Integrated information structure collapses with anesthetic loss of conscious arousal in *Drosophila melanogaster*

Integrated information structure collapses during anesthesia

Angus Leung[1,*], Dror Cohen[1,2], Bruno van Swinderen[3], Naotsugu Tsuchiya[1,2,4,5,*]
1 School of Psychological Sciences, Monash University, Melbourne, Australia
2 Center for Information and Neural Networks (CiNet), National Institute of Information and Communications Technology (NICT), Osaka, Japan
3 Queensland Brain Institute, The University of Queensland, Brisbane, Australia
4 Monash Institute of Cognitive and Clinical Neuroscience (MICCN), Monash University, Melbourne, Australia
5. Advanced Telecommunications Research Computational Neuroscience Laboratories, 2-2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288, Japan.

Corresponding authors: AL, NT
angus.leung1@monash.edu
naotsugu.tsuchiya@monash.edu

Abstract
The physical basis of consciousness remains one of the most elusive concepts in current science. One influential conjecture is that consciousness is to do with some form of causality, measurable through information. The integrated information theory of consciousness (IIT) proposes that conscious experience, filled with rich and specific content, corresponds directly to a hierarchically organised, irreducible pattern of causal interactions; i.e. an integrated informational structure among elements of a system. Here, we tested this conjecture in a simple biological system (fruit flies), estimating the information structure of the system during wakefulness and general anesthesia. We found that causal interactions among populations of neurons during wakefulness collapsed to isolated clusters of interactions during anesthesia. We used classification analysis to quantify the accuracy of discrimination between wakeful and anesthetised states, and found that informational structures inferred conscious states with greater accuracy than a scalar summary of the structure, a measure which is generally championed as the main measure of IIT. Spatially, we found that the information structures collapsed rather uniformly across the fly brain. Our results speak to the potential utility of the novel concept of an “informational structure” as a measure for level of consciousness, above and beyond simple scalar values.

Author summary
The physical basis of consciousness remains elusive. Efforts to measure consciousness have generally been restricted to simple, scalar quantities which summarise the complexity of a system, inspired by integrated information theory, which links a multi-dimensional, informational structure to the contents of experience in a system. Due to the complexity of the definition of the structure, assessment of its utility as a measure of conscious arousal in a system has largely been ignored. In this manuscript we evaluate the utility of such an information structure in measuring the level of consciousness in the fruit fly. Our results indicate that this structure can be more informative about the level of consciousness in a system than even the scalar summary proposed by the theory itself. These results may push consciousness research towards the notion of multi-dimensional informational structures, instead of traditional summaries.

Introduction

The question of how subjective, conscious experience arises from physical interactions has been pondered by philosophers for centuries [1,2], and now has moved into the domain of cognitive neuroscience [3–5]. Because we are only able to experience our own individual consciousness, exact inference of others' conscious contents (i.e., what it is like to be a bat [1]) seems intractable. However, broader inference on levels of consciousness, ranging from low during coma and deep anesthesia to high in wakeful states seems possible across animals. Behaviors of animals, ranging from humans to insects, all seem to change in a similar manner from highly active wakefulness with marked high-level cognitive capability to loss of consciousness with negligible cognitive functions. Indeed, such inferences have been widely accepted across various losses of consciousness in brain damaged patients [6] and non-human mammals [7,8], and are now becoming possible to apply to insects [9–11].

One influential view is that consciousness arises from “self-existence”, as measured through “integrated information” [12]. Integrated information, distinct from the standard notion of Shannon information [13], is defined as “differences that make a difference within a system” [14–16]. In other words, integrated information is concerned with how elements of a system causally influence each other such that information is accessible to the system itself (extrinsic information, conversely, concerns how states of a system casually influence states of another, separate system; see supporting information in [17]). Integrated information theory (IIT; [18,16,19]) proposes that integrated information is critical for consciousness to arise.

IIT provides a mathematical quantification of integrated information. Specifically, IIT describes how hierarchically organised elements uniquely and causally interact with other elements within a system in an integrated manner to produce information accessible to the system itself. According to IIT, the “maximally irreducible conceptual structure” [16] is hypothesised to directly correspond to the quantity and quality of consciousness. That is, the richer and more specific the informational structure of the system, the higher the level of consciousness in a system, and the richer the contents that the system consciously experiences.

While IIT offers a compelling theoretical account linking integrated information and consciousness, empirical tests of the theory remain rare [20]. Thus, whether *empirically* estimated integrated information structures relate to conscious arousal remains largely unknown. Here we bridge this gap by addressing the following empirical questions. First, how can we estimate informational structures from neural activity? Second, does manipulating the level of consciousness using general anesthesia shrink the estimated informational structures as compared to wakefulness? And third, can estimating the informational structures give us any biological hints as to the “interesting” regions of the brain that correlate with the level of consciousness? We address these questions by analyzing neural recordings from a simple biological system (fruit flies), collected during wakefulness and isoflurane anesthesia [11,21].

We propose a novel theoretical construct, which we call “integrated information structures” (IIS) to capture both the contents and level of consciousness in a system with reasonable computation time. We found that these collapsed during anesthesia. Critically, they were better at classifying conscious states than with a scalar summary of the structure, a measure which is usually championed as “integrated information” in IIT, with their collapse occuring all throughout the fly brain. Our results demonstrate the utility of information structure as a measure for level of consciousness, above and beyond simple scalar values, and open the door for improved clinical measures of consciousness.

Results

Constructing integrated information structures from fly local field potentials

In this paper, we construct integrated information structures (IIS), based on neural recordings from flies. According to the axioms of IIT, a set of structured and integrated

causal interactions among elements of the system, called a “maximally irreducible conceptual structure” in IIT [16], is hypothesized to be “identical” to the contents of conscious experience.

To construct the IIS, we used local field potentials (LFPs; hereafter referred to as “channels”) recorded from the fruit fly brain (Fig 1A; see Methods; [11]). LFPs were recorded using a linear multi-electrode array, such that the 15 channels covered both peripheral and central regions of the brain. We operationally defined the discrete state of each channel at each time by binarizing it with respect to the median voltage of that channel (Fig 1B).

Fig 1C-I illustrates the steps to estimate the IIS of two channels, A and B. From the empirically observed time course of the two discretized channels, we first construct a transition probability matrix (TPM; Fig 1C). The TPM characterizes how the whole system (A and B) evolves over time, containing all necessary information for unfolding how subsets of the system (A, B, and AB) causally interact to irreducibly specify the state of the whole system (AB).

Fig 1D considers how subset A’s current state (A=’0’ at time t) specifies the future state of any subset of the system AB (at time $t+\tau$). For brevity we will refer to the subset whose current status is analyzed (red circles in Fig 1D-G) as a “mechanism”, and the affected subset (white circles in Fig 1D-G) as a “purview”, following IIT terminology; [16]. Based on the TPM, we can compute a probability distribution over past and future purview states (the bar graphs in Fig 1D-H), given the current state of the mechanism in consideration; for example, the bar graph in Fig 1D shows that if mechanism A is in state ‘0’ at time t, AB is more likely to be ‘00’ or ‘01’ than ‘10’ or ‘11’ at time $t+\tau$.

Fig 1E illustrates the procedure to find “irreducibility” of the causal interaction from mechanism A to purview AB. To estimate how much the purview is irreducible, or uniquely determined by integrative interactions between A and the purview (according to IIT’s integration axiom), we estimate probability distributions assuming that some causal interactions are disconnected (i.e. statistically noised). We quantify the degree of causal interactions by computing the distance between the two probability distributions (distance is measured using the earth mover’s distance, with probabilities being moved as “earth”; EMD; [22]). The distance between the full (Fig 1D) and disconnected distribution which best approximates (i.e. is closest to) the original full distribution (Fig 1E) quantifies integrated information φ. Here, φ of A on AB can be understood as the degree to which mechanism A generates information about purview AB, above and beyond independent parts. In Fig 1E, the disconnection from A to B minimally affects the distribution out of all the possible cuts, giving φ of 0.0191.

Next, Fig 1F illustrates the identification of the purview over which A generates the most integrated information, as dictated by IIT’s exclusion axiom (the exclusion axiom in this context means that only the maximal information specified by A should be considered in order to avoid information being multiplied beyond necessity). The purview for which A generates the most information is referred to as A’s “core effect”. Here, mechanism A has a set of candidate purviews: A, B and AB. Based on the current state of A (‘0’), we repeat the

process of measuring distances between full distributions and disconnected distributions among all purviews (i.e. all subsets which are potentially affected by A). In this particular case, purview A is the core effect (φ of A on A is 0.0632, compared to φ of A on B and φ of A on AB both being 0.0191). Next, we perform similar operations on the TPM, but now looking at information the mechanism generates about a purview's *past*, instead of future. This is done to estimate the core *cause* of A. According to the intrinsic existence axiom of IIT, we consider A's overall influence (i.e. the information it generates for the system), to be the minimum of A's cause and effect.

