



# Collective Computation in Animal Fission-Fusion Dynamics

Gabriel Ramos-Fernandez<sup>1,2\*</sup>, Sandra E. Smith Aguilar<sup>3</sup>, David C. Krakauer<sup>4</sup> and Jessica C. Flack<sup>4</sup>

<sup>1</sup>Departamento de Modelación Matemática de Sistemas Sociales, Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, Ciudad de México, Mexico, <sup>2</sup>Unidad Profesional Interdisciplinaria en Ingeniería y Tecnologías Avanzadas, Instituto Politécnico Nacional, Ciudad de México, Mexico, <sup>3</sup>Conservación Biológica y Desarrollo Social A.C., Ciudad de México, Mexico, <sup>4</sup>Santa Fe Institute, Santa Fe, NM, United States

## OPEN ACCESS

### Edited by:

Daniel Polani,  
University of Hertfordshire,  
United Kingdom

### Reviewed by:

Deborah M. Gordon,  
Stanford University, United States  
Matthew Lutz,  
Max Planck Institute of Animal  
Behaviour, Germany  
Heiko Hamann,  
University of Lübeck, Germany

### \*Correspondence:

Gabriel Ramos-Fernandez  
ramosfer@alumni.upenn.edu

### Specialty section:

This article was submitted to  
Computational Intelligence in  
Robotics,  
a section of the journal  
Frontiers in Robotics and AI

**Received:** 31 October 2019

**Accepted:** 05 June 2020

**Published:** 21 July 2020

### Citation:

Ramos-Fernandez G, Smith  
Aguilar SE, Krakauer DC and Flack JC  
(2020) Collective Computation in  
Animal Fission-Fusion Dynamics.  
*Front. Robot. AI* 7:90.  
doi: 10.3389/frobt.2020.00090

Recent work suggests that collective computation of social structure can minimize uncertainty about the social and physical environment, facilitating adaptation. We explore these ideas by studying how fission-fusion social structure arises in spider monkey (*Ateles geoffroyi*) groups, exploring whether monkeys use social knowledge to collectively compute subgroup size distributions adaptive for foraging in variable environments. We assess whether individual decisions to stay in or leave subgroups are conditioned on strategies based on the presence or absence of others. We search for this evidence in a time series of subgroup membership. We find that individuals have multiple strategies, suggesting that the social knowledge of different individuals is important. These stay-leave strategies provide microscopic inputs to a stochastic model of collective computation encoded in a family of circuits. Each circuit represents an hypothesis for how collectives combine strategies to make decisions, and how these produce various subgroup size distributions. By running these circuits forward in simulation we generate new subgroup size distributions and measure how well they match food abundance in the environment using transfer entropies. We find that spider monkeys decide to stay or go using information from multiple individuals and that they can collectively compute a distribution of subgroup size that makes efficient use of ephemeral sources of nutrition. We are able to artificially tune circuits with subgroup size distributions that are a better fit to the environment than the observed. This suggests that a combination of measurement error, constraint, and adaptive lag are diminishing the power of collective computation in this system. These results are relevant for a more general understanding of the emergence of ordered states in multi-scale social systems with adaptive properties—both natural and engineered.

**Keywords:** social systems, distributed computing, inductive game theory, social information, animal foraging, collective intelligence

## 1. INTRODUCTION

In an influential framework for studying animal social organization, Hinde (1976) stressed that both animal and human societies are multiscale. Short-term interactions between pairs of individuals lead to longer-term social relationships and social structures, with social relationships arising as individuals generalize from a history of social interactions. Hinde noted that individuals

classify social relationships into types (kin, matriline, etc.) regardless of the individuals involved. The idea that primates use abstraction to make sense of their world has been shown in a number of studies subsequent to Hinde (1976) (e.g., Cheney and Seyfarth, 1990, 2008).

Over a series of papers, Flack et al. (Flack, 2012, 2017a,b; Flack et al., 2013; Daniels et al., 2017; Brush et al., 2018) have been developing a theory of collective computation (inspired in part by Hopfield's collective computation in neural networks Hopfield, 1982, 1984; Tank and Hopfield, 1988). In the context of animal behavior, this work links Hinde's (1976) generalization and abstraction processes to the formation of collectives. In Flack and Krakauer's formulation, components (for the purposes of this paper, individuals) reduce uncertainty about the environment or state of a system by coarse-graining fast microscopic behavior (Flack, 2017a). An example of uncertainty reduction would be over the cost of social interaction (Flack, 2012). When coarse-grainings converge (meaning the estimates of regularities are largely shared by individuals), this can produce a coherent mesoscale (e.g., a social network or circuit). This can then function like an information bottleneck (Tishby et al., 2000; Tishby and Zaslavsky, 2015; Flack, 2017a): the strategies, as coarse-grainings, capture regularities individuals perceive in the physical or social environment. The way individuals combine strategies to make decisions in the collective captures the regularities they perceive as most important. Emergent from these slowly changing mesoscopic individual strategies and collective metastrategies is social structure. As a social structure consolidates and individuals start to "reference it" for decision-making, it feeds back through effective downward causation (Flack, 2017a) to modulate the cost of social interaction or interaction with the environment. Once complete, this process can give rise to a new scale, and under suitable conditions, novel functions.

To make this concrete, consider as an example the collective computation of power structure in macaque societies (reviewed in Flack, 2012, 2017a). Individuals summarize fight histories using unidirectional signals. The sender emits the signal once it perceives it is likely to lose a fight. The signal reduces uncertainty in the receiver that the sender agrees to subordination—willingness to yield in future interactions. Encoded in the consolidating network or circuit of signals between group members is information about the distribution of power. Hence the power structure is computed as individuals estimate regularities about fighting abilities and share these opinions with the receiver and other group members via signals. Through this process, different levels of organization arise at successively slower timescales: fights (fast), signaling (slow), and power structure (slowest). The process of generating coarse-grained, slow variables (the signals, properties of the circuits) is the outcome of individual strategic computations (interaction and signaling decisions) that aggregate into an output collectively estimated to fit the state of the environment (Flack, 2017a,b). This two-part process of information accumulation and aggregation makes up collective computation (Daniels et al., 2017; Flack, 2017a).

Among other examples in the animal behavior literature that might result from collective computation are coordinated foraging and predator avoidance in animal groups (Couzin et al., 2003; Gordon, 2016; Sosna et al., 2019), rapid direction changes during collective motion in fish schools and bird flocks (Hein et al., 2015), and distributed foraging in social insects (Gordon, 2016).

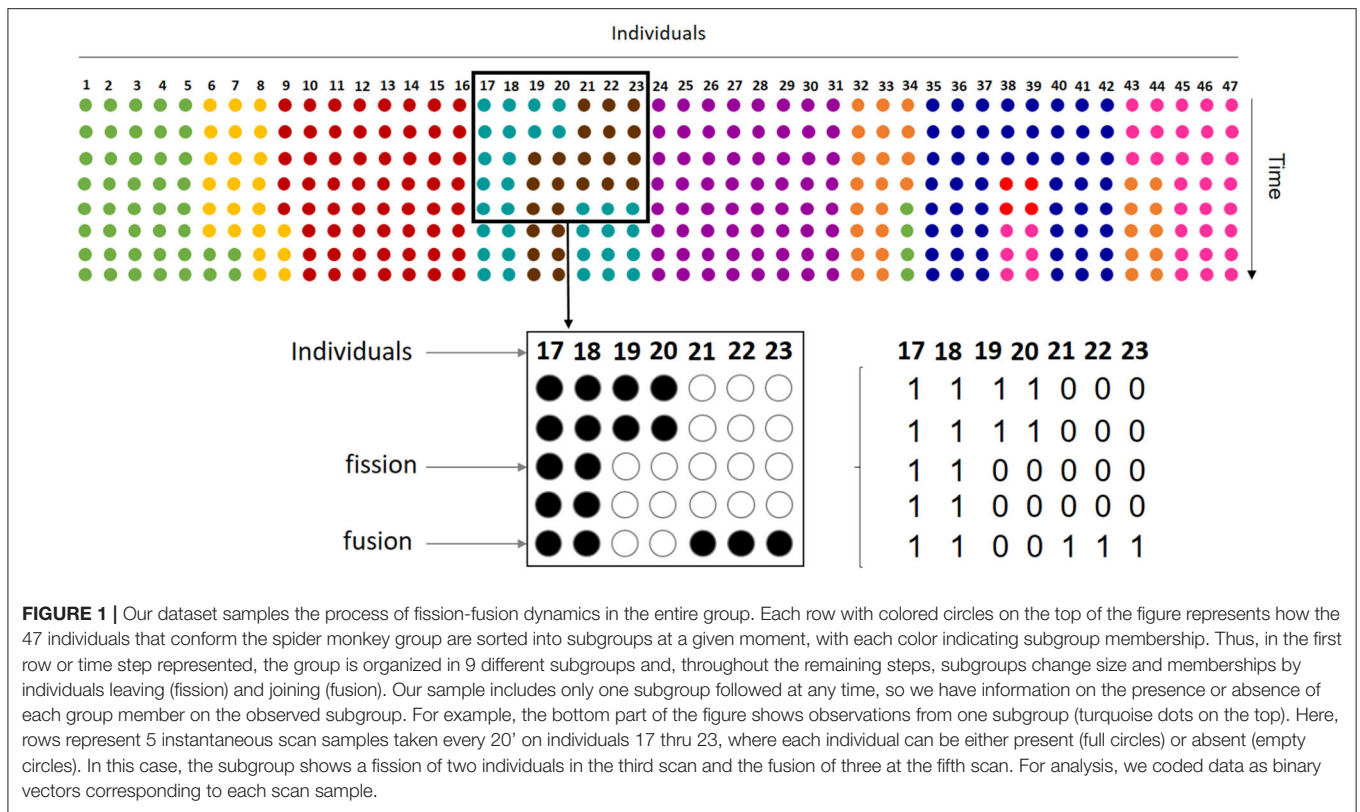
Fission-fusion social dynamics, in which individuals fission and fuse into subgroups of varying size, is a collective pattern arising from individual decisions (Sueur et al., 2011; Ramos-Fernández et al., 2018). These dynamics are thought to be adaptive, as they allow individuals to forage more efficiently in heterogeneous environments, share information about the location of resources, and adjust the size of their subgroups to resource availability (Aureli et al., 2008; Sueur et al., 2011; Palacios-Romo et al., 2019). The individual, strategic decisions to leave or join subgroups, how these decisions influence subgroup size distributions, and whether these are a good fit or even predicted by environmental states, are open questions. Previous work on spider monkeys suggests individuals change their strategies based on environmental states to include the rate at which they encounter fruit and the presence of knowledgeable individuals in social networks (Ramos-Fernández and Morales, 2014; Palacios-Romo et al., 2019).

