Male

N = 59
~21 x 10^6 data points

Female

N = 51
~18 x 10^6 data points
Wilcoxon rank sum test (p < .01)

Region-Normalised PDF

Female PDF - Male PDF

Mode #1

Mode #2

Mode #6

Mode #7

Female- preferred
Male- preferred
Cande et al, eLife, 2018
Cande et al, eLife, 2018
Fig. 84

Results.

This content downloaded from 128.112.66.66 on Sat, 16 Aug 2014 23:45:06 UTC

Richmond & Sachs, *Behavior*, 1980
Markov ($\tau=100$)  

Data ($\tau=100$)  

Data ($\tau=1,000$)

FIG. 2. Long time scale transition matrices and non–Markovian dynamics. (A) Markov model transition matrix for behaviors. Furthermore, the transition clusters are consistent with classical ideas of postural facilitation. Animals transition between locomotion gates systematically by other grooming behaviors of close-by body parts and the patterning of behaviors in the map is a conserved property of the system. In particular, the matrix appears modular, with most transitions out of similar behaviors [9, 20–22].

Complete characterization of the system requires the solution of the Markov process, which we achieve using an information bottleneck formalism (see below). Plotting time, as shown in more detail below.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.

The matrix $M$ does not display a nearly block-diagonal structure, $\mu$ is the eigenvalue, and $v$ and $u$ are the left and right eigenvectors, respectively, and $\lambda$ is the eigenvalue.
tiguous in the behavioral space, defining gross categories of behaviors. Furthermore, the transition clusters are consistent with classical ideas of postural facilitation and previous observations that transitions largely occur within similar phases of behavior and the patterning of behaviors in the map is a consequence of these transitions. The most elementary transitions are always equal to one, \( T_{ii} = 1 \), and the other eigenvalues have magnitudes less than one, \( \mu_j < 1 \), where \( j \neq i \). This is the noise floor, in this case the largest decay rate that we can resolve above the random structures present in our finite sample. When the behavior is Markovian then we can calculate the stationary distribution over states at long times. All flies, and thicknesses represent the standard error of the mean. Dashed lines are the predictions for the Markov model.

### FIG. 2. Long time scale transition matrices and non–Markovian dynamics.

- **A** Markov model transition matrix for \( T = 100 \), from Eq (3).
- **B** Transition matrices for all flies, and thicknesses represent the standard error of the mean. Dashed lines are the predictions for the Markov model.
- **C** Transition matrices for all flies, and thicknesses represent the standard error of the mean. Dashed lines are the predictions for the Markov model.

<table>
<thead>
<tr>
<th>Initial State</th>
<th>Final State</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle/Slow</td>
<td>Idle</td>
</tr>
<tr>
<td>Idle</td>
<td>Wing</td>
</tr>
<tr>
<td>Wing</td>
<td>Abdomen</td>
</tr>
<tr>
<td>Abdomen</td>
<td>Thorax</td>
</tr>
<tr>
<td>Thorax</td>
<td>Head</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Initial State</th>
<th>Final State</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle/Slow</td>
<td>Idle</td>
</tr>
<tr>
<td>Idle</td>
<td>Wing</td>
</tr>
<tr>
<td>Wing</td>
<td>Abdomen</td>
</tr>
<tr>
<td>Abdomen</td>
<td>Thorax</td>
</tr>
<tr>
<td>Thorax</td>
<td>Head</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Transitions</th>
<th>Eigenvalues of Transition Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 10^0 )</td>
<td>( 1 )</td>
</tr>
<tr>
<td>( 10^1 )</td>
<td>( \mu = 2 )</td>
</tr>
<tr>
<td>( 10^2 )</td>
<td>( \mu = 3 )</td>
</tr>
<tr>
<td>( 10^3 )</td>
<td>( \mu = 4 )</td>
</tr>
</tbody>
</table>

**Solid: Data**

**Dashed: Markov**

- \( \mu = 2 \)
- \( \mu = 3 \)
- \( \mu = 4 \)
- \( \mu = 5 \)
- \( \mu = 6 \)
- Random
Across Age Groups in Fruit Flies

Katherine Overman, Data from Daniel Choi & Joshua Shaevitz
Social bonding in prairie voles

FIRST TRIAD

SECOND TRIAD

REFERENCES:

RESULTS

ACKNOWLEDGEMENTS:
This work is supported by NIMH 1R01MH115831-01

HYPOTHESIS

• Predict the underlying neural dynamics driving animals’ movements
• Combine behavioral space mapping with electrophysiology.
• Apply methods to fully interacting voles without a barrier.