Repeating the procedure (Fig 1D-F) for all candidate mechanisms (A, B, and AB), Fig 1G characterizes how all possible elements of the system specify the set of structured and integrated causal interactions, listing a full set of core causes and effects of all the mechanisms. The full set of distributions for all core causes and effects for all mechanisms, and their associated integrated information values is referred to in IIT as the minimally irreducibly conceptual structure (MICS). The MICS is purported to reflect the contents of conscious experience of a system.

Finally, Fig 1H explains how IIT arrives at a purported measure of level of consciousness, system-level integrated information Φ, through a system-level disconnection. The process of identifying core causes and effects for each mechanism is repeated after making unidirectional disconnections to the full system, in the same manner as disconnecting mechanisms from purviews. Once again, as there are many possible ways of disconnecting the system, we select the disconnection which best approximates the MICS of the full, whole system. In the case of the 2-channel system AB, the minimal disconnection is the disconnection from B to A. This disconnected MICS is used to assess system-level integrated information (for details, see Methods).

One difficulty with Φ is the high computational cost due to the combinatorial explosion of all possible system cuts. For this reason, we considered an alternative to Φ. Specifically, we assessed a set of φ values, which we term Integrated Information Structure (IIS; Fig 1I) as an alternative measure for discriminating the level of consciousness. A set of mechanism-level φ values are faster to compute, as they are already obtained as part of the computation of Φ.

As a first step to thorough understanding of the characteristics of Φ and IIS, we used every possible combination of 4 channels as a "system". This provides a good balance between spatial coverage for each set of channels and computation time. Fig 2 shows an example IIS obtained from 1 fly, 1 channel set during both wakefulness and anesthesia, when extending this process to the 4-channel case. Based on IIT, we tested two hypotheses in this paper: 1) system-level integrated information (Φ) should be reduced by general anesthesia, and 2) a set of mechanism-level integrated information (φ) values, the IIS, should also collapse during general anesthesia, reflected by reduced φ values for each mechanism (opposed to e.g. increased φ for some mechanisms). While IIT does not explicitly predict the latter, we reasoned that the level of consciousness should generally correlate with the richness of contents of consciousness.

System-level integrated information reduces globally due to general anesthesia

We first checked the prediction that system-level integrated information (Φ), IIT's proposed measure of level of consciousness, was reduced during anesthesia. Using linear mixed effects analysis (to account for intra-fly channel set correlations; see Methods), we indeed found system-level integrated information to be significantly reduced during anesthesia (Fig 3, Fig 4A-B; $t(35488) = -85.57$, $p < .001$).

IIT advocates searching for the channel set which gives the greatest system-level integrated information. Thus, we next looked at whether anesthesia's effect depended on spatial location of the channel sets. Given the linear arrangement of channels from our recording setup, we characterised two features of each channel set, 1) the average location of channels in the set (relative to channel 1, the most central channel) and 2) the distance among channels in the set (the sum of pairwise distances between channel labels; total path distance).

As central regions of the brain have been identified as responsible for integrating feedforward inputs from peripheral regions [23], we expected that more central channel sets would tend to have greater system-level integrated information. However, we found a weak trend towards the opposite, where peripheral channel sets tended to have slightly but significantly greater system-level integrated information ($\beta = 1.750 \times 10^{-2}$, $\chi^2(1) = 39.31$, $p < .001$). This trend was stronger during anesthesia, as indicated by a significant interaction between channel set location ($\beta = 2.613 \times 10^{-2}$, $\chi^2(1) = 43.80$, $p < .001$). Thus, centrally located channel sets seemed to be more affected by anesthesia, despite having less system-level integrated information than peripheral channels.

We also expected more "global" channel sets consisting of widely spaced channels to have greater system-level integrated information, reflecting integration across the whole brain. Indeed, we found system-level integrated information to increase with greater distance among channels ($\beta = 1.364 \times 10^{-3}$, $\chi^2(1) = 5.351$, $p < .021$), with the direction of the trend being reversed during anesthesia (significant interaction between anesthesia and channel distance; $\beta = -2.714 \times 10^{-3}$, $\chi^2(1) = 10.59$, $p < .001$). Thus, during wakefulness, the more global sampling of channels tended to yield larger system-level integrated information, while anesthesia disrupted this effect to some extent.

Overall, system-level integrated information was reduced regardless of spatial location or distance among channels. So, for analysis on the multi-dimensional IIS, we analysed all channel sets together without dividing into groups based on location or distance among channels.

Integrated informational structure collapses due to general anesthesia

We next compared integrated information (φ) for each mechanism during wakefulness to those during anesthesia. We found anesthesia to significantly reduce integrated information values across mechanisms ($\chi^2(1) = 3.092 \times 10^4$, $p < .001$; Fig 4C). Integrated information values also significantly depended on the size of the mechanism ($\chi^2(3) = 1.512 \times 10^5$, $p <$

.001). Further, there was a significant interaction between anesthesia and mechanism size ($\chi^2(3) = 1.203 \times 10^4$, $p < .001$), indicating that the extent to which integrated information was reduced due to anesthesia depended on mechanism sizes. We were surprised to find that 1-channel mechanisms had the greatest integrated information (*t*-test against 4-channel mechanisms: $t(88723) = 95.10$). Other than the unexpected 1-channel integrated information, we confirmed that the larger the size of the mechanism, the larger the integrated information (LME with two levels of mechanism size at a time, see Methods for details: 2-channel << 3-channel: $t(1.775 \times 10^5) = 141.63$, p<.001; 3-channel << 4-channel: $t(88723) = 178.16$, $p < .001$). A potential explanation for the unexpected result for 1-channel mechanisms is that 1-channel mechanisms are inherently irreducible to smaller parts. We return to this in the Discussion.

General anesthesia affects smaller mechanisms more than larger mechanisms

Given the significant interaction between anesthesia and mechanism size, we next investigated how the different mechanism sizes were differentially affected by anesthesia. We expected that integrated information for larger mechanisms (consisting of more channels) would be affected more by anesthesia than smaller mechanisms. This is because anesthesia is known to preferentially disrupt global communication [11,21], and so its effect should be reflected more strongly in larger mechanisms involving many channels. To further illustrate, consider two pairs of strongly connected neurons, [AB] and [CD], where there is a very weak connection between the two pairs (i.e., [AB]- -[CD]). In such a case, integrated information for both the 2-channel pairs ([AB] and [CD]) and the 4-channel mechanism ([ABCD]) could be high. If during anesthesia the connections between the pairs are disrupted, then 2-channel integrated information of the individual pairs could remain high while the overall 4-channel integrated information would reduce to zero.

To test if higher order mechanisms were more greatly affected by anesthesia, we first analyzed the degree of reduction in integrated information as a function of mechanism size. To account for the variation in integrated information among mechanism sizes, we compared the ratio of wakeful to anesthetized integrated information. A larger ratio corresponds to a larger decrease in integrated information due to anesthesia. We verified that the ratio of wakeful to anesthetized integrated information was also significantly different among mechanism sizes ($\chi^2(3) = 2.229 \times 10^4$, $p < .001$; Fig 4D). However, instead of finding higher order mechanisms to have larger relative reductions in integrated information due to anesthesia, we found the opposite - larger mechanisms had smaller relative reductions ($p < .001$ for all comparisons, $t(1.775 \times 10^5) = -106.8$, $t(1.775 \times 10^5) = -34.22$, and $t(88723) = -22.38$, for comparing 1-, to 2-, 2- to 3-, and 3- to 4-channel mechanisms respectively).

Integrated information structure better distinguishes conscious level than system-level integrated information

Given that the IIS indeed collapsed during anesthesia, we next sought to determine whether larger mechanisms give better discriminability of conscious level than smaller mechanisms. In other words, given that integrated information is reduced during anesthesia, is this decrease more reliable for larger mechanisms? We also sought to determine whether

considering the entire IIS allows for better discrimination conscious level than just consideration of single mechanisms, i.e. is the pattern of integrated information useful above and beyond considering independent integrated information values in isolation? As IIT proposes the scalar system-level integrated information value as the measure of conscious level (whereas the multi-dimensional IIS should represent experiential contents), we further compared this to the reliability of the decrease in system-level integrated information, where we expected system-level integrated information to achieve the best classification accuracy.

To compare the reliability of decreased integrated information, the collapse of the IIS, and decreased system-level integrated information, we used classification analysis. This allowed us to compare the reliability of one-dimensional changes of integrated information and system-level integrated information with multidimensional changes of the IIS. We used support vector machines (SVMs) to classify the conscious arousal level of individual epochs within each fly (within-fly classification, repeated for each fly; leave-one-paired-epoch-out cross-validation for each channel set; see Methods). To compare integrated information of different sized mechanisms, we averaged accuracies obtained across mechanisms of the same size.