We study how individual spider monkeys use social knowledge (information accumulation) to collectively compute adaptive subgroup size distributions (information aggregation). We use inductive game theory (DeDeo et al., 2010; Krakauer et al., 2010) to extract stay-leave probabilistic strategies from a time series of subgroup composition. The strategies constitute the microscopic input to the collective computation. From the microscopic input we construct a family of circuits in which nodes correspond to individuals and edges, weighted by probabilities obtained from the data, specify probabilistic rules—strategies—for remaining in or leaving a subgroup. Circuits capture variation in the way individuals integrate over their strategies (see section 3) to decide to stay or go.

Each circuit serves as a mesoscopic hypothesis for how strategies combine to produce decisions and how decisions combine to compute subgroup size distributions. In a computational language, the inputs (individual strategies) combine to produce an output (a subgroup size distribution). We run the circuits forward in simulation to determine how individuals combine strategies and hence how many information sources they take into account to make decisions. We construct a food abundance index based on the size and abundance of fruiting trees and calculate the transfer entropy between this index and the distribution of subgroup size in order to determine whether the circuit that best recovers the observed subgroup size distribution is also optimally computing the state of the environment.

## 2. DATA

Subgroup composition data were collected in Punta Laguna, Yucatan, Mexico, as part of a long-term study of social behavior



using identified individuals (details about study site and subjects can be found in the **Supplementary Information**). Data consist of scan samples of subgroup composition, taken every 20' during an average of 5 h. per day throughout 2 years (Jan. 2013–Dec. 2014), for a total of 5,780 scan samples. A total of 47 known adult, sub-adult and juvenile individuals were observed during this period (see **Supplementary Table 1**). Thus, each sample is a vector of 47 binary digits, with 0 corresponding to an absence of the individual in the *i*th position and 1 corresponding to a presence (**Figure 1**). Continuous series of scans, averaging 8.4 scan samples ( $\pm 3.9$  SD), include uninterrupted follows of a subgroup in which at least one individual remained during the full series. Given that the typical duration of a subgroup is 1.5 h. (Pinacho-Guendulain and Ramos-Fernández, 2017), a subgroup may persist over multiple scans. The temporal resolution of this sampling regime was maintained in the analysis in order to obtain a sufficient number of continuous series of scans. Had we resampled the original dataset at a larger temporal scale, we would have lost an important number of continuous series. Also, the persistence of a subgroup over several scans implies that individuals in a subgroup are tolerating one another, which is informative about the weight of their mutual influence (see below).

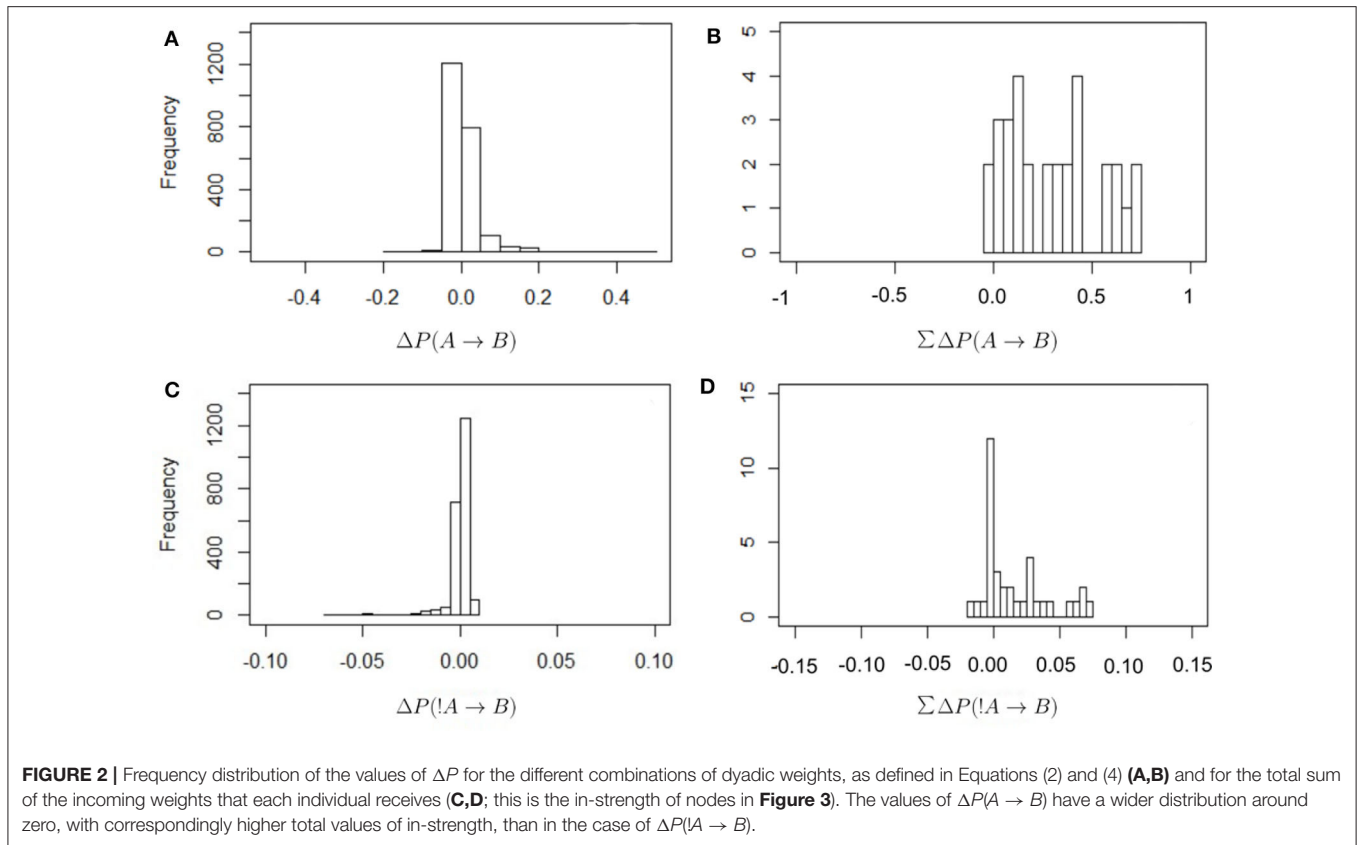
The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

### 3. MICROSCOPIC STRATEGY EXTRACTION AND DISTRIBUTION

We distinguish between strategies and decisions. A decision is binary: to leave or stay in a subgroup (in the original inductive game theory work, to join or avoid a fight, DeDeo et al., 2010). Strategies (called  $\Delta P$ , as in previous work, DeDeo et al., 2010) are “above-null” probabilities (see below for calculation) describing the weight of individual A’s presence or absence in the *current* subgroup (as determined by scan sampling, see section 2) on individual B’s decision to stay or go from the subgroup in the subsequent sample. Here and in previous work (DeDeo et al., 2010), multiple individuals can influence individual B. Hence B will have multiple strategies and, in the limit, a strategy for every other group member. We address how B integrates strategies to reach a decision in section 4. Here we quantitatively describe how we define and extract strategies from the time series. We end up with a list of pair-wise strategies for which our extraction method indicates above-null support in the time series. We do not consider higher order strategies as in DeDeo et al. (2010).

For all pairs of individuals {A:B, A:C, A:D,...}, we calculate the probability an individual B is present or absent in a sample if individual A was present in the previous sample within the same continuous series of scans:

$$P(A \rightarrow B) = \frac{N(B_{t+1} | A_t)}{N(A)} \tag{1}$$



where  $N(B_{t+1} | A_t)$  is the total number of times B was present at time  $t+1$  given that A was present at time  $t$  within a continuous series of scans and  $N(A)$  is the number of times A was present in all samples.

As with previous work (DeDeo et al., 2010), to remove time-independent effects from the transition probabilities (for example, due to general differences in gregariousness), we calculate the difference between the probability inferred from the data and a null expectation:

$$\Delta P(A \rightarrow B) = \frac{N(B_{t+1} | A_t) - N_{null}(B_{t+1} | A_t)}{N(A)}, \quad (2)$$

where  $N_{null}(B_{t+1} | A_t)$  is the average number of times B is present at time  $t+1$  given that A is present at time  $t$  within a continuous series of scans, calculated from 1,000 bootstrapped permutations of the data.

Similarly, we consider the weight of A's absence on the presence of another individual B in a subsequent sample:

$$P(!A \rightarrow B) = \frac{N(B_{t+1} | !A_t)}{N(!A)}, \quad (3)$$

and

$$\Delta P(!A \rightarrow B) = \frac{N(B_{t+1} | !A_t) - N_{null}(B_{t+1} | !A_t)}{N(!A)}, \quad (4)$$

where  $N(B_{t+1} | !A_t)$  is the number of times B is present in a sample when A is absent in the previous sample within a continuous series of scans,  $N(!A)$  is the number of times A is absent in all samples, and  $N_{null}(B_{t+1} | !A_t)$  is the average of the same number for 1,000 bootstrapped versions of the original data.

These  $\Delta P$  constitute the pair-wise weight of each group member on a given individual's binary decision to leave or join a subgroup.

**Figure 2** shows the frequency distribution of the values of  $\Delta P$  as defined in Equations (2) and (4). In all cases values are centered around zero, with the values of  $\Delta P(!A \rightarrow B)$  closer to zero than in other cases. This is because the denominator in Equation (4) is larger than in Equation (2), as it includes all instances of individual A being absent from the observed scan. There are proportionally fewer cases in which B is present after an absence of A because there are many cases where A is absent. Thus, these values of  $\Delta P(!A \rightarrow B)$  should be interpreted with care. It is also the case that most values of the total sum of weights received are positive. In other words, most individuals receive a total positive weight from the presence or absence of strategically connected individuals. Only a few cases show a total negative weight of the presence or absence of others.

We identified significantly positive dyadic weights as values of  $\Delta P$  higher than the 95% percentile of the permuted values for each dyad. Accordingly, significantly negative dyadic weights were values of  $\Delta P$  lower than the 5% percentile of the permuted values for each dyad.