CONCLUSION

Animals varied across social context (e.g. the map of a male subject’s interaction with the partner is different from that of the male subject in the stranger trials). Between the two triads, we observed that behavioral maps for the same animal varied depending on the social context. The behavioral maps from two different triads correspond to distinct stereotyped behaviors and inferences about stereotyped behaviors during the sessions. This space provides a representation based on two triads of voles. This space provides a representation of the animals’ stereotyped behaviors during the sessions.

REFERENCES:

FUTURE DIRECTIONS

Behavioral Space Probability Density

Each colored region coarsely groups similar behavior patterns. Movements

Sena Agezo
Robert Liu
Larry Young
Transitions in human eCoG states

Yating Yang, Emory ’18 (now Cornell NBB)

Data From: Cory Inman & Jon Willie, Emory Neurosurgery

- Data
- HMM
- Random
- Markov

|λ_2|
---
1
0.8
0.6
0.4
0.2
0

Number of Transitions

10^0
10^1
10^2
10^3

Katherine Overman
Level of the "consummatory act"

- chasing
- biting
- threatening

- digging
- testing of materials
- boring
- gluing

- zigzag dance
- leading female to nest
- showing entrance
- quivering
- fertilizing the eggs

- care for offspring
- fanning
- rescuing eggs

etc.

reproductive instinct
Hierarchical organisation:
a candidate principle for ethology

RICHARD DAWKINS (1976)

THE NEED FOR GENERAL PRINCIPLES: SOFTWARE EXPLANATION
Accordingly, we can infer that the organism must have invested a great deal of non-randomness. This observation can be made more precise by looking at the contributions from the slowest time scale in the system, the slowest time scale in the system is determined by solving for the eigenvalue $\lambda$ which should be constant for a Markovian system. For very long times, there is an exponential drop in complexity.

At $t = 1$, the Markov model retains essentially no information, as demonstrated by the fact that $I(Z(t); X(t))$ is zero. At $t = 2 \times 10^3$, the complexity information is lost in the noise (Figure 2E). Similar patterns appear in higher modes, but we have more limited data to draw from.

The modular structure of the flies' transition matrix, which consists of 117 behavioral states, suggests that we might be able to group similar behavioral states into clusters that preserve much of the information about future actions (predictive information). Because we have made the process minimal by comparing the lack of structure within each fly and averaging, we find that structure extends beyond each fly or even $h$. Calculating these eigenvalues for $\mu < 0$ for all but the leading eigenvalue, we see that $\lambda$ loses all information about the current state in a Markovian system.

Because of this, we can simplify the process by looking at the leading mode, the apparent decay rate falls by nearly a factor of two orders of magnitude before the corresponding eigenvalue. For very long times, structure extends beyond the leading mode, the apparent decay rate, $\tau$, is given by $\tau = \frac{1}{\mu}$.

Calculating these eigenvalues for $\mu = 0.004$, we find that $\lambda(1)$ is 29 and $\lambda(2) = 100$, which are the eigenvalues that are closest to zero. If the contributions from the transitions have been randomized. Conversely, although the similarity between all of the rows, implying that all of the information is lost in the noise, we see that some of the block–diagonal structure from Fig. 1B has been dissipated, we see that structure extends beyond the leading mode, the apparent decay rate, $\tau$, is given by $\tau = \frac{1}{\mu}$.
V. PREDICTABILITY AND HIERARCHY

The partitioning, we optimize Equation 4 and plot the resulting complexity of \( I(Z(t); X(t)) \) fixed, we wish to maximize

\[
I(Z(t); X(t)) = \sum_{i,j} p(i,j) \log \frac{p(i,j)}{p(i) p(j)}
\]

where \( p(i,j) \) is the joint probability of being in states \( i \) and \( j \), \( p(i) \) is the probability of being in state \( i \), and \( p(j) \) is the probability of being in state \( j \).

Murta et al [25], is that if our representation is perfectly "treeness" metric, ever, by quantifying the degree of hierarchy in our rep-

hierarchically. It is important to note that these results happen is strong evidence that fly behavior is organized behav-

behaviors from many other clusters, That this does not result in any particular behavior. Thus, in a Markovian system, the slowest time scale in the system is deter-

and instead reflects the average probabilities of perform-

transitions have been randomized. Conversely, although the similarity between all of the rows, implying that all states that we do not directly observe, even though are required to model behavioral sequences, even in this range of time scales that we need a logarithmic axis for observing them.

\[
\tau = 100
\]

These results are direct evidence that many time scales are constant for a Markovian system. For \( \tau \) range of time scales that we need a logarithmic axis for great deal of non-randomness.

Calculating these eigenvalues for \( \tau \) transitions into the future is di-
ermanent, resulting in a characteristic decay time

\[
\| \mu > \| = 29
\]

for all but the leading eigenvalue, \( |\mu_1| \) for \( \tau = 100 \) to \( \tau = 1000 \) retrain a model.