We were able to discriminate wakefulness from anesthesia in the majority of channel sets, using either integrated information values or system-level integrated information (Fig 5A). Further, classification accuracy varied significantly depending on what measure was used (LME testing for main effect of mechanism size (1- to 4-channels), IIS and system-level integrated information; $\chi^2(5) = 1.300 \times 10^4$, $p < .001$). Unexpectedly, system-level integrated information performed slightly worse than integrated information of 1-channel mechanisms ($t(6823) = -5.204$, $p < .001$). 1-channel mechanisms classified with the highest accuracy ($t(13648) = 71.56$, $t(10918) = 38.26$, $t(6823) = 12.19$, compared to 2-channel, 3-channel, and 4-channel mechanisms respectively, $p < .001$ for all comparisons), consistent with 1-channel mechanisms having the largest decrease in integrated information due to anesthesia (Fig 4C). However, 2- and 3-channel mechanisms performed worse than 4-channel mechanisms ($t(9553) = -34.30$, and $t(6823) = -14.76$, $p < .001$) despite having larger relative decreases in integrated information due to anesthesia. Meanwhile, the full multi-dimensional IIS outperformed integrated information of individual mechanisms and system-level integrated information ($t(6823) = 52.02$ compared to individual 1-channel mechanisms, and $t(2728) = 65.11$ compared to system-level integrated information, $p < .001$), implying that the structure of integrated information may reflect quantity of consciousness better than the simple summary provided by system-level integrated information.

We also tested whether the reductions in integrated information were reliable across flies. This is important because in certain clinical contexts, such as traumatic brain injury, there may be no baseline measurements available, ruling out within-subject assessment. We repeated the decoding analysis, this time using leave-one-fly-out cross-validation. We found that the trend of results for discriminating wakefulness from anesthesia, among mechanism sizes, was similar to within-fly classification, though accuracies and performance differences were overall reduced (Fig 5B). As before, classification accuracy varied depending on what measure was used ($\chi^2(5) = 4451$, $p < .001$). As we found for the within-fly analysis, the IIS outperformed system-level integrated information, though to a lesser extent ($t(2728) =$

-7.198, $p < .001$). In contrast to the within-fly analysis, we found that (1) 1-channel integrated information only outperformed 2-channel integrated information ($t(13648) = 22.22$, $p < .001$), with 4-channel integrated information achieving greater accuracies than 1-, 2-, and 3-channel integrated information ($t(6823) = 15.74$, $t(9553) = 35.90$, and $t(6823) = 13.45$ respectively, $p < .001$), and (2) system-level integrated information outperformed the highest performing mechanism size, ($t(2728) = 9.815$, $p < .001$, compared to 4-channel integrated information).

Integrated information structure reliably collapses globally in space

Finally, we tested if the reliability of using the IIS to distinguish wakefulness from anesthesia depended on spatial features (Fig 5C-D). Similar to the trends for the raw system-level integrated information values previously, we found significant trends between classification accuracy and channel set location. Classification accuracy increased as channel sets moved closer to the central brain (with channel 1 being the most central in the brain), for both within- and across-fly classification (Table 1), and decreased as channels became more spaced out, also for both within- and across-fly classification. These same trends were present for classification when using the system-level integrated information values (Fig. 5E-F), though the trend of decreasing accuracy with more spaced out channels was not significant. Overall, using the IIS to discriminate level of consciousness in the fly brain yielded better classification accuracies, while maintaining the same spatial pattern of results as system-level integrated information.

Discussion

In this paper, we demonstrated the construction of integrated information structures (IIS), operationalised based on IIT 3.0, from real neural data to measure level of conscious arousal. We investigated how both system-level integrated information, the primary measure of conscious level put forward by IIT, and these information structures, consisting of a subset of the maximally irreducible conceptual structure (MICS) proposed by IIT as corresponding to the contents of phenomenal experience, varied with change in conscious level.

Other measures, inspired by IIT [24,8,25,26], have been previously tested in distinguishing conscious arousal states in neural data. However, rigorous assessment of IIT ultimately requires assessing its proposed measures, not proxies thereof. So far, empirical testing of IIT has been lacking in this regard. Instead, research has instead focussed on comparing varying operationalizations of system-level integrated information with regards to theoretical requirements [27,17,28–30] or specific network architectures [31–33]. Meanwhile, there are relatively few papers on testing system-level integrated information as a measure of consciousness in neural data [34]. Further, to our knowledge only one paper has empirically investigated the notion of information structures, but in the context of the correspondence between the structures and conscious content rather than conscious level [20].

Consistent with IIT's predictions, we found system-level integrated information to be reduced during anesthesia, and this was accompanied with the collapse of the information structures

as reflected by loss of integrated information across all mechanism sizes. Further, we found that the collapse in the information structure during anesthesia was more reliable than the reduction in system-level integrated information, allowing us to classify wake from anesthesia with greater accuracy than using the scalar measure. Finally, we found that both the reduction in system-level integrated information and the collapse of the information structures was fairly uniform across all the channel sets which we considered, as was the reliability of their respective reduction and collapse. Overall, these results demonstrate the utility of using information structures to assess conscious level, over a scalar measure such as system-level integrated information.

Why are 1-channel mechanisms more affected by anesthesia?

We encountered unexpected results with regards to 1-channel mechanisms. They had larger decreases in integrated information due to anesthesia (Fig 4C) and higher classification accuracy (Fig 5A) than larger mechanisms composed of 2, 3, or 4 channels. We had suspected that, as integrated information is supposed to measure information which is generated above and beyond separate parts, it would reflect the strength of long-range connectivity, which has been shown to be disrupted by isoflurane anesthesia, in humans, rats, and flies [35–37,21]. Given this background, we had expected that larger mechanisms, which are more likely to reflect long range connectivity, would be much more reduced and more reliable in classifying conscious states than 1-channel mechanisms, which we thought would primarily reflect the strength of self connections.

We see two ways of interpreting this. First, if we consider 1-channel mechanisms as providing information to the rest of the channel set, then disrupting communication among individual channels inevitably leads to disruption of larger mechanisms. A second interpretation is that the large decrease in integrated information for 1-channel mechanisms may primarily reflect disruption of strong self-connections present during wakefulness, rather than communication with other channels. Having said that, we note that 1-channel integrated information is not a well-developed theoretical construct. In fact 1-channel integrated information isn't defined for earlier versions of IIT and its approximations [18,38,27,17,29]. The bimodal distribution we observed for 1-channel mechanism integrated information (Fig 4C) may also be a consequence of this. Further theoretical investigation on how 1-channel mechanisms should be treated in IIT is necessary (see [39], for related discussion).

Why did spatial features have only small effects on system-level integrated information and the IIS?

Given the general literature on higher order, more centrally located brain areas being responsible for integration [23], we were surprised to find slightly greater system-level integrated information at more peripheral channels (Fig 3A). According to the known physiology of fly brains, we expected the strength of feedforward and feedback connectivity to be stronger towards the central areas of the brain [40,23]. By design, system-level integrated information should be stronger for those areas that have stronger feedforward and feedback connectivity as a whole. Taken together, we had expected that system-level integrated information to be stronger in the central brain, which was not the case.

Upon careful consideration, however, our finding may not necessarily be unexpected. One possible explanation would be the complexity of numerous brain structures in the centre of the brain versus the relative simplicity of fewer structures in the periphery [41]. Signals from a mix of many different structures may have cancelled each other out at the raw LFP level, which could have hampered our efforts in assessing the absolute level of system-level integrated information for more channel sets. Nonetheless, these central structures may have been more sensitive to the effects of anesthesia. Indeed, system-level integrated information tended to be more reliably reduced (Fig, 5A). Further investigation, especially with recording at a higher spatial resolution than LFP, such as optical imaging, may be able to resolve this issue.

Caveats of applying IIT empirically

We acknowledge several theoretical caveats with relation to system-level integrated information and the IIS. First, IIT requires that the TPM be constructed by perturbing the system into all possible states, instead of letting it naturally evolve over time as we did here. Second, IIT requires searching for the ideal spatial and temporal scales for describing the system, such that system-level integrated information is maximised [16,42,43,33], while here for simplicity and feasibility of computation we used only one spatial scale (that of LFPs), and one time scale. Thirdly, IIT requires identification of the “complex”, the set of parts which maximise system-level integrated information [44–46]. In addition to these core IIT requirements, we also arbitrarily operationalised the states of the system (through median-binarisation of voltages). These issues are related to practical limitations regarding empirical observation and manipulation at fine spatial scales as well as high computational cost, and thus apply to any empirical data for which IIT is being applied to and tested on.

Conclusion and future outlook

Our work opens up several future directions for empirically assessing mathematical approaches to consciousness, especially for IIT. It will be also important to test the generality of our finding across different modulations of consciousness, such as considering graded levels of anesthesia or sleep, as well as across datasets from different systems, such as in the more complex mammalian brain. While we investigated the use of information structures in determining conscious level, IIT links these structures more directly to contents of consciousness. Flies can demonstrate complicated behaviors, including attentional selection [47]. It would be an interesting future research to see if the contents of consciousness in flies, inferred from behaviors or under no-report paradigms [48], would correlate with the structures of integrated information as in humans [20].