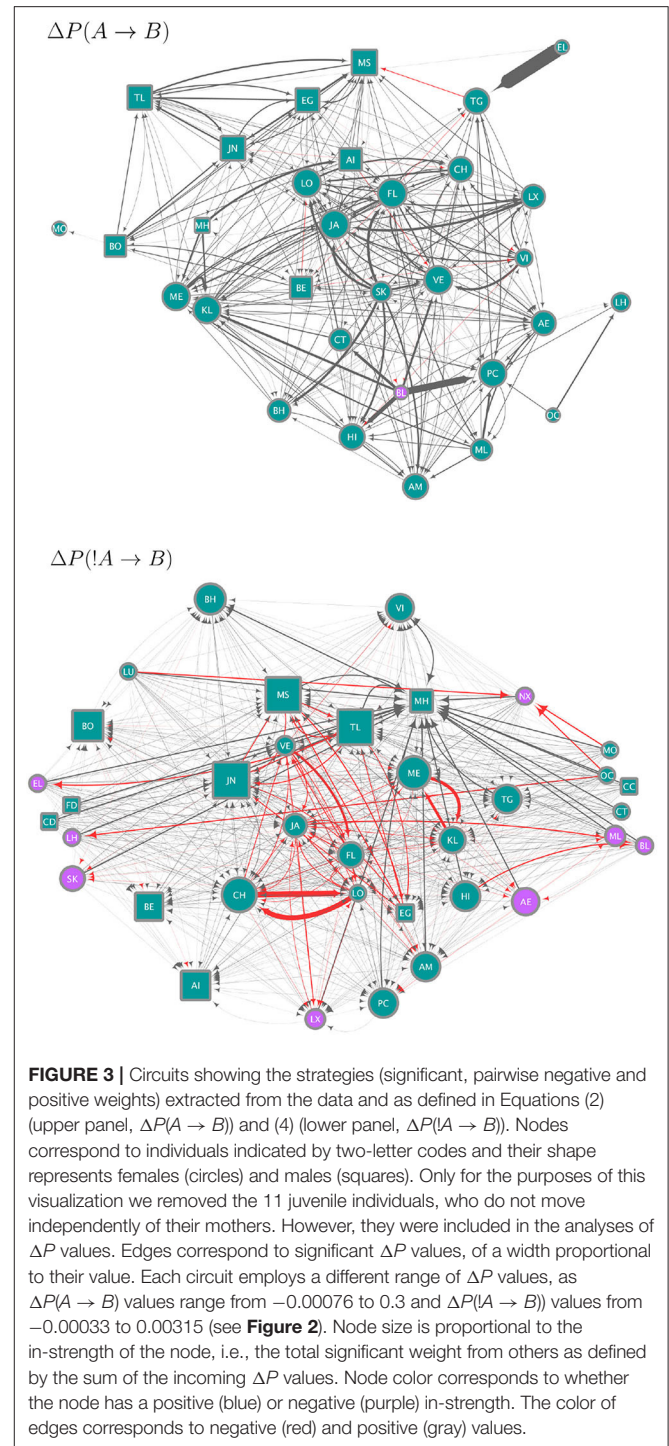


## 4. MESOSCOPIC CIRCUIT CONSTRUCTION

We use the strategies obtained from the data to construct circuits (*i.e.* the set of all significant  $\Delta P$  values as weights between all pairs of individuals; this is the mesoscopic level of our analysis) each of which is a hypothesis for (1) how individuals integrate over their strategies to arrive at a binary decision to join or leave a subgroup and, (2) specify how the resulting decisions combine to produce the distribution of subgroup size. The circuits in **Figure 3** give a qualitative summary of significant strategies. For each individual, there are 46 potential weights (significant  $\Delta P$  values) from either the presence or absence of others at scan time  $t$ , which could determine its presence or absence at scan time  $t + 1$ . The circuits in **Figure 3** show only 31 individual nodes for  $\Delta P(A \rightarrow B)$  and 36 for  $\Delta P(!A \rightarrow B)$ , who were involved in significant weights. On average, each individual in these circuits is linked to 20.25 ( $\pm 1.98$  SE) other individuals in the  $\Delta P(A \rightarrow B)$  and to 31.67 ( $\pm 1.40$  SE) other individuals in the  $\Delta P(!A \rightarrow B)$  circuit (**Figure 3**). Similarly, whereas each of the circuits in **Figure 3** could have up to 1,260 links, the  $\Delta P(A \rightarrow B)$  circuit has 314 and the  $\Delta P(!A \rightarrow B)$  circuit 570 links. **Supplementary Figure 1** shows the values of all significant weights included in these circuits, as well as the individual in-strength and outstrength.

The circuit for  $\Delta P(A \rightarrow B)$  (upper panel in **Figure 3**) represents significant weights of the presence of individual A at scan  $t$  on the presence of individual B at scan  $t + 1$ . Most of the values of  $\Delta P(A \rightarrow B)$  were positive or close to zero (see **Figure 2A**), therefore this circuit contains mostly positive weights (gray links), corresponding to weights of attraction. There is an apparent homophily by sex in this circuit, with individuals influencing other individuals of the same sex more than those of the other. Other attractive interactions are those between some pairs of adult females and their subadult daughters (e.g., females VE-VI and JA-LX in the upper panel of **Figure 3**, CH-LO and ME-KL in the lower panel). Individuals differ in their in-strength values (as can be observed in **Figure 2B**) with the individuals with the highest values of in-strength receiving many different weights, some with high values of  $\Delta P$ , both females and males. Only one individual (female BL) had a negative in-strength value, implying that it received a total negative  $\Delta P(A \rightarrow B)$  higher than the total positive  $\Delta P(A \rightarrow B)$ .

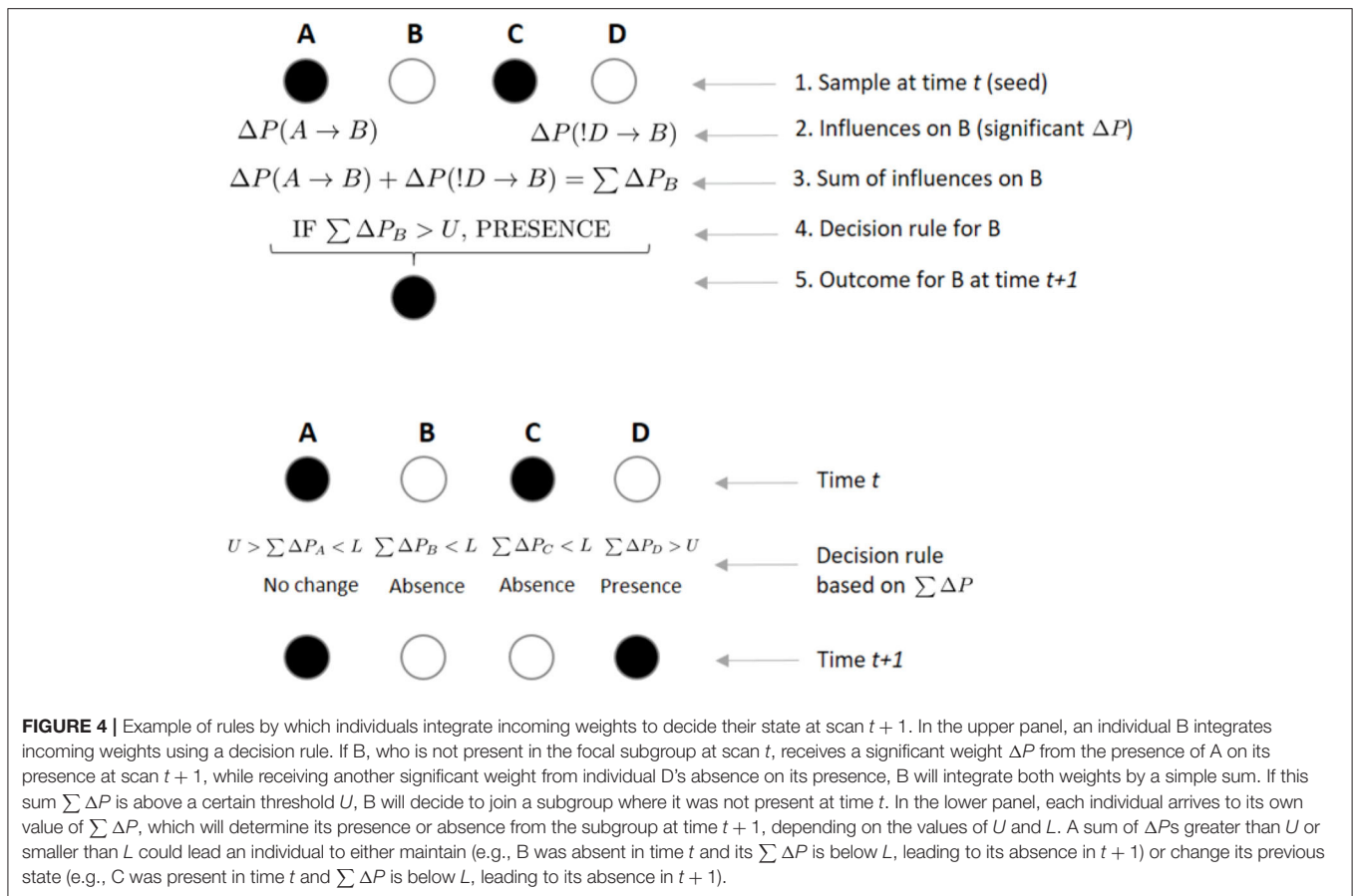
The circuit for  $\Delta P(!A \rightarrow B)$  shows a different picture (lower panel in **Figure 3**). Here values were skewed below zero, although overall they were much closer to zero than the values of  $\Delta P(A \rightarrow B)$  (**Figure 2**). Even considering that the variation around zero is small, this circuit contains both positive and negative weights, corresponding to repulsion and attraction, respectively, but the most important links are negative or attractive. There is, as in the previous circuit, evidence of some degree of homophily, with individuals of the same sex influencing each other through negative links more than those of the opposite sex. Conversely, a high proportion of positive or repulsive links occur between the sexes. Both males and females have high values of in-strength, although those with a negative in-strength (receiving many negative, attractive weights) in this circuit were all females. Individuals with the highest values of



**FIGURE 3** | Circuits showing the strategies (significant, pairwise negative and positive weights) extracted from the data and as defined in Equations (2) (upper panel,  $\Delta P(A \rightarrow B)$ ) and (4) (lower panel,  $\Delta P(!A \rightarrow B)$ ). Nodes correspond to individuals indicated by two-letter codes and their shape represents females (circles) and males (squares). Only for the purposes of this visualization we removed the 11 juvenile individuals, who do not move independently of their mothers. However, they were included in the analyses of  $\Delta P$  values. Edges correspond to significant  $\Delta P$  values, of a width proportional to their value. Each circuit employs a different range of  $\Delta P$  values, as  $\Delta P(A \rightarrow B)$  values range from  $-0.00076$  to  $0.3$  and  $\Delta P(!A \rightarrow B)$  values from  $-0.00033$  to  $0.00315$  (see **Figure 2**). Node size is proportional to the in-strength of the node, *i.e.*, the total significant weight from others as defined by the sum of the incoming  $\Delta P$  values. Node color corresponds to whether the node has a positive (blue) or negative (purple) in-strength. The color of edges corresponds to negative (red) and positive (gray) values.

positive in-strength (corresponding to a total sum of positive or repulsive weights in this network) were males.