We can make this di

increase in the height of these curves for small decrease in the height of these curves for small

further into the future. We also observe a relatively rapid

FIG. 4. Information bottleneck partitioning of behavioral space for t = 1 to 3

FIG. 5. Hierarchical organization for optimal solutions with

\( \# \text{ of Clusters} = 5 \)

\( \# \text{ of Clusters} = 6 \)

\( \# \text{ of Clusters} = 7 \)

\( \# \text{ of Clusters} = 25 \)

GJB et al, PNAS (2016)
As expected, the optimal curves move downward as the time lag increases, implying that the ability to predict the behavioral state of the animal decreases as we look further into the future. We also observe a relatively rapid decrease in the height of these curves for small $\tau$, followed by increasingly-closely spaced optimal curves as the lag length increases. It is this slowing that is indicative of the long time-scales in behavior.

Along each of these trade-off curves lie partitions of the behavioral space that contain an increasing number of clusters. We can make several observations about these data. First, in agreement with our investigation of the single-step transition matrix, we find that the clusters are spatially contiguous in the behavioral map as exemplified in Figure 4 for $\tau = 67$. Thus, even when we add in the long time-scale dynamics, we find that transitions predominantly occur between similar behaviors. Second, these spatially-contiguous clusters separate hierarchically as we increase the number of clusters, i.e. new clusters largely result from subdividing existing clusters instead of emerging from multiple existing clusters. One example of this can be seen in Figure 5, where the probability flow between partitions of increasing size subdivide in a tree-like manner. It is important to note that these results are not built in to the information bottleneck algorithm: we can solve the bottleneck problem for different numbers of clusters independently, and hence (in contrast to hierarchical clustering) this method could have found non-hierarchical evolution with new clusters comprised of behaviors from many other clusters. That this does not happen is strong evidence that fly behavior is organized hierarchically.

We can go beyond this qualitative description, however, by quantifying the degree of hierarchy in our representation as the number of clusters increases using a "treeness" metric, $T$. The idea behind this metric, which is similar to the one introduced by Corominas–Murtra et al [25], is that if our representation is perfectly hierarchical, then each cluster has precisely one "parent" in a partitioning with a smaller number of clusters. Thus, the better our ability to distinguish the lineage of a cluster as it splits through increasingly complex partitionings implies a higher value of $T$. More precisely, the treeness index is given by the relative reduction in entropy going...
We have measured the behavioral repertoires for dozens of flies, indicative of internal states that carry memory that persists thousands of transitions into the future. Thus, by finding optimally-partitions of the behavioral space, we were able to discover a perfect hierarchy.

The metric used to measure the hierarchy was defined as:

\[ T = \frac{\mathcal{H}_f - \mathcal{H}_b}{\mathcal{H}_f} \]

where \( T \) is the compression ratio, \( \mathcal{H}_f \) is the entropy of the full dataset, and \( \mathcal{H}_b \) is the entropy of the compressed representation.

The graph shows the compression ratio \( T \) as a function of the number of transitions in the behavioral space. Each line represents a different partitioning of the behavioral transitions, with colors indicating the number of transitions. The hierarchy was modeled, and our results suggest that such hierarchical neuroanatomical organization will be found in the brain with local circuitry within and emerging from the tor cortex, where groupings of neurons from millimeter up to centimeter scales are specific to tasks.

ACKNOWLEDGMENTS

The authors acknowledge the financial support of National Institutes of Health (GM098090, GM071508), The National Science Foundation, and The John Templeton Foundation. They also acknowledge the Aspen Center for Physics, where the ideas for this work were first discussed.

Drosophila melanogaster was maintained at 25°C on a 12:12 h light-dark cycle and fed standard cornmeal media. They were subsequently allowed 5 minutes for adaptation before the experiments. Behaviors were assigned by smoothing the embedded points and performing a watershed transform. For details, see the Methods section.

The authors also acknowledge the contributions of T. Dang, X. Tan, M. Seo, E. Lee, and J. Kim, who helped with the experiments and data analysis. They also thank A. Mayer, J. Bower, and P. Miller for access to the fly facility and the fly strains.

The authors declare no competing financial interests.
Fig. 8. The per minute rates of occurrence of seven behaviours as a function of four 'basic' slow processes, $v_1, \ldots, v_4$. The thickness of the line connecting behaviour $j$ with the slow process $v_k$ is proportional to the value $c'_{jk}$ in Table I. $c'$-coefficients with absolute values smaller than 0.05, which included all negative coefficients, were discarded.


RANDOM PROCESSES DESCRIBING THE OCCURRENCE OF BEHAVIOURAL PATTERNS IN A CICHLID FISH

BY WALTER HEILIGENBERG

Max-Planck-Institut für Verhaltensphysiologie, 8131 Seewiesen West Germany
As expected, the optimal curves move downward as the time lag increases, implying that the ability to predict the behavioral state of the animal decreases as we look further into the future. We also observe a relatively rapid decrease in the height of these curves for small $\tau$, followed by increasingly closely spaced optimal curves as the lag length increases. It is this slowing that is indicative of the long time-scales in behavior.