Methods

Experimental procedure

Thirteen female laboratory-reared *Drosophila melanogaster* flies (Canton S wild type, 3-7 days post eclosion) were collected under cold anaesthesia and glued dorsally to a tungsten

rod. As the data have been published in [21], here we detail methods directly relevant to the dataset used presently.

Linear silicon probes with 16 electrodes (Neuronexus Technologies) were inserted laterally into the fly's eye. Probes had an electrode site separation of 25 μm. Recordings were made using a Tucker-Davis Technologies multichannel data acquisition system with a 25 kHz sampling rate. Isoflurane was delivered from an evaporator onto the fly through a connected rubber hose. Actual concentration near the fly body was either 0 vol% (awake condition) or 0.6 vol% (isoflurane condition). Flies in the awake condition responded to air puffs by moving their legs and abdomen, but were rendered inert under the isoflurane condition. Importantly, they regained responsiveness when isoflurane was subsequently removed, ensuring that flies were alive during the anesthesia recording.

An experiment, as conducted in [21] consisted of two blocks: one for the 0% isoflurane (air condition, followed by one for the isoflurane condition. Each block started with a series of air puffs, followed by 18 s of rest, 248 s of visual stimuli, another 18 s of rest, and finally a second series of air puffs. Isoflurane was administered immediately after completion of the first block (i.e. after the last air puff), and flies were left for 180 s to adjust to the new concentration before beginning the second block. We used data obtained in the 18 s period between the end of the first series of air puffs and the beginning of the visual stimuli.

Local field potential preprocessing

LFPs were downsampled to 1000 Hz from their original sampling rate of 25 kHz. Downsampled LFPs were bipolar re-referenced by subtracting neighbouring electrodes, resulting in 15 signals which we refer to as "channels". The 18 s of data for each condition was split into 2.25 s segments, giving 8 epochs of 2250 time-samples each. We removed line noise at 50 Hz using the function rmlinesmovingwinc.m function of the Chronux toolbox (http://chronux.org/; [49]) with three tapers, a windows size of 0.7 s, and a step size of 0.35 s. We then linearly detrended and *z*-scored each channel by removing the mean and dividing by the standard deviation across time samples within each epoch. Finally, we binarised voltages by taking the median voltage for each channel across all time-samples within a 2.25 s epoch, and then converting each time-sample in the epoch to 'on' if the voltage for that time-sample was greater than the median, and 'off' otherwise.

IIS computation

Data processing for computing the IIS and system-level integrated information was conducted using Python 3.6.0 in MASSIVE (Multi-modal Australian ScienceS Imaging and Visualisation Environment), a high-performance computing facility. We calculated the measures using PyPhi (version 0.8.1; [22]), publicly available from https://github.com/wmayner/pyphi. Complete details of all the calculations can be found in [16,22].

To compute the IIS, transition probability matrices (TPMs) describing how the set of channels evolves through time are required. To estimate these, we first select a set of *n*

channels of interest. As we used binary time-series, there are 2^n possible states for this set. For each channel X_n, we computed the empirical probability of being "on" at time $t+\tau$ given the state of the system at time t. In this manner, we computed the TPMs for every possible, 4-channel subset out of the 15 channels (15choose4 = 1365 channel sets), repeating this procedure for each fly and epoch. We selected 4 channels as this gave a reasonable balance between system-level integrated information and the IIS's strength of being a multivariate measure and their weakness of exponentially growing computation cost with system size (Mayner 2018).

To compute the IIS and system-level integrated information for a given set of 4 channels, we submitted its associated transition probabilities to PyPhi. Conceptually, PyPhi finds distances between the probability distribution of transitions specified by the full system with that of the disconnected system (Fig 1G-H). As there are 2^n possible states for a set of n-channels (16 states for 4-channels), we computed the IIS and system-level integrated information for every state. To obtain a single IIS and system-level integrated information value for each channel set, we averaged integrated information and system-level integration across all states, weighting by the number of occurrences of each state [31], and then averaged across epochs.

In Fig 1G-H, we explained system-level integrated information (Φ) as the sum of distances between cause and effect probability distributions specified by the full and (minimally) disconnected system (i.e. the full MICS and the disconnected MICS). Distances for each mechanism are weighted by the mechanism's φ value, as φ is the "earth" which is being moved from the full to the disconnected system. Any differences in φ values between the full and disconnected system, such as for mechanism AB in Fig 1G-H, are "moved" to the maximally uninformative distributions (red asterisks between Fig 1G-H). This weighted summation is depicted in between Fig 1G-H, in the solid box.

Classification analysis

To assess the reliability of the effects of anesthesia on the IIS, we conducted classification analysis, which allows us to holistically compare the full, multivariate IIS with single mechanism integrated information and system-level integrated information values. We trained and tested SVMs for each channel set using LIBLINEAR (using default options, i.e. L2-regularized L2-loss support vector classification (dual) [50]) at two levels: a) classifying epochs within each fly (within-fly classification, repeated for each fly), and b) classifying flies at each trial (across-fly classification, repeated at each epoch).

For each measure (integrated information of individual mechanisms, IIS, or system-level integrated information), we conducted nested leave-one-pair-out cross-validation [51,52]. At each outer validation, we conducted an inner-cross-validation procedure on 7 epoch-pairs / 12 fly-pairs (within-fly classification / across-fly classification; where a pair consists of one wakeful and one anesthetized epoch / fly), where we trained SVMs on 6 epoch-pairs / 11 fly-pairs at a time, and validated performance on the remaining epoch-pair / fly-pair. Training features (integrated information values or system-level integrated information values) were each z-scored before training, and testing features were z-scored using the mean and

standard deviation of the training set. This was repeated at different cost hyperparameters (2^{-50} to 2^{50}, in steps of powers of 10). We then trained a SVM on all 7 epochs / 12 flies at the cost hyperparameter value which gave the greatest validation performance (in cases of tie conditions, we took the lower cost value), and tested the overall classifier on the remaining fly / epoch. For the majority of validations (~74% for within-fly classification), the lowest cost of 2^{-50} was selected as the cost value. This process was repeated for each fly / epoch (within-fly classification / across-fly classification), and we averaged across repeats to obtain a final classification accuracy for the channel set and measure.

Statistical analyses

We used linear mixed effects analysis (LME; [53,54]) to test for significant differences. LME allows us to account for within-fly correlations among channel sets and avoid averaging across either channel sets or flies. Thus we always included random intercepts for fly and the interaction between fly and channel set as random effects, unless otherwise specified. To test for statistical significance of an effect, we employed likelihood ratio tests, where we compared the log-likelihood of the full model with a model with the effect of interest removed. As the likelihood ratio statistic is χ^2 distributed when two models are equivalent, we report the likelihood ratio statistic with the associated degrees of freedom ($\chi^2(d.o.f.)$) corresponding to the difference in number of coefficients between the full model with the model with the effect of interest removed, as well as the corresponding *p*-value. To conduct pairwise comparisons (e.g. to compare 1-channel to 2-channel integrated information), we limited the effect of interest to two levels at a time and report the two-tailed *t*-statistic and *p*-values associated with the regression coefficient.

We first employed LME to compare system-level integrated information, Φ, between wakefulness and anesthesia, using the following model (in Wilkinson notation [54]):

(1) Φ ~ condition + (1|fly) + (1|fly:set)

Where condition is level of conscious arousal (wake or anesthesia; dummy coded to be treated as a categorical variable), fly is individual flies (treated as a nominal variable), and set is channel set (treated as a nominal variable). As condition is the only fixed effect here and only has two levels, we report its *t*-statistic from the regression instead of the likelihood ratio statistic.

To test for a relationship between system-level integrated information (Φ) values and channel set location or total path distance among channels, we regressed system-level integrated information values onto channel set location and distance among channels:

(2) Φ ~ condition + location + distance + location:condition + distance:condition + (1|fly) + (1|fly:set)

Where “location:condition” and “distance:condition” denote interaction terms between channel set location and condition, and distance among channels and condition, respectively. We describe the relationship between Φ and channel set location or distance

among channels by reporting regression coefficients from *z*-scored Φ values in addition to the significance of the effect of location from the likelihood ratio test.

To compare integrated information (φ) values of the IIS between wakefulness and anesthesia and among mechanism orders, we used the model:

(3) φ ~ condition + size + condition:size + (1|fly) + (1|fly:set)

Where size is mechanism size (1, 2, 3, or 4, dummy coded to be treated as a categorical variable). The number of observations among mechanism sizes differed due to each order having a different number of possible mechanisms (4, 6, 4, and 1, respectively for 1-, 2-, 3-, and 4-channel mechanisms). The term “condition:size” denotes an interaction between level of conscious arousal and mechanism size.