Each individual can have multiple strategies, and they can be in conflict (DeDeo et al., 2010), with some weights positive and others negative. In addition, the weight or importance (given by  $\Delta P$ ) of each strategy varies. Hence individuals must integrate over their set of strategies to make a decision about whether



**FIGURE 4 |** Example of rules by which individuals integrate incoming weights to decide their state at scan  $t + 1$ . In the upper panel, an individual B integrates incoming weights using a decision rule. If B, who is not present in the focal subgroup at scan  $t$ , receives a significant weight  $\Delta P$  from the presence of A on its presence at scan  $t + 1$ , while receiving another significant weight from individual D's absence on its presence, B will integrate both weights by a simple sum. If this sum  $\sum \Delta P$  is above a certain threshold  $U$ , B will decide to join a subgroup where it was not present at time  $t$ . In the lower panel, each individual arrives to its own value of  $\sum \Delta P$ , which will determine its presence or absence from the subgroup at time  $t + 1$ , depending on the values of  $U$  and  $L$ . A sum of  $\Delta P$ s greater than  $U$  or smaller than  $L$  could lead an individual to either maintain (e.g., B was absent in time  $t$  and its  $\sum \Delta P$  is below  $L$ , leading to its absence in  $t + 1$ ) or change its previous state (e.g., C was present in time  $t$  and  $\sum \Delta P$  is below  $L$ , leading to its absence in  $t + 1$ ).

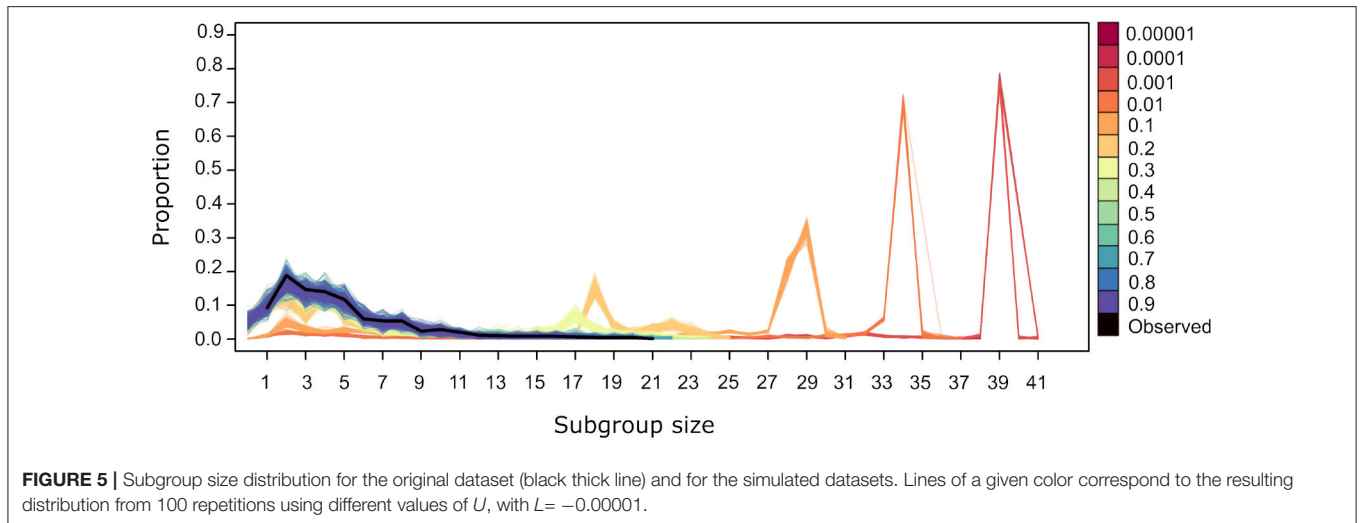
to join or leave the subgroup. **Figures 2B,D** show frequency histograms for these incoming values, corresponding to the in-strength of the nodes in **Figure 3**. These in-strength values can be understood as the likelihood that an individual will be influenced by others: an individual with a high in-strength is more likely to decide to be present due to another individual's presence (in the case of  $\Delta P(A \rightarrow B)$  values, upper panel in **Figure 3**) or absence (in the case of  $\Delta P(!A \rightarrow B)$  values, lower panel in **Figure 3**) than another individual with a lower in-strength.

We further assume that at any given time  $t$ , if the sum of significant  $\Delta P$  values  $\sum \Delta P$  directed toward an individual B is positive and greater than a threshold  $U$ , B will be present on the sample at  $t + 1$  (irrespective of whether it was present or absent in the previous sample; **Figure 4**). Conversely, if  $\sum \Delta P$  is negative and smaller than a threshold  $L$ , individual B will be absent from the following sample (again, independently of whether it was present or absent in the previous sample). However, if  $L < \sum \Delta P < U$ , then there is no effect from others and B remains in the same state as in the previous sample (i.e., present if it was present at time  $t$ , absent if it was absent; **Figure 4**). Thus,  $U$  is a threshold parameter controlling how likely it is for individuals to be present in a subgroup based on the weight of others. The value  $L$  controls the opposite, i.e., how likely it is that individuals will be absent in a subgroup based on the weight of others. Note that the total sum  $\sum \Delta P$  includes both the  $\Delta P(A \rightarrow B)$

and the  $\Delta P(!A \rightarrow B)$  values, such that an individual would be integrating the weights it receives across both circuits shown in **Figure 3**. At higher values of  $U$ , the presence of an individual in a subgroup is less likely to be influenced by others. In that sense, high values of  $U$  imply less interdependence of individuals in their decisions to be present or not in a subgroup. Conversely,  $L$  controls the opposite end of the range of values of  $\sum \Delta P$ , such that at more negative values of  $L$ , an individual should be less likely to be absent from a subgroup due to the previous weight from others. We tested  $U = \{0.0001, 0.001, 0.01, 0.1, 0.2, \dots, 0.9\}$  and  $L = \{-0.9, -0.8, \dots, -0.1, -0.01, \dots, -0.00001\}$ .

Different individuals could actually be using a different value of the  $U$  and  $L$  thresholds, or the values could change over time, depending on slower ecological variables such as the dry and wet seasons or even longer timescales related to the ecological succession of the forest in the spider monkey's habitat. In this work we assume, as a first approximation, a single value of the threshold parameters for all individuals and seasons.

There are also subtle points here concerning how strategies are aggregated by individuals to produce binary decisions. In previous work (DeDeo et al., 2010), higher order (triadic—C only joins current fight if both A and B were present in the previous fight) as well as pair-wise strategies (A joins if B was previously present) were extracted from time series data and a circuit was constructed for each strategy class. Preliminary



analyses in that work suggest these triadic strategies are non-decomposable into two pair-wise strategies (i.e., not reducible to additive individual or pair-wise interactions; Daniels et al., 2016; Chen et al., 2019). Individuals typically had multiple higher-order strategies and so, as with pair-wise, higher-order strategies were pushed through gates to produce binary decisions. Here we allow for the possibility that individuals take into account multiple strategies and hence be under the influence of multiple individuals, but we do not explore whether the interactions are pair-wise or higher-order.

We use these circuits to generate, by simulation, new datasets from the original dataset. In what follows, we restricted our analyses and simulations to a subset of the original dataset that included the same months for which food abundance data was available (Sep. 2013–Sep. 2014; see section 6), corresponding to 3,032 scan samples. We started by randomly choosing a scan sample (subgroup) that serves as the “seed” or first scan of a sequence of  $n$  samples, where  $n$  is randomly drawn from the frequency distribution of the number of samples per continuous observation period in the original biweekly period. Thus, the seed establishes which of the 47 monkeys in the group are present or absent in the first sample. Because the seed and the duration of continuous observation periods are selected within observation periods, simulated data contain information about the variation in subgroup size and composition between bi-weekly periods. If an individual  $A$  is present in the first scan, the simulation looks at values of  $\Delta P(A \rightarrow B)$  and considers any significant values or weights of  $A$  on others. If, on the contrary,  $A$  is not in the seed, then the simulation looks for significant values of  $\Delta P(!A \rightarrow B)$ . This applies to all 47 individuals.

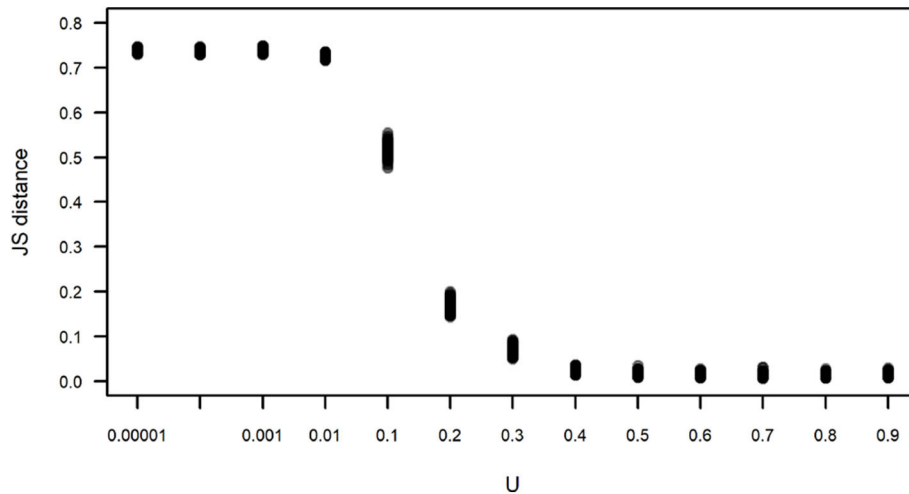
These rules are used to determine subgroup composition of the  $n$  samples in the continuous observation period. This is repeated for 633 sequences, corresponding to the number of continuous observation periods in the original dataset. In total, we generated 100 simulated datasets for each combination of thresholds  $U$  and  $L$ .

## 5. TESTING CIRCUITS IN SIMULATION

Here we assess how individuals integrate strategies to make decisions  $\Delta P$  and how decisions combine to compute the subgroup size distribution. We do so by asking which circuit, given an integration threshold, produces a simulated data set with a distribution of subgroup size that best recovers the observed one. We used each set of 100 simulated datasets with different values of  $U$  to evaluate the set of subgroup size distributions that is in closest correspondence to the observed. We only show the effects of varying  $U$  at  $L = -0.00001$ , since the variation in  $L$  for any value of  $U$  does not have an effect on the subgroup size distribution. This is likely because values of  $\sum \Delta P$  are mostly positive (Figures 2C,D), so very few values are below the  $L$  threshold. In other words, even the smallest negative value of  $L$  has no effect on the tendency of individuals to modify their presence based on the presence or absence of others.