Along each of these trade-off curves lie partitions of the behavioral space that contain an increasing number of clusters. We can make several observations about these data. First, in agreement with our investigation of the single-step transition matrix, we find that the clusters are spatially contiguous in the behavioral map as exemplified in Figure 4 for $\tau = 67$. Thus, even when we add in the long time-scale dynamics, we find that transitions predominantly occur between similar behaviors. Second, these spatially-contiguous clusters separate hierarchically as we increase the number of clusters, i.e. new clusters largely result from subdividing existing clusters instead of emerging from multiple existing clusters. One example of this can be seen in Figure 5, where the probability flow between partitions of increasing size subdivide in a tree-like manner. It is important to note that these results are not built in to the information bottleneck algorithm: we can solve the bottleneck problem for different numbers of clusters independently, and hence (in contrast to hierarchical clustering) this method could have found non-hierarchical evolution with new clusters comprised of behaviors from many other clusters. That this does not happen is strong evidence that fly behavior is organized hierarchically.

We can go beyond this qualitative description, however, by quantifying the degree of hierarchy in our representation as the number of clusters increases using a "treeness" metric, $T$. The idea behind this metric, which is similar to the one introduced by Corominas-Murtra et al [25], is that if our representation is perfectly hierarchical, then each cluster has precisely one "parent" in a partitioning with a smaller number of clusters. Thus, the better our ability to distinguish the lineage of a cluster as it splits through increasingly complex partitionings implies a higher value of $T$. More precisely, the treeness index is given by the relative reduction in entropy going Idle
Slow
Locomotion
Anterior
Posterior
Side Legs

FIG. 4. Information bottleneck partitioning of behavioral space for $\tau = 67$ (approximately twice the longest time scale in the Markov model). Borders from the previous partitions are shown in black. For 25 clusters (bottom right), the partitions, still contiguous, are denoted by dashed lines.

FIG. 5. Hierarchical organization for optimal solutions with lag $\tau = 100$ ranging from 1 cluster to 25. The displayed clusterings are those that have the largest value of $I(Z; S(n+\tau))$ for that number of clusters. The length of the vertical bars are proportional to the percentage of time a fly spends in each of the clusters, and the lines flowing horizontally from left to right are proportional in thickness to the flux from the clustering on the left to the clustering on the right. Fluxes less than .01 are suppressed for clarity.

Data:

Network Output:
Idle

Locomotion

Anterior

Posterior

Abdomen

Leg Twitches

Slow
Maximum Likelihood Dimensionality Estimation (Levina & Bickel, NIPS, 2005)
Murta et al. [25], is that if our representation is perfectly
ric, which is similar to the one introduced by Corominas–
treeness” metric,
presentation as the number of clusters increases using a
ever, by quantifying the degree of hierarchy in our rep-
hierarchically.
behaviors from many other clusters, That this does not
non-hierarchical evolution with new clusters comprised of
bers of clusters independently, and hence (in contrast
we can solve the bottleneck problem for di
are not built in to the information bottleneck algorithm:
like manner. It is important to note that these results
between partitions of increasing size subdivide in a tree-
of this can be seen in Figure 5, where the probability flow
of emerging from multiple existing clusters. One example
largely result from subdividing existing clusters instead
as we increase the number of clusters, i.e. new clusters
these spatially-contiguous clusters separate hierarchically
predominantly occur between similar behaviors. Second,
in the long time-scale dynamics, we find that transitions
plified in Figure 4 for
are spatially contiguous in the behavioral map as exem-
bers of clusters. We can make several observations about these
of clusters. We can go beyond this qualitative description, how-
Along each of these trade-o
-0.5
# of Clusters = 5
# of Clusters = 1
# of Clusters = 6
# of Clusters = 2
# of Clusters = 3
# of Clusters = 4
# of Clusters = 7
# of Clusters = 25

FIG. 4. Information bottleneck partitioning of behavioral space for

Markov model). Borders from the previous partitions are shown in black. For 25 clusters (bottom right), the partitions, still

FIG. 5. Hierarchical organization for optimal solutions with

ters as it splits through increasingly complex partitionings
the better our ability to distinguish the lineage of a clus-

Locomotion
Anterior
Side Legs
Posterior

slow down as the
length increases. It this slowing that is indicative of the
by increasingly-closely spaced optimal curves as the lag

further into the future. We also observe a relatively rapid
contiguous, are denoted by dashed lines. More precisely, the treeness
implies a higher value of

index is given by the relative reduction in entropy going

ter as it splits through increasingly complex partitionings

suppressed for clarity.