When comparing classification accuracy across flies across the different feature types (i.e. 1-, 2-, 3-, and 4-channel φ, the IIS, and Φ), classification accuracy was not nested within fly, thus we only included random intercepts for each channel set:

(4) accuracy ~ feature + (1|set)

Where feature was dummy coded to be one of 1-, 2-, 3-, or 4-channel φ, the full IIS, or Φ.

To test for a relationship between classification performance and channel set location or distance among channels, we regressed accuracies onto the two spatial features:

(5) accuracy ~ location + distance + (1|set)

Where accuracy is classification accuracies, averaged across flies or epochs (for within-fly and across-fly classification, respectively). As previously for the relationship between Φ and the spatial features, we describe the relationship between accuracies and the spatial features by reporting regression coefficients on *z*-scored accuracies in addition to the significance of the effect of location from the likelihood ratio test.

Acknowledgements
We would like to thank Andrew Haun, Leonardo Barbosa, and William Mayner for their comments on our manuscript. This work was supported by computational resources provided by the Australian Government through MASSIVE under the National Computational Merit Allocation Scheme.

References

1. Nagel T. What Is It Like to Be a Bat? Philos Rev. 1974;83: 435–450. doi:10.2307/2183914
2. Chalmers DJ. The conscious mind: In search of a fundamental theory. Oxford University Press; 1996.
3. Dehaene S. Consciousness and the brain: Deciphering how the brain codes our thoughts. Penguin; 2014.

4. Lamme V. The crack of dawn: perceptual functions and neural mechanisms that mark the transition from unconscious processing to conscious vision. Open MIND. Frankfurt am Main: MIND Group; 2014.
5. Koch C, Massimini M, Boly M, Tononi G. Neural correlates of consciousness: progress and problems. Nat Rev Neurosci. 2016;17: 307–321.
6. Laureys S, Gosseries O, Tononi G. The neurology of consciousness: cognitive neuroscience and neuropathology. Academic Press; 2015.
7. Yanagawa T, Chao ZC, Hasegawa N, Fujii N. Large-scale information flow in conscious and unconscious states: an ECoG study in monkeys. PloS One. 2013;8: e80845.
8. Hudetz AG, Mashour GA. Disconnecting Consciousness: Is There a Common Anesthetic End-Point? Anesth Analg. 2016;123: 1228–1240.
9. Zalucki O, Van Swinderen B. What is unconsciousness in a fly or a worm? A review of general anesthesia in different animal models. Conscious Cogn. 2016;44: 72–88.
10. Shaw PJ, Cirelli C, Greenspan RJ, Tononi G. Correlates of sleep and waking in Drosophila melanogaster. Science. 2000;287: 1834–1837.
11. Cohen D, van Swinderen B, Tsuchiya N. Isoflurane Impairs Low Frequency Feedback but Leaves High Frequency Feedforward Connectivity Intact in the Fly Brain. eNeuro. 2018; ENEURO.0329-17.2018.
12. Tegmark M. Life 3.0: Being human in the age of artificial intelligence. Knopf; 2017.
13. Shannon CE. A mathematical theory of communication. Bell Syst Tech J. 1948;27: 379–423.
14. Bateson G. Steps to an ecology of mind: Collected essays in anthropology, psychiatry, evolution, and epistemology. University of Chicago Press; 1972.
15. Tononi G. Consciousness as Integrated Information: a Provisional Manifesto. Biol Bull. 2008;215: 216–242. doi:10.2307/25470707
16. Oizumi M, Albantakis L, Tononi G. From the phenomenology to the mechanisms of consciousness: integrated information theory 3.0. PLoS Comput Biol. 2014;10: e1003588.
17. Oizumi M, Amari S, Yanagawa T, Fujii N, Tsuchiya N. Measuring Integrated Information from the Decoding Perspective. PLoS Comput Biol. 2016;12: e1004654. doi:10.1371/journal.pcbi.1004654
18. Tononi G. An information integration theory of consciousness. BMC Neurosci. 2004;5: 42. doi:10.1186/1471-2202-5-42
19. Tononi G, Boly M, Massimini M, Koch C. Integrated information theory: from consciousness to its physical substrate. Nat Rev Neurosci. 2016;17: 450–461. doi:10.1038/nrn.2016.44
20. Haun AM, Oizumi M, Kovach CK, Kawasaki H, Oya H, Howard MA, et al. Conscious Perception as Integrated Information Patterns in Human Electrocorticography. eNeuro. 2017;4: ENEURO.0085-17.2017. doi:10.1523/ENEURO.0085-17.2017
21. Cohen D, Zalucki OH, van Swinderen B, Tsuchiya N. Local Versus Global Effects of Isoflurane Anesthesia on Visual Processing in the Fly Brain. eneuro. 2016;3: ENEURO.0116-16.2016. doi:10.1523/ENEURO.0116-16.2016
22. Mayner WG, Marshall W, Albantakis L, Findlay G, Marchman R, Tononi G. PyPhi: A toolbox for integrated information theory. PLoS Comput Biol. 2018;14: e1006343.
23. Barron AB, Klein C. What insects can tell us about the origins of consciousness. Proc Natl Acad Sci. 2016;113: 4900–4908.
24. Abásolo D, Simons S, da Silva RM, Tononi G, Vyazovskiy VV. Lempel-Ziv complexity of cortical activity during sleep and waking in rats. J Neurophysiol. 2015;113: 2742–2752.
25. Casali AG, Gosseries O, Rosanova M, Boly M, Sarasso S, Casali KR, et al. A Theoretically Based Index of Consciousness Independent of Sensory Processing and

Behavior. Sci Transl Med. 2013;5: 198ra105. doi:10.1126/scitranslmed.3006294
26. Sarasso S, Rosanova M, Casali AG, Casarotto S, Fecchio M, Boly M, et al. Quantifying Cortical EEG Responses to TMS in (Un)consciousness. Clin EEG Neurosci. 2014;45: 40–49. doi:10.1177/1550059413513723
27. Barrett AB, Seth AK. Practical Measures of Integrated Information for Time-Series Data. PLoS Comput Biol. 2011;7: e1001052. doi:10.1371/journal.pcbi.1001052
28. Oizumi M, Tsuchiya N, Amari S. Unified framework for information integration based on information geometry. Proc Natl Acad Sci. 2016;113: 14817–14822. doi:10.1073/pnas.1603583113
29. Tegmark M. Improved Measures of Integrated Information. PLoS Comput Biol. 2016;12: e1005123. doi:10.1371/journal.pcbi.1005123
30. Sevenius Nilsen A, Juel BE, Marshall W. Evaluating Approximations and Heuristic Measures of Integrated Information. Entropy. 2019;21: 525. doi:10.3390/e21050525
31. Albantakis L, Hintze A, Koch C, Adami C, Tononi G. Evolution of Integrated Causal Structures in Animats Exposed to Environments of Increasing Complexity. PLoS Comput Biol. 2014;10: e1003966. doi:10.1371/journal.pcbi.1003966
32. Haun A, Tononi G. Why Does Space Feel the Way it Does? Towards a Principled Account of Spatial Experience. Entropy. 2019;21: 1160. doi:10.3390/e21121160
33. Marshall W, Albantakis L, Tononi G. Black-boxing and cause-effect power. PLoS Comput Biol. 2018;14: e1006114. doi:10.1371/journal.pcbi.1006114
34. Kim H, Hudetz AG, Lee J, Mashour GA, Lee U, the ReCCognition Study Group, et al. Estimating the Integrated Information Measure Phi from High-Density Electroencephalography during States of Consciousness in Humans. Front Hum Neurosci. 2018;12: 42. doi:10.3389/fnhum.2018.00042
35. Ku S-W, Lee U, Noh G-J, Jun I-G, Mashour GA. Preferential Inhibition of Frontal-to-Parietal Feedback Connectivity Is a Neurophysiologic Correlate of General Anesthesia in Surgical Patients. PLoS ONE. 2011;6: e25155. doi:10.1371/journal.pone.0025155
36. Lee U, Ku S, Noh G, Baek S, Choi B, Mashour GA. Disruption of Frontal-Parietal Communication by Ketamine, Propofol, and Sevoflurane. Anesthesiology. 2013;118: 1264–1275. doi:10.1097/ALN.0b013e31829103f5
37. Alkire MT, Hudetz AG, Tononi G. Consciousness and Anesthesia. Science. 2008;322: 876–880. doi:10.1126/science.1149213
38. Balduzzi D, Tononi G. Integrated Information in Discrete Dynamical Systems: Motivation and Theoretical Framework. PLoS Comput Biol. 2008;4: e1000091. doi:10.1371/journal.pcbi.1000091
39. Albantakis L, Marshall W, Hoel E, Tononi G. What caused what? A quantitative account of actual causation using dynamical causal networks. Entropy. 2019;21: 459.
40. Pfeiffer K, Homberg U. Organization and Functional Roles of the Central Complex in the Insect Brain. Annu Rev Entomol. 2014;59: 165–184. doi:10.1146/annurev-ento-011613-162031
41. Shih C-T, Sporns O, Yuan S-L, Su T-S, Lin Y-J, Chuang C-C, et al. Connectomics-Based Analysis of Information Flow in the Drosophila Brain. Curr Biol. 2015;25: 1249–1258. doi:10.1016/j.cub.2015.03.021
42. Hoel EP, Albantakis L, Tononi G. Quantifying causal emergence shows that macro can beat micro. Proc Natl Acad Sci. 2013;110: 19790–19795. doi:10.1073/pnas.1314922110
43. Hoel EP, Albantakis L, Marshall W, Tononi G. Can the macro beat the micro? Integrated information across spatiotemporal scales. Neurosci Conscious. 2016;2016: niw012. doi:10.1093/nc/niw012
44. Hidaka S, Oizumi M. Fast and exact search for the partition with minimal information