For values of  $U = 0.4$  and above the subgroup size distribution from simulated datasets is similar to the observed (Figure 5). Values of  $U < 0.4$  generate distributions where small subgroups are underrepresented and larger subgroups are overrepresented. This is due to the fact that, at lower values of  $U$ , individuals are more likely to be influenced by others, both through the significant values of  $\Delta P(A \rightarrow B)$  and  $\Delta P(!A \rightarrow B)$ . The former dominate the dynamics of subgroup size change because they have higher and positive values overall (Figure 2). Thus, when  $U < 0.4$ , individuals are aggregating more frequently, deciding to join subgroups at higher frequency as in the observed data. Values of  $U < 0.4$  give rise to subgroups converging at a single size for each value of  $U$  (Figure 5). This may be due to all individuals deciding to join subgroups, even those without significant weights, as must be the case in subgroups larger than 36, the number of nodes in the largest network in Figure 3 that depicts all individuals that are involved in significant weights.

We compared the observed subgroup size distribution and those obtained by simulation under different values of  $U$  using



**FIGURE 6 |** Jensen-Shannon (JS) distance between the observed and simulated subgroup size distributions shown in **Figure 5**. Each dot corresponds to the JS distance between an instance of 100 simulations for each value of  $U$ . For all simulations,  $L = -0.00001$ .

the Jensen-Shannon distance (**Figure 6**). This distance between two random variables  $x$  and  $y$  is defined as:

$$JS(x|y) = H\left[\frac{x+y}{2}\right] - \frac{1}{2}[H(x) + H(y)] \quad (5)$$

where  $H$  is the entropy of each variable,  $p(x)\frac{1}{\log p(x)}$  and  $X$  and  $Y$  are, in this case, the observed subgroup size and the subgroup size obtained in one run of a simulation, respectively. **Figure 6** corroborates what is apparent in **Figure 5**, that simulations run with  $U \geq 0.4$  yield subgroup size distributions that are closer and indistinguishable from the observed distribution, with  $JS$  values that are close to zero, while simulations run with  $U < 0.4$  have an increasing  $JS$  with respect to the observed. Simulations run with all values of  $L$  for  $U=0.4$  yield subgroup size distributions that are equally close to the observed (data not shown).

## 6. FIT OF OUTPUT TO ENVIRONMENT

A central question is whether the collective computation output is adaptive (Flack, 2017a; Brush et al., 2018). Previous studies of spider monkeys suggest there is a weak relationship between subgroup size and food abundance (Symington, 1988; Pinacho-Guendulain and Ramos-Fernández, 2017). In general, subgroups tend to be larger during periods of high food abundance. This suggests that subgroup size can track the abundance of resources. Here, we investigate whether subgroup size distribution is predicted by the relative abundance of fruiting trees.

We use data from a 1-ha plot where all the trees (diameter at breast height,  $D > 10$  cm) from the 15 most consumed species by the monkeys, were monitored bi-weekly for a year from September 2013 to September 2014, comprising 25 monitoring periods. A total of 487 trees were identified, their  $D$  was recorded, and every 2 weeks they were assessed for the presence of fruit. The data obtained were used to calculate the proportion of trees with fruit available in a given period expressed in terms of the total tree

$D$  rather than tree number. To do so we calculated the sum of the  $D$  values of all the trees with fruit ( $D_f$ ) in period  $p$  divided by the sum of  $D$  values for all the trees in the plot ( $D_i$ ), giving an index of food abundance for a period  $p$ ,  $IFA_p = \sum D_f / \sum D_i$ .

**Figure 7** shows the time series for the IFA and subgroup size during one year. As mentioned above, maintaining the temporal resolution of the subgroup size time series was important in order to maintain a sufficient number of continuous series of observations. Despite the different temporal resolution of each time series, it seems that subgroup size increases together with IFA during the second wet season.

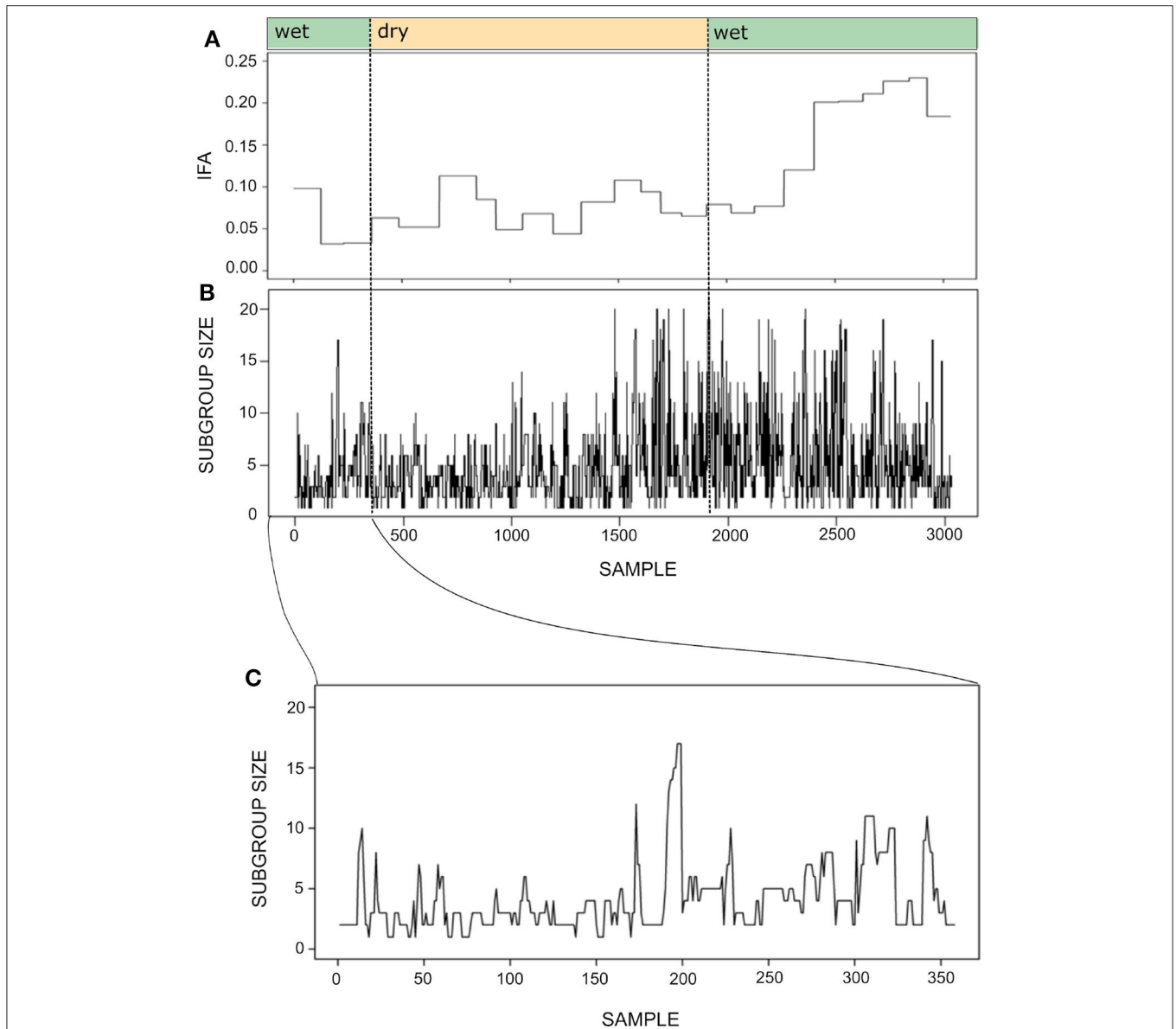
In previous work, the match between the collective computation output and the environment was evaluated using mutual information (Brush et al., 2018). Here we use transfer entropy:

$$T_{x \rightarrow y}(t) = H(y_t|y_{t-1}) - H(y_t|y_{t-1}, x_{t-1}) \quad (6)$$

This is a measure of how much uncertainty in a variable  $y$  is reduced given past states of both  $y$  and a variable  $x$  that is assumed to be independent of  $y$ . This dependence is over and above the uncertainty about  $y$  reduced by consideration of its own past state. Here transfer entropy is measuring how much subgroup size uncertainty is reduced by considering past states of subgroup size and IFA, conditioned on the uncertainty reduction by the past states of subgroup size alone. Given the difference in time resolution for the two time series (**Figure 7**), this implies that, within a given bi-weekly period, we are measuring the transfer entropy between a constant value of IFA and varying values of subgroup size. We used the JIDT package (Lizier, 2014) in R (R Core Team, 2017) to estimate the transfer entropy between time series, using the Kraskov estimator with the number of closest neighbors  $k = 4$ . The two observed time series have a  $T_{IFA \rightarrow SGS}(t) = 0.036$  nats.

To explore whether spider monkeys collectively compute a subgroup size distribution that is a good match to the distribution



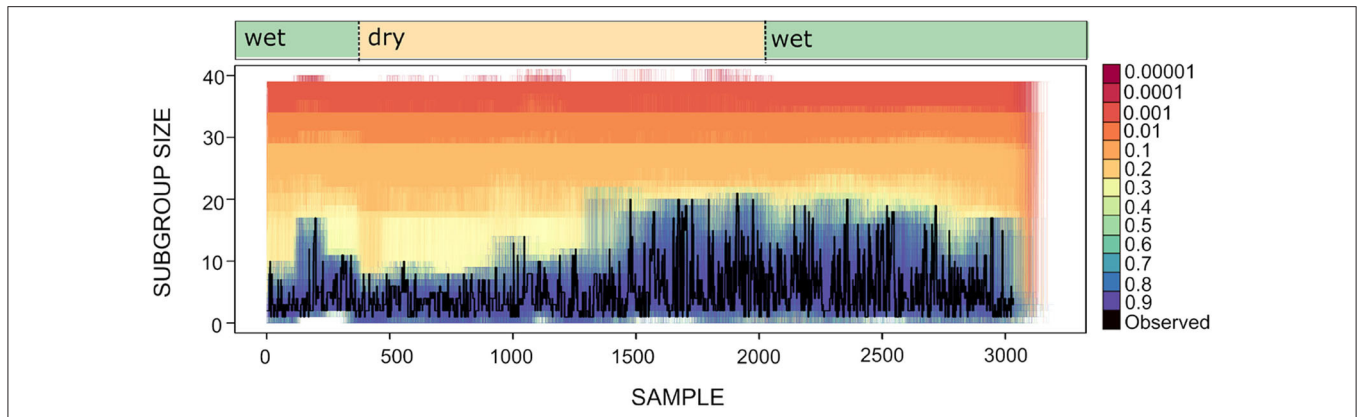


**FIGURE 7 |** Time series for the index of food abundance (IFA; **A**) and subgroup size (**B**). The IFA measures the overall abundance of fruit in the spider monkey's habitat, considering their most preferred species, their fruiting status and the abundance and relative size of trees (see section 6). The temporal resolution of the subgroup size data is 20 min, whereas food abundance was monitored biweekly. Thus, the IFA series has the same value throughout a given biweekly period, while subgroup size fluctuates at a much finer temporal scale. Noted above are the seasons (wet or dry) to which each sample belongs. Panel (**C**) presents a fragment of the subgroup size time series showing its variation between September 30 and October 31st 2013. Note that the time series was constructed with sets of scan samples taken every 20' collected throughout 4–8 h periods and that subgroups followed in consecutive days were not necessarily the same. Therefore, the spikes and drops observed in the curve do not always reflect fission or fusion events.

