, R., & Bialek, W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*, *440*(7087), 1007–1012. https://doi.org/10.1038/nature04701

Shadlen, M. N., & Newsome, W. T. (1994). Noise, neural codes and cortical organization. *Current Opinion in Neurobiology*, *4*(4), 569–579. https://doi.org/10.1016/0959-4388(94)90059-0

Shlens, J., Field, G. D., Gauthier, J. L., Grivich, M. I., Petrusca, D., Sher, A., … Chichilnisky, E. J. (2006). The structure of multi-neuron firing patterns in primate retina. *The Journal of Neuroscience*, *26*(32), 8254–8266. https://doi.org/10.1523/JNEUROSCI.1282-06.2006

Tang, A., Jackson, D., Hobbs, J., Chen, W., Smith, J. L., Patel, H., … Beggs, J. M. (2008). A maximum entropy model applied to spatial and temporal correlations from cortical networks in vitro. *The Journal of Neuroscience*, *28*(2), 505–518. https://doi.org/10.1523/JNEUROSCI.3359-07.2008

Traub, R. D., & Wong, R. (1982). Cellular mechanism of neuronal synchronization in epilepsy. *Science*, *216*(4547), 745–747. https://doi.org/10.1126/science.7079735

Van Vreeswijk, C., & Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, *274*(5293), 1724–1726. https://doi.org/10.1126/science.274.5293.1724

Xu, Z. Q. J., Bi, G., Zhou, D., & Cai, D. (2017). A dynamical state underlying the second order maximum entropy principle in neuronal networks. *Communications in Mathematical Sciences*, *15*(3), 665–692. https://doi.org/10.4310/CMS.2017.v15.n3.a5

Xu, Z. Q. J., Crodelle, J., Zhou, D., & Cai, D. (2019). Maximum entropy principle analysis in network systems with short-time recordings. *Physical Review E*, *99*(2), 022409. https://doi.org/10.1103/PhysRevE.99.022409

Xu, Z. Q. J., Zhou, D., & Cai, D. (2019). Dynamical and coupling structure of pulse-coupled networks in maximum entropy analysis. *Entropy*, *21*(1), 76. https://doi.org/10.3390/e21010076

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Xu Z-QJ, Gu X, Li C, Cai D, Zhou D, McLaughlin DW. Neural networks of different species, brain areas and states can be characterized by the probability polling state. *Eur J Neurosci*. 2020;52:3790–3802. https://doi.org/10.1111/ejn.14860