loss. PLoS ONE. 2018;13. doi:10.1371/journal.pone.0201126
45. Kitazono J, Kanai R, Oizumi M. Efficient Algorithms for Searching the Minimum Information Partition in Integrated Information Theory. Entropy. 2018;20: 173. doi:10.3390/e20030173
46. Toker D, Sommer FT. Information integration in large brain networks. PLoS Comput Biol. 2019;15: e1006807. doi:10.1371/journal.pcbi.1006807
47. Van Swinderen B. Competing visual flicker reveals attention-like rivalry in the fly brain. Front Integr Neurosci. 2012;6: 96. doi:10.3389/fnint.2012.00096
48. Tsuchiya N, Wilke M, Frässle S, Lamme VAF. No-Report Paradigms: Extracting the True Neural Correlates of Consciousness. Trends Cogn Sci. 2015;19: 757–770. doi:10.1016/j.tics.2015.10.002
49. Mitra PP, Bokil HS. Observed Brain Dynamics. Oxford University Press; 2007.
50. Fan R-E, Chang K-W, Hsieh C-J, Wang X-R, Lin C-J. LIBLINEAR: A library for large linear classification. J Mach Learn Res. 2008;9: 1871–1874.
51. Tsamardinos I, Rakhshani A, Lagani V. Performance-estimation properties of cross-validation-based protocols with simultaneous hyper-parameter optimization. Int J Artif Intell Tools. 2015;24: 1540023.
52. Varoquaux G, Raamana PR, Engemann DA, Hoyos-Idrobo A, Schwartz Y, Thirion B. Assessing and tuning brain decoders: cross-validation, caveats, and guidelines. NeuroImage. 2017;145: 166–179.
53. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. J Stat Softw. 2015;67: 1–48. doi:10.18637/jss.v067.i01
54. Harrison XA, Donaldson L, Correa-Cano ME, Evans J, Fisher DN, Goodwin CE, et al. A brief introduction to mixed effects modelling and multi-model inference in ecology. PeerJ. 2018;6: e4794.

Figures

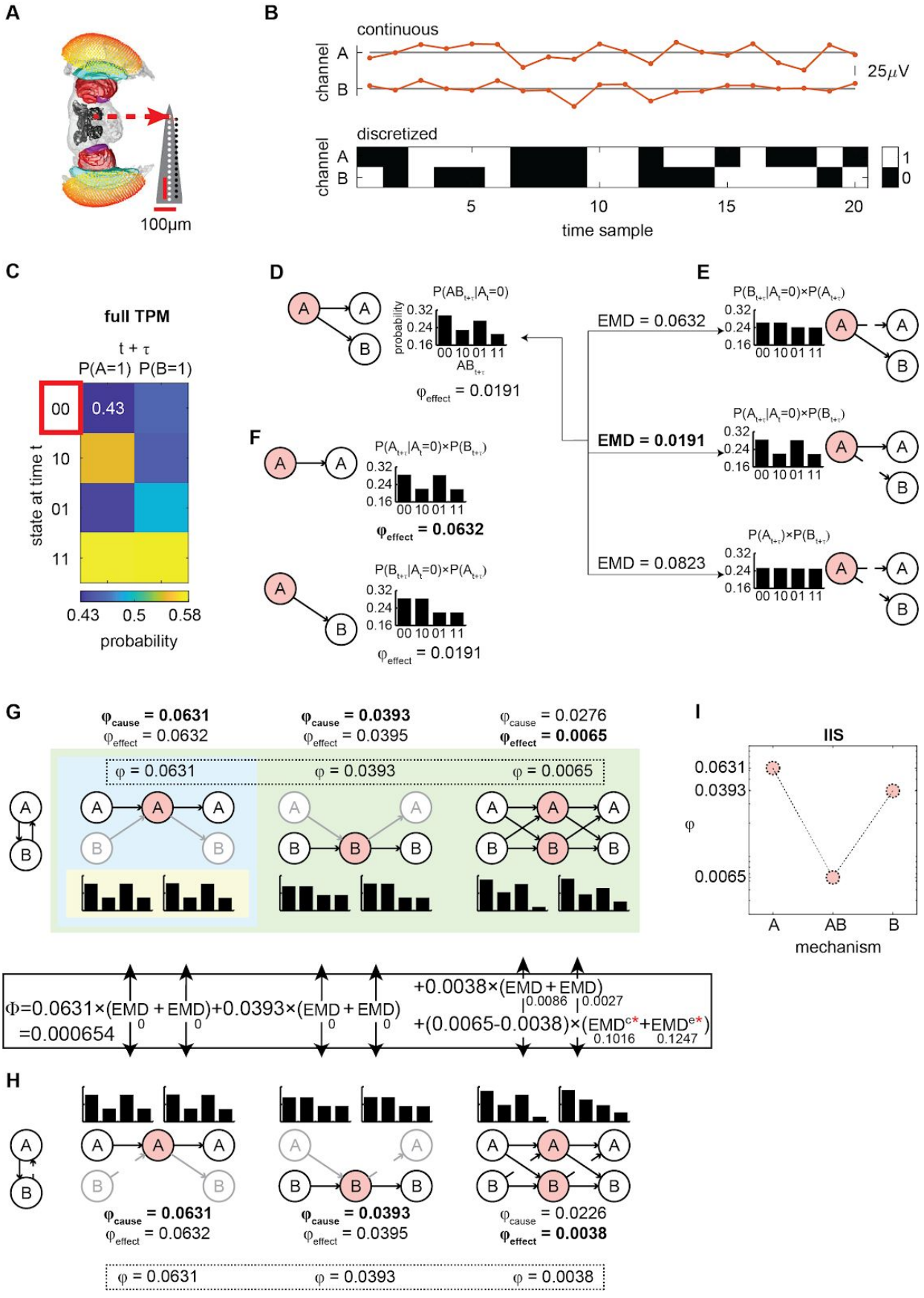

Fig 1. Summary of IIT and processing pipeline for computing the IIS from LFPs. (**A**) Multi-electrode probe recording of LFPs from the fly. (**B**) Continuous LFPs (red, top) are

discretized (black/white, bottom) by comparing to the median voltage for each trial. Displayed is an example of 20 samples for a set of two channels A and B. (**C**) A transition probability matrix (TPM) describes how the state of a system transitions from time *t* to time $t+\tau$ (τ = 4 ms). A full TPM is constructed by considering a state of all channels in the system (i.e., A and B) at time *t* and computing the probability of each channel's state at time $t+\tau$. For example, the top left entry of the full TPM is 0.43, which represents the probability of channel A being '1' at time $t+\tau$ given that channel A and B were both '0' at time *t*. (**D**) At a given state (e.g. A='0' and B='0' at time *t*, outlined in red in **C**, the effect information specified by a subset ("mechanism"; here A, in light red) over the future states of another subset ("purview"; here A and B, in white), is given by the probability distribution of the purview conditioned on the current state of the mechanism. (**E**) To compute integrated information (φ_{effect}) of mechanism A over purview AB, we find the minimal disconnections between the mechanism and the purview (indicated by broken arrows) which best approximate the original probability distribution. We compare the disconnected probability distributions to the original distribution using the earth mover's distance (EMD; treating probabilities as "earth" to be moved). We interpret the minimum EMD (bolded) as irreducible information generated over the purview by the mechanism (i.e. φ_{effect}). (**F**) We compute φ_{effect} for every possible purview (A, B, and AB as in **D**, with values 0.0632, 0.0191, and 0.0191 respectively), and select the purview and its associated probability distribution which gives the maximally integrated effect (bold). As probability distributions (bar graphs), we display the distribution over both channels A and B, assuming the maximum entropy distribution and independence on the channels outside of the purview. (**G**) φ_{cause} is determined in the same manner as φ_{effect}, except looking at possible past states of the purview (at $t-\tau$). Both φ_{cause} and φ_{effect}, and their associated probability distributions, are determined for every mechanism (A, B, and AB; left and right distributions are cause and effect probability distributions of the selected purviews; channels outside of the purview are greyed out). The overall φ generated by a mechanism is the minimum of φ_{cause} and φ_{effect} (bolded and in the dotted box). Yellow, blue, and green backgrounds indicate correspondence with the IIT terminology of "Maximally Irreducible Cause-Effect repertoire" (MICE), "concept", and "Maximally Irreducible Conceptual Structure" (MICS), respectively. (**H**) All φ values and associated probability distributions are re-computed for each possible uni-directional cut separating the channels into a feedforward interaction from one subset of channels to the remainder of the system. Broken lines here depict the cut removing channel B's input to A. System-level integrated information (Φ) is the sum of distances between cause and effect probability distributions specified by the full and (minimally) disconnected system, weighted by φ values from the full system (hence Φ is the minimum across all possible system level cuts; solid box between **G** and **H**, see Methods for details). Note that distances between 1-channel mechanisms were 0, not contributing to Φ, which we found to generally often be the case. (**I**) We take the φ values of each mechanism (within the dotted box in **G**) to form the integrated information structure (IIS).