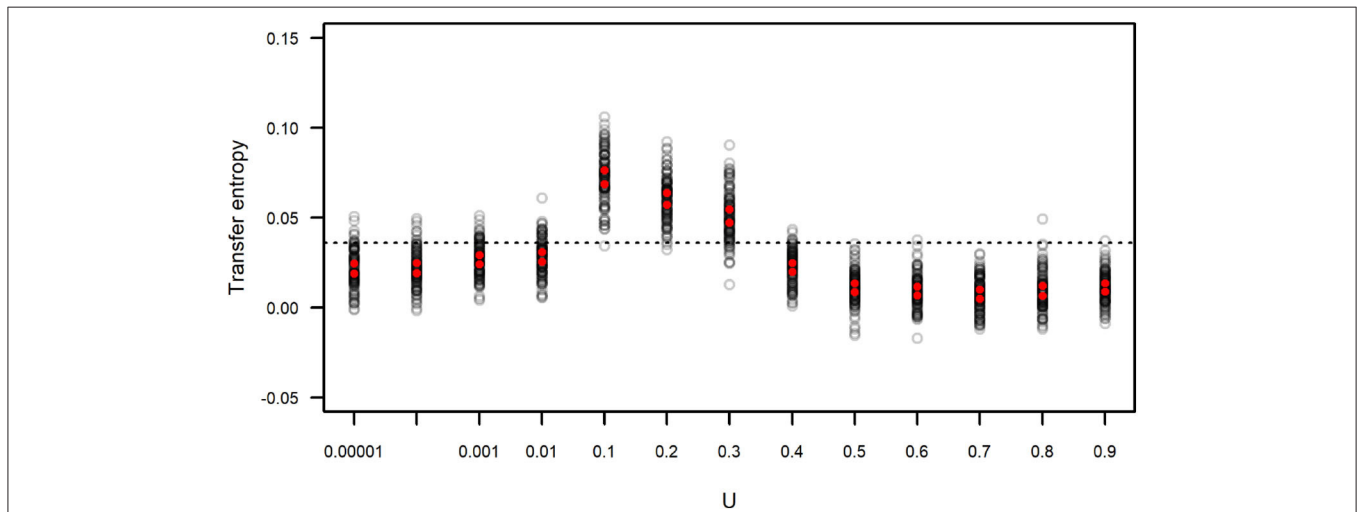
of fruiting trees, we assess which of our circuits with different strategy integration rules (described in section 4), computes a distribution of subgroup size that is a good fit to the current abundance of fruiting trees. Shown in **Figure 8** is the time series for the subgroup size values together with the subgroup size time series of all simulated data sets generated for different values of  $U$ . **Figure 8** shows what was already apparent in the subgroup size distributions shown in **Figure 5**, but in the

form of a time series: simulated data sets with  $U \geq 0.4$  generate a subgroup size distribution that is closest to the observed distribution.

We calculated the transfer entropy between the IFA time series and its corresponding subgroup size time series. We generated simulated data sets that included the same values of IFA as in the original dataset, but because the observation period length could vary (as the length of each observation period,  $n$ , was



**FIGURE 8 |** Time series for subgroup size as observed (black line) and simulated (lines of varying color). Each colored line corresponds to an instance of 100 simulations for different values of  $U$  and  $L = -0.00001$ . Wet and dry seasons are noted above.



**FIGURE 9 |** Transfer entropy between simulated IFA and simulated subgroup size. Each gray circle corresponds to an instance of 100 simulations run with varying values of  $U$  and  $L = -0.00001$ . Red dots indicate the upper and lower limits of 99 percent confidence intervals of the mean. The dotted line corresponds to the value of transfer entropy found for the observed IFA and subgroup size data in **Figure 7**.

sampled from the distribution of observed  $n$ ) there is a certain degree of variation around the observed data. Each simulated IFA series was compared to its corresponding subgroup size series. These values of  $T_{IFA \rightarrow SGS}(t)$  are presented in **Figure 9**, which also shows the value of  $T_{IFA \rightarrow SGS}(t)$  obtained for the observed IFA and subgroup size time series (**Figure 7**). The results suggest simulated subgroup size data sets with  $0.01 < U < 0.4$  match the temporal variation in IFA values better than the empirically observed subgroup size distribution and better than the simulated distributions computed with  $U \geq 0.4$ .

## 7. DISCUSSION

Social structure typically changes slowly compared to the interactions giving rise to it. As such, social structure, whether optimal for the environment or not, reduces uncertainty about the future state of the system and provides a relatively stable

background against which individuals can tune their own strategies (Flack, 2017a). Hence there are two challenges for a group computing its social structure: that it changes slowly enough to remain informative for decision-making and that it adaptively tracks the environment.

Frugivorous spider monkeys are faced with two significant sources of uncertainty related to foraging—to discover the location of fruiting trees and to distribute themselves over these fruiting trees to minimize conflict (Aureli et al., 2008) and the costs associated with large groups (Asensio et al., 2009), as well as to maximize resource intake (Symington, 1988). We have used a theory of collective computation (see references in the introduction) to explore how fission-fusion dynamics arises in spider monkey groups and whether the resulting distribution of subgroup size is a good match to the environment. We found spider monkey collectives appear to be able to partially match subgroup size to resource abundance. Our results suggest

however that the collective computation of subgroup size is not optimal with respect to food availability as measured by our index.

In simulating the circuits of subgroup-joining strategies we discover values of a sensitivity parameter  $U$  (a measure of the degree of consensus among the incoming weights required for an individual to make a decision about whether to stay or go) leading to a distribution of subgroup size that is a better match (than the observed distribution of subgroup size) to the observed abundance of fruiting trees. This suggests collective computation is under constraint and the system is experiencing adaptive lag—that is, still learning the best collective strategy to integrate information accumulated by group members. The deviation might instead be spurious—an outcome of (1) the way in which we calculate the food abundance index, (2) the fact that the data used to construct the two distributions are noisy and have different time resolutions: food abundance was measured at a bi-weekly scale while subgroup size was observed every 20 min, or (3) other factors besides social knowledge and relationships contributing to subgroup size decision-making.

We should also be cautious in interpreting the power of the collective computation at small  $U$  values. In these limits subgroups converge to a constant size where food abundance is expected to be somewhat predictive of size simply because both values remain constant during each bi-weekly period. These caveats aside, whereas collective computation in this system is not optimal, it remains nonetheless predictive and able to capture information about the environment. Specifically, the circuits that capture subgroup joining strategies can aggregate information about the environment. Although we did not study longer timescales, the slowly changing structure of groups provides a means for storing information accumulated by individuals about food availability across years (Palacios-Romo et al., 2019). With individuals that are more than 30 years old (see **Supplementary Information**), who are using spatial memory for their foraging decisions (Valero and Byrne, 2007), the information made available to the group through their experience is likely an important element to track long-term changes in the foraging environment.

Some means by which computations can be refined maximizing the match between group behavior and the abundance of food, includes individuals changing the way they accumulate information and/or compute strategies for staying or leaving, tuning how individuals integrate over those strategies, and tuning how the strategies interact in the circuit to produce subgroup size distributions. For example, are some individuals' strategies (perhaps because they influence many others) exerting a disproportionate effect on the output or do many individuals contribute in small ways? The problem of how collectives achieve optimal information processing is an important one in biology (Tkačik and Bialek, 2016), and near optimal information processing has been discovered in a number of biological systems (e.g., Petkova et al., 2019). However, these examples tend to be relatively simple developmental mechanisms such as segmentation during development of the fruit fly larval body plan. The circuit approach allows the question of tuning to obtain optimal information

processing to be addressed through simulation in more complicated systems.

Additional factors that could affect decision-making, thereby shifting the subgroup distribution from optimal to suboptimal, are a variety of social variables like sex and age, the previous history of interactions, and kinship relationships (Ramos-Fernández et al., 2009; Busia et al., 2017). However, because we are extracting individual strategies directly from the data, these modulating factors are already included in the weights between individuals. Other factors that are currently implicit include the risk of predation or location within the group's home range, which could also affect the subgroup size.

Our results shed light on how a group can best acquire and share information about patchy and dynamic environments. While individual foraging strategies based on spatial knowledge have been well-documented (Janson and Byrne, 2007; Fagan et al., 2013), group foraging strategies are less well-known outside of social insects (Gordon, 2016; cf. Gil et al., 2018). Exchanging information about available patches when foragers disperse and learning about the location and availability of different patches increases the foraging success of the whole group (Falcón-Cortés et al., 2019). The circuit of individual strategies that we infer here is, at least in part, a reflection of information sharing about available patches. Following another individual when ignorant is a simple mechanism of information sharing (Palacios-Romo et al., 2019), that could be reflected in the dyadic weights we have measured. This would lead to a fully connected circuit with information about food sources promoting a flexible grouping pattern that matches heterogeneity in the environment.

It is interesting to compare our approach to that of optimal foraging theory, which would postulate an optimal subgroup size distribution, based on a set of constraints and the best compromise between costs and benefits, which for most cases are unknown (Fretwell and Lucas, 1970; Stephens and Krebs, 1986). An empirical test of this postulate would consist of the match or lack thereof of the observed distribution to the food abundance and this would be interpreted in terms of the unknown mechanisms for how subgroup size comes about (e.g., Chapman et al., 1995). Our approach is more mechanistic: we observe a series of stay-leave decisions resulting from the interactions between individuals and construct a circuit of strategies that serves as a hypothesis for how the subgroup size distribution could emerge. We measure how similar these emerging distributions are to the observed and then test how well the time series matches the environmental variation. That we find alternative circuits that could produce a better match to the environment implies that the system is not necessarily constrained, as would be postulated by optimal foraging theory.