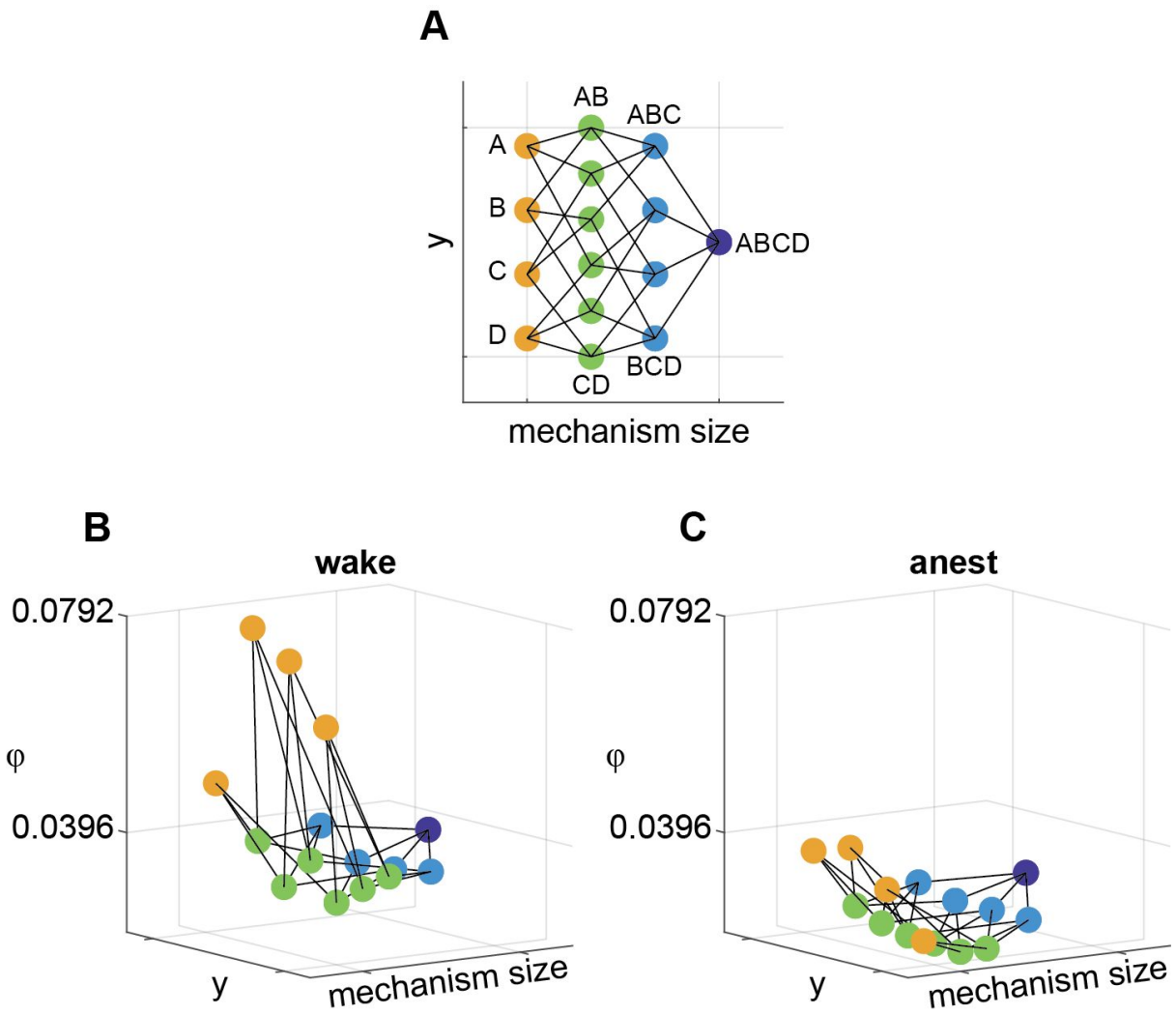

Fig 2. Integrated information structure (IIS) for one channel set for one fly in the awake state. (**A**) Top-down view of the IIS. Mechanism size refers to the number of channels that constitute each mechanism (yellow, green, light blue, and dark blue dots indicate mechanisms consisting of 1, 2, 3, and 4 channels respectively). The y-axis is arbitrarily set to give equal spacing between mechanisms. Lines indicate inclusion relations (e.g., mechanism AB consists of A and B). (**B**) An exemplar IIS from a single fly and channel set, during wakefulness. (**C**) An IIS from the same fly and channel set as in **B** during anesthesia. A 3D rotation video of the IIS is available at http://dx.doi.org/10.26180/5eb952457b48f.

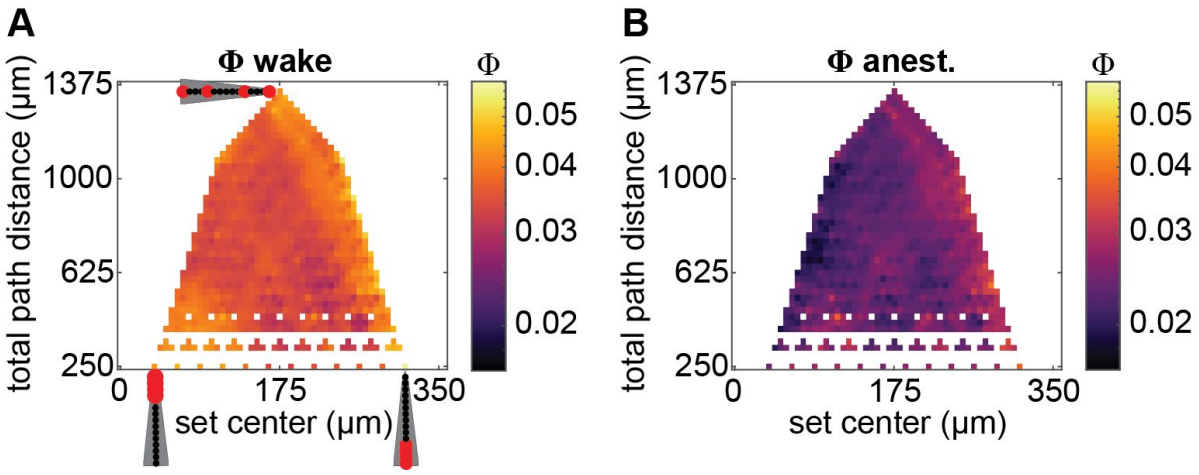

Fig 3. Spatial map of system-level integrated information Φ. (**A**) System-level integrated information Φ values during wakefulness, averaged across flies, as a function of average channel location relative to the position of the most central channel (x-axis; larger values indicate channel sets which on average are more peripherally located), and sum of pairwise distances between each pair of channels (total path distance; y-axis) within each channel set. Channel arrays, as in Fig 1A, indicate example locations of channels (in red) and their spacing along the two axes. Channel sets with identical centers and path distances were averaged. A subset of otherwise unfilled values in the map were linearly interpolated to reduce gaps in the map. (**B**) System-level integrated information Φ values during anesthesia.