Mutual information, as a measure of uncertainty reduction, has some nice properties. It provides a robust way to study how near optimal a collective behavior is, and this provides a proxy for adaptiveness. We can also study different kinds of uncertainty reduction: an endogenous one, that involves collective computation of social structure that makes the world more predictable for individuals within a system (e.g., Brush et al., 2018); and an exogenous one, whereby collective computation produces social structure that encodes

knowledge about resource availability in the environment (this paper). Uncertainty reduction is consistent with a cost-benefit framework without requiring costs and benefits to be estimated. And quantification of the quality of the output of collective computation in information theoretic terms builds a technical bridge to Boltzmann's and von Neumann's ideas about the role of entropy in generating ordered states (Krakauer et al., 2020) that can form the basis of new levels of individuality, even at the social level.

In addition to assessing whether the output matches the environment, we studied the mechanics of collective computation. Previous work suggests spider monkeys preferentially follow food-aware individuals (Palacios-Romo et al., 2019). In the time series we find evidence in support of this result: we are able to extract significant (above-null) pair-wise probabilistic strategies used by individuals to decide to stay in or leave subgroups. Each individual had 20-30 strategies of varying strength (out of 46 possible). Generally the  $\Delta P$  were larger for "stay" strategies than "leave" strategies, suggesting possible food presence is a more important factor to spider monkeys than possible food absence. This emphasis on "attraction" might also be important for maintaining cohesion in fission-fusion dynamics in the context of a heterogeneous foraging environment with multiple alternative foraging options (Ramos-Fernández, 2005; Sueur et al., 2011). The strategies we find also recover well-known social patterns for *Ateles* spp., in particular—same sex based homophily for joining and repulsive tendencies between individuals of different sex (Fedigan and Baxter, 1984; Ramos-Fernández et al., 2009). It remains to be determined whether further, more fine-grained patterns like the frequency of dyadic interactions are also recovered by these strategies.

We used the extracted strategies to construct a family of circuits that vary in how individuals integrate these strategies to produce binary decisions to join or leave a subgroup. Individuals can have both repulsion (leave) and attraction (join) strategies. In previous work (DeDeo et al., 2010), strategies were passed through an AND or OR gate that captured conflict averse (all strategies have to say "go" to join a fight) and conflict prone dispositions (one "go" strategy was sufficient to join). Here we use thresholds. To recover the observed subgroup size distribution in simulation requires sums over strategies ( $\sum \Delta P \geq U = 0.4$ ) much larger than the strength of individual strategies (the majority of individual  $\Delta P$  values are below 0.05). This suggests individual-level decisions, as well as the aggregate output, require that individuals take into account relationships and social knowledge of many group members. If so, this would suggest that spider monkeys rely on social information from the wisdom of crowds (e.g., Jayles et al., 2017; Moreno-Gómez et al., 2017; Kao et al., 2018) to make decisions. These decisions are aggregated to collectively compute subgroup size distributions.

Mesoscale strategic circuits are summaries or average tendencies and therefore provide an economical way to process information. Slow variables, encoded in individual strategies, are compressed summaries of noisy interactions (Flack, 2017b). The idea that the mesoscale circuit is a compressed representation

of microscopic dynamics has parallels in multiplex networks, which have proven to be a better representation of the dynamics of many systems than the simple aggregation of different layers (De Domenico et al., 2015; Smith-Aguilar et al., 2019). Moreover, this way of compressing information may allow the social structure of spider monkeys to be flexible enough to track a dynamic environment, and, at the same time, be robust to disturbances. This has parallels to neural processing (Bassett et al., 2011; Daniels et al., 2017). As we have discussed elsewhere (see Brush et al., 2013, 2018; Daniels et al., 2017; Flack, 2017a) compression and related principles of collective computation have implications for engineered systems, such as web search and swarm robotics (e.g., Bonabeau et al., 1999; Seth, 2001; Young et al., 2013), as well as pattern recognition by artificial neural networks and human reputation networks.

How spider monkeys collectively compute fission-fusion social structure and how these computations can be tuned to realize adaptive variants raises many questions. Using longer time series, we could ask whether collective computation and fit to the environment are being refined and improved over time. With higher resolution data on strategies, and using methods from information theory (e.g., Rosas et al., 2019), it should be possible to quantify the degree to which the output is irreducibly encoded in the circuit as opposed to decomposable. Is social knowledge processed in a pairwise manner or do individuals perceive synergistic interactions among group members (e.g., does individual's A perception of individuals B and C contribute non-additively to its social knowledge)?

Understanding how a natural social system carries out adaptive computations could help to improve the performance of artificial systems. For instance, our results could provide insight into the mechanisms underlying learning through backpropagation in artificial neural networks. The way in which individuals adjust their strategic signaling in computing an appropriate power structure that feeds back to provide information about social interaction cost might be analogous to unsupervised learning (i.e., where the target is endogenous to the system) (Flack, 2017a; Brush et al., 2018). A system like the one we study here, with fission-fusion dynamics that can adjust to environmental conditions like the availability of fruiting trees, might be analogous to supervised learning (i.e., where the target is exogenous to the system). In both cases, feedback might share features with backpropagation in the strong and weak senses—the connection weights in the circuits/networks appear to be adjusted with a combination of vector (Brush et al., 2018) and scalar feedback (Flack et al., 2006) to minimize the network's error function when learning a task (Rumelhart et al., 1986; Lillicrap et al., 2020). This is just one of many exciting comparisons that could be made to better understand how different types of feedback, through tuning (Daniels et al., 2017) and downward causation (Flack, 2017a), shape the ability of the circuit to learn. And, as described in the Introduction, collective coarse-graining can produce a coherent mesoscale functioning as an information bottleneck, an ideal that is at least conceptually similar to the information bottleneck described by Tishby and colleagues to explain how deep neural networks



encode information parsimoniously (Tishby et al., 2000; Tishby and Zaslavsky, 2015; Flack, 2017a).

We have studied how a natural social system collectively computes. This is achieved through feedback among different scales of social organization, as proposed by Hinde's (1976) early paradigm and made explicit in Flack (2017a) and Flack (2017b). Studying collective computation should also find a range of different applications in the engineering of distributed, adaptive systems (Bonabeau et al., 1999).

## DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

## ETHICS STATEMENT

The animal study was reviewed and approved by the corresponding authorities in Mexico: the Direccion General de Vida Silvestre, Secretaria de Medio Ambiente y Recursos Naturales.

## AUTHOR CONTRIBUTIONS

All authors conceived the idea for this study. JF and DK developed the theory. GR-F and SS designed the data collection and performed the analysis and simulations. SS collected the data. All authors discussed the results and

contributed to the manuscript and giving approval to the final version.

## FUNDING

Data collection was aided by a grant from the Mexican 1431 Council of Science and Technology (CONACYT CB157656). JF thanks the Proteus Foundation and the Bengier Family Foundation for support during the project. JF also acknowledges JTF 60501/St. Andrews sub award 13337 for support during the project.

## ACKNOWLEDGMENTS

We thank the following people and organizations: Augusto, Macedonio, and Eulogio Canul for their invaluable assistance with data collection; Heiko Hamann, Deborah M. Gordon, and Matthew Lutz for their thorough review of a previous version of this article; Filippo Aureli and Colleen Schaffner for sharing the management of the field project. Instituto Politecnico Nacional and the Center for Complexity Science (C3-UNAM) for their logistical support.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frobt.2020.00090/full#supplementary-material>

## REFERENCES

- Asensio, N., Korstjens, A. H., and Aureli, F. (2009). Fissioning minimizes ranging costs in spider monkeys: a multiple-level approach. *Behav. Ecol. Sociobiol.* 63, 649–659. doi: 10.1007/s00265-008-0699-9
- Aureli, F., Schaffner, C. M., Boesch, C., Bearder, S. K., Call, J., Chapman, C. A., et al. (2008). Fission-fusion dynamics: new research frameworks. *Curr. Anthropol.* 49, 627–654. doi: 10.1086/586708
- Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J., Carlson, J. M., and Grafton, S. T. (2011). Dynamic reconfiguration of human brain networks during learning. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7641–7646. doi: 10.1073/pnas.1018985108
- Bonabeau, E., Dorigo, M., and Theraulaz, G. (1999). *Swarm Intelligence: From Natural to Artificial Systems*. Oxford: Oxford University Press.
- Brush, E. R., Krakauer, D. C., and Flack, J. C. (2013). A family of algorithms for computing consensus about node state from network data. *PLoS Comput. Biol.* 9:e1003109. doi: 10.1371/journal.pcbi.1003109
- Brush, E. R., Krakauer, D. C., and Flack, J. C. (2018). Conflicts of interest improve collective computation of adaptive social structures. *Sci. Adv.* 4:e1603311. doi: 10.1126/sciadv.1603311
- Busia, L., Schaffner, C. M., and Aureli, F. (2017). Relationship quality affects fission decisions in wild spider monkeys (*Ateles geoffroyi*). *Ethology* 123, 405–411. doi: 10.1111/eth.12609
- Chapman, C. A., Chapman, L. J., and Wrangham, R. W. (1995). Ecological constraints on group size: an analysis of spider monkey and chimpanzee subgroups. *Behav. Ecol. Sociobiol.* 36, 59–70. doi: 10.1007/s002650050125
- Chen, X., Randi, F., Leifer, A. M., and Bialek, W. (2019). Searching for collective behavior in a small brain. *Phys. Rev. E* 99:052418. doi: 10.1103/PhysRevE.99.052418
- Cheney, D. L., and Seyfarth, R. M. (1990). The representation of social relations by monkeys. *Cognition* 37, 167–196. doi: 10.1016/0010-0277(90)90022-C
- Cheney, D. L., and Seyfarth, R. M. (2008). *Baboon Metaphysics: The Evolution of a Social Mind*. Chicago, IL: University of Chicago Press. doi: 10.7208/chicago/9780226102429.001.0001
- Couzin, I. D., Krause, J., et al. (2003). Self-organization and collective behavior in vertebrates. *Adv. Study Behav.* 32, 10–1016. doi: 10.1016/S0065-3454(03)01001-5
- Daniels, B. C., Ellison, C. J., Krakauer, D. C., and Flack, J. C. (2016). Quantifying collectivity. *Curr. Opin. Neurobiol.* 37, 106–113. doi: 10.1016/j.conb.2016.01.012
- Daniels, B. C., Flack, J. C., and Krakauer, D. C. (2017). Dual coding theory explains biphasic collective computation in neural decision-making. *Front. Neurosci.* 11:313. doi: 10.3389/fnins.2017.00313
- De Domenico, M., Lancichinetti, A., Arenas, A., and Rosvall, M. (2015). Identifying modular flows on multilayer networks reveals highly overlapping organization in interconnected systems. *Phys. Rev. X* 5:011027. doi: 10.1103/PhysRevX.5.011027
- DeDeo, S., Krakauer, D. C., and Flack, J. C. (2010). Inductive game theory and the dynamics of animal conflict. *PLoS Comput. Biol.* 6:e1000782. doi: 10.1371/journal.pcbi.1000782
- Fagan, W. F., Lewis, M. A., Auger-Méthé, M., Avgar, T., Benhamou, S., Breed, G., et al. (2013). Spatial memory and animal movement. *Ecol. Lett.* 16, 1316–1329. doi: 10.1111/ele.12165
- Falcón-Cortés, A., Boyer, D., and Ramos-Fernández, G. (2019). Collective learning from individual experiences and information transfer during group foraging. *J. R. Soc. Interface* 16:20180803. doi: 10.1098/rsif.2018.0803

- Fedigan, L. M., and Baxter, M. J. (1984). Sex differences and social organization in free-ranging spider monkeys (*Ateles geoffroyi*). *Primates* 25, 279–294. doi: 10.1007/BF02382267
- Flack, J. C. (2012). Multiple time-scales and the developmental dynamics of social systems. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 1802–1810. doi: 10.1098/rstb.2011.0214
- Flack, J. C. (2017a). Coarse-graining as a downward causation mechanism. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 375:20160338. doi: 10.1098/rsta.2016.0338
- Flack, J. C. (2017b). “Life’s information hierarchy,” in *From Matter to Life: Information and Causality*, eds S. I. Walker, P. C. Davies, and G. F. Ellis (Cambridge: Cambridge University Press), 283–302. doi: 10.1017/9781316584200.012
- Flack, J. C., Erwin, D., Elliot, T., and Krakauer, D. C. (2013). “Timescales, symmetry, and uncertainty reduction in the origins of hierarchy in biological systems,” in *Evolution of Cooperation and Complexity*, eds K. Sterelny, B. Calcott, and R. Joyce (Boston, MA: MIT Press), 45–74.
- Flack, J. C., Girvan, M., de Waal, F. B. M., and Krakauer, D. C. (2006). Policing stabilizes construction of social niches in primates. *Nature* 439, 426–429. doi: 10.1038/nature04326
- Fretwell, S., and Lucas, H. (1970). On territorial behaviour and other factors influencing habitat distribution in birds. *Acta Biotheoretica* 19, 1–6. doi: 10.1007/BF01601953
- Gil, M. A., Hein, A. M., Spiegel, O., Baskett, M. L., and Sih, A. (2018). Social information links individual behavior to population and community dynamics. *Trends Ecol. Evol.* 33, 535–548. doi: 10.1016/j.tree.2018.04.010
- Gordon, D. M. (2016). The evolution of the algorithms for collective behavior. *Cell Syst.* 3, 514–520. doi: 10.1016/j.cels.2016.10.013
- Hein, A. M., Rosenthal, S. B., Hagstrom, G. I., Berdahl, A., Torney, C. J., and Couzin, I. D. (2015). The evolution of distributed sensing and collective computation in animal populations. *eLife* 4:e10955. doi: 10.7554/eLife.10955
- Hinde, R. A. (1976). Interactions, relationships and social structure. *Man* 11, 1–17. doi: 10.2307/2800384
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A.* 79, 2554–2558. doi: 10.1073/pnas.79.8.2554
- Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. U.S.A.* 81, 3088–3092. doi: 10.1073/pnas.81.10.3088
- Janson, C. H., and Byrne, R. (2007). What wild primates know about resources: opening up the black box. *Anim. Cogn.* 10, 357–367. doi: 10.1007/s10071-007-0080-9
- Jayles, B., Kim, H.-R., Escobedo, R., Cezera, S., Blanchet, A., Kameda, T., et al. (2017). How social information can improve estimation accuracy in human groups. *Proc. Natl. Acad. Sci. U.S.A.* 114, 12620–12625. doi: 10.1073/pnas.1703695114
- Kao, A. B., Berdahl, A. M., Hartnett, A. T., Lutz, M. J., Bak-Coleman, J. B., Ioannou, C. C., et al. (2018). Counteracting estimation bias and social influence to improve the wisdom of crowds. *J. R. Soc. Interface* 15:20180130. doi: 10.1098/rsif.2018.0130
- Krakauer, D., Bertschinger, N., Olbrich, E., Flack, J. C., and Ay, N. (2020). The information theory of individuality. *Theory Biosci.* 139, 209–223. doi: 10.1007/s12064-020-00313-7
- Krakauer, D. C., Flack, J. C., Dedeo, S., Farmer, D., and Rockmore, D. (2010). “Intelligent data analysis of intelligent systems,” in *Advances in Intelligent Data Analysis IX*, eds P. R. Cohen, N. M. Adams, and M. R. Berthold (Berlin: Springer), 8–17. doi: 10.1007/978-3-642-13062-5\_3
- Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., and Hinton, G. (2020). Backpropagation and the brain. *Nat. Rev. Neurosci.* 21, 335–346. doi: 10.1038/s41583-020-0277-3
- Lizier, J. T. (2014). JIDT: an information-theoretic toolkit for studying the dynamics of complex systems. *Front. Robot. AI* 1:11. doi: 10.3389/frobt.2014.00011
- Moreno-Gómez, S., Sorg, R. A., Domenech, A., Kjos, M., Weissing, F. J., van Doorn, G. S., et al. (2017). Quorum sensing integrates environmental cues, cell density and cell history to control bacterial competence. *Nat. Commun.* 8:854. doi: 10.1038/s41467-017-00903-y
- Palacios-Romo, T., Castellanos, F., and Ramos-Fernandez, G. (2019). Uncovering the decision rules behind collective foraging in spider monkeys. *Anim. Behav.* 149, 121–133. doi: 10.1016/j.anbehav.2019.01.011
- Petkova, M. D., Tkačik, G., Bialek, W., Wieschaus, E. F., and Gregor, T. (2019). Optimal decoding of cellular identities in a genetic network. *Cell* 176, 844–855. doi: 10.1016/j.cell.2019.01.007
- Pinacho-Guendulain, B., and Ramos-Fernández, G. (2017). Influence of fruit availability on the fission-fusion dynamics of spider monkeys (*Ateles geoffroyi*). *Int. J. Primatol.* 38, 466–484. doi: 10.1007/s10764-017-9955-z
- R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: <https://www.R-project.org/>
- Ramos-Fernández, G. (2005). Vocal communication in a fission-fusion society: do spider monkeys stay in touch with close associates? *Int. J. Primatol.* 26, 1077–1092. doi: 10.1007/s10764-005-6459-z
- Ramos-Fernández, G., Boyer, D., Aureli, F., and Vick, L. G. (2009). Association networks in spider monkeys (*Ateles geoffroyi*). *Behav. Ecol. Sociobiol.* 63, 999–1013. doi: 10.1007/s00265-009-0719-4
- Ramos-Fernández, G., King, A. J., Beehner, J. C., Bergman, T. J., Crofoot, M. C., Di Fiore, A., et al. (2018). Quantifying uncertainty due to fission-fusion dynamics as a component of social complexity. *Proc. R. Soc. B Biol. Sci.* 285:20180532. doi: 10.1098/rspb.2018.0532
- Ramos-Fernández, G., and Morales, J. M. (2014). Unraveling fission-fusion dynamics: how subgroup properties and dyadic interactions influence individual decisions. *Behav. Ecol. Sociobiol.* 68, 1225–1235. doi: 10.1007/s00265-014-1733-8
- Rosas, F. E., Mediano, P. A. M., Gastpar, M., and Jensen, H. J. (2019). Quantifying high-order interdependencies via multivariate extensions of the mutual information. *Phys. Rev. E* 100:032305. doi: 10.1103/PhysRevE.100.032305
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature* 323, 533–536. doi: 10.1038/323533a0
- Seth, A. K. (2001). Modeling group foraging: Individual suboptimality, interference, and a kind of matching. *Adapt. Behav.* 9, 67–89. doi: 10.1177/105971230200900204
- Smith-Aguilar, S. E., Aureli, F., Busia, L., Schaffner, C., and Ramos-Fernández, G. (2019). Using multiplex networks to capture the multidimensional nature of social structure. *Primates* 60, 277–295. doi: 10.1007/s10329-018-0686-3
- Sosna, M. M. G., Twomey, C. R., Bak-Coleman, J., Poel, W., Daniels, B. C., Romanczuk, P., and Couzin, I. D. (2019). Individual and collective encoding of risk in animal groups. *Proc. Natl. Acad. Sci. U.S.A.* 116, 20556–20561. doi: 10.1073/pnas.1905585116
- Stephens, D. W., and Krebs, J. R. (1986). *Foraging Theory*. Princeton, PA: Princeton University Press. doi: 10.1515/9780691206790
- Sueur, C., King, A. J., Conradt, L., Kerth, G., Lusseau, D., Mettke-Hofmann, C., Schaffner, C. M., et al. (2011). Collective decision-making and fission-fusion dynamics: a conceptual framework. *Oikos* 120, 1608–1617. doi: 10.1111/j.1600-0706.2011.19685.x
- Symington, M. M. (1988). Food competition and foraging party size in the black spider monkey (*Ateles Paniscus Chamek*). *Behaviour* 105, 117–132. doi: 10.1163/156853988X00476
- Tank, D., and Hopfield, J. (1988). Collective computation in neuronlike circuits. *Sci. Am.* 257, 104–114. doi: 10.1038/scientificamerican1287-104
- Tishby, N., Pereira, F. C., and Bialek, W. (2000). The information bottleneck method. *arXiv preprint physics/0004057*.
- Tishby, N., and Zaslavsky, N. (2015). “Deep learning and the information bottleneck principle,” in *2015 IEEE Information Theory Workshop (ITW)* (Jerusalem: IEEE), 1–5. doi: 10.1109/ITW.2015.7133169
- Tkačik, G., and Bialek, W. (2016). Information processing in living systems. *Annu. Rev. Condens. Matter Phys.* 7, 89–117. doi: 10.1146/annurev-conmatphys-031214-014803
- Valero, A., and Byrne, R. W. (2007). Spider monkey ranging patterns in mexican subtropical forest: do travel routes reflect planning? *Anim. Cogn.* 10, 305–315. doi: 10.1007/s10071-006-0066-z

Young, G. F., Scardovi, L., Cavagna, A., Giardina, I., and Leonard, N. E. (2013). Starling flock networks manage uncertainty in consensus at low cost. *PLoS Comput. Biol.* 9:e1002894. doi: 10.1371/journal.pcbi.1002894

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Copyright © 2020 Ramos-Fernandez, Smith Aguilar, Krakauer and Flack. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*