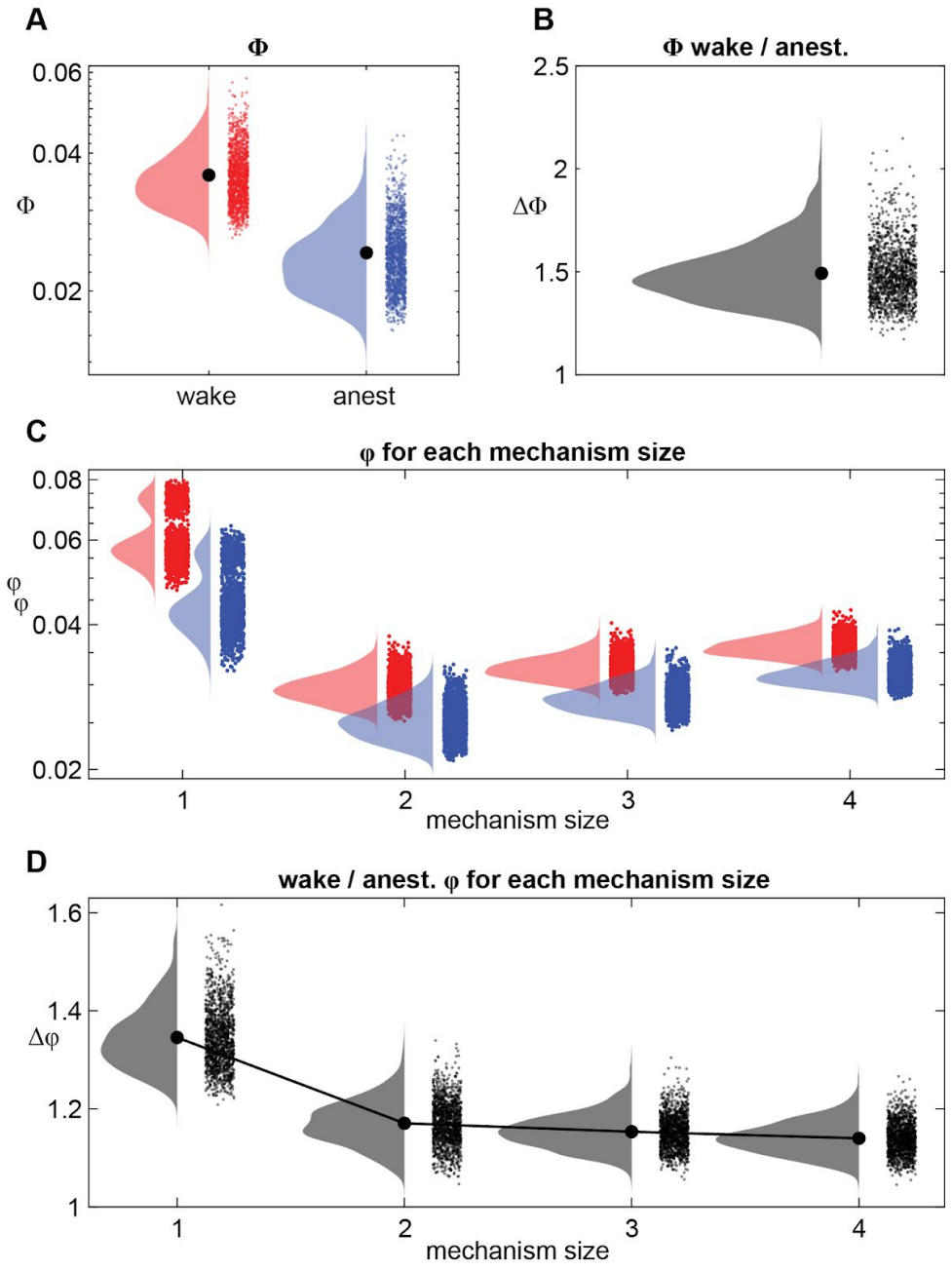

Fig 4. Effect of anesthesia on system-level integrated information (Φ) and the integrated information structure (IIS: a set of φ values). (**A**) Φ values during wakefulness (red) and anesthesia (blue) for each of 1365 channel sets, averaged across flies. (**B**) Ratio of Φ (wakeful / anesthetized), for all channel sets, averaged across flies. (**C**) φ values from the IIS for each mechanism size, for wake (red) and anesthesia (blue). We show the average ratio for each of 1365 channel sets averaged across flies for each mechanism size. (**D**) Ratio of wakeful φ to anesthetized φ (averaged across flies) for the full IIS.

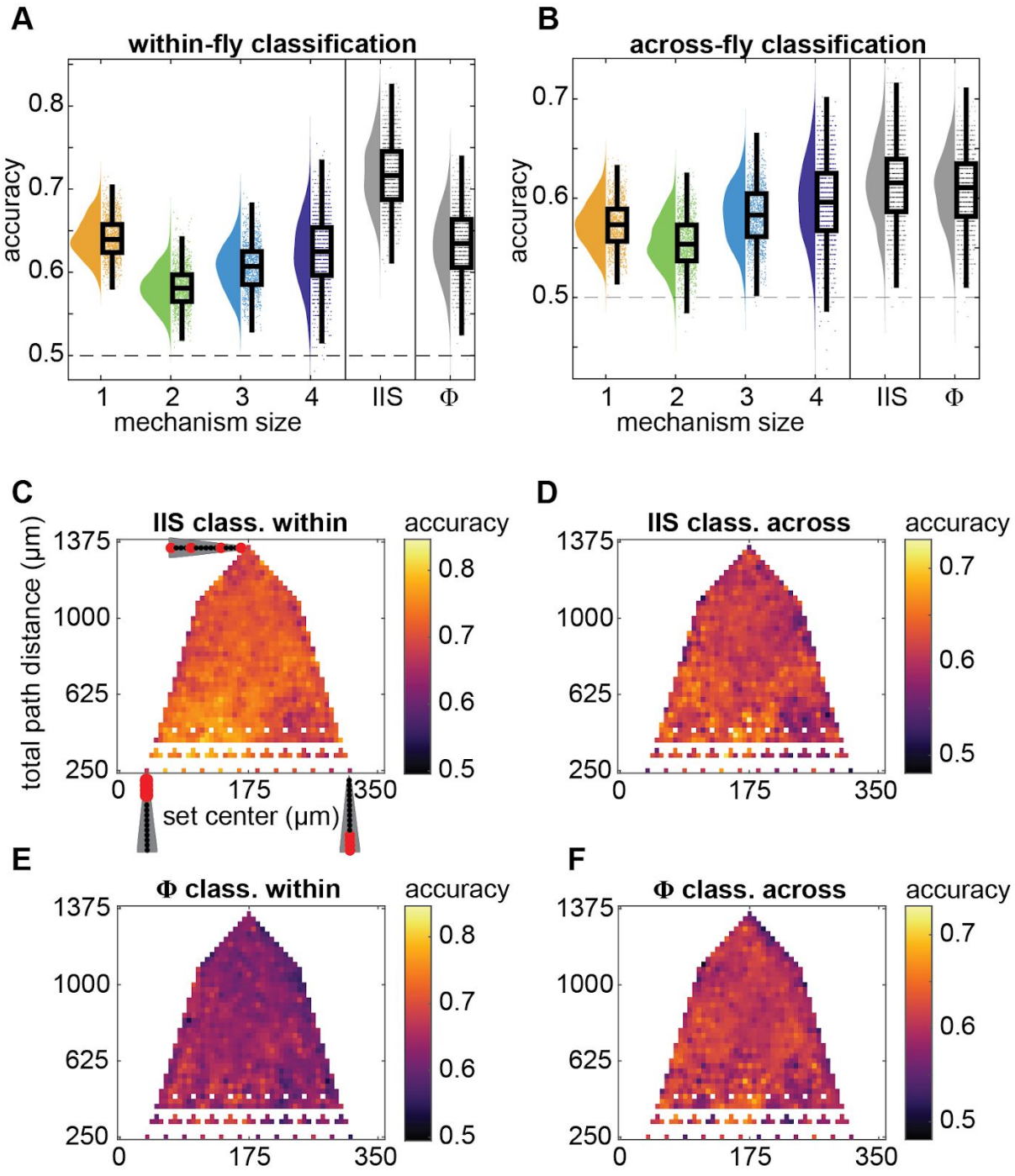

Fig 5. Classification of wakeful vs. anesthetized conditions using mechanism-level φ, system-level Φ, or the integrated information structure (IIS: a set of φ's) . (**A**) Within-fly and (**B**) across-fly classification at each individual channel set using each φ values for each mechanism size (orange, green, pale and dark blue are 1-, 2-, 3-, and 4-channel mechanisms, respectively; single-feature classification), when using the IIS (i.e. all mechanisms together, 15-feature classification), and when using Φ (single-feature classification). Individual points are classification accuracy of each channel set, after averaging accuracies across all mechanisms within the same mechanism order. Boxplots show median, 25th-75th percentiles, and whiskers are 1.5 interquartile below and above respectively. (**C**-**D**) Within-fly (**C**) and across-fly (**D**) classification accuracies when using the IIS. (**E**-**F**) Within-fly (**E**) and across-fly (**F**) classification accuracies when using Φ.

Tables

Table 1. Dependence of system-level integrated information and classification of wake/anesthesia on channel set location and distance among channels

	Location			Distance		
Regressand	**β**[b]	**χ^2(1)**[c]	***p***	**β**[b]	**χ^2(1)**[c]	***p***
Φ WF	-9.16	44.30	< .001	-2.10	52.11	< .001
Φ AF	-5.76	16.77	< .001	-0.47	2.56	.110
IIS WF	-7.89	32.71	< .001	-2.09	51.05	< .001
IIS AF	-9.06	42.03	< .001	-0.90	9.31	.002

Location, average location of channels in a channel channel set. Distance, sum of pairwise distances between channels within a channel set. Φ, system-level integrated information; WF, within-fly classification accuracy; AF, across-fly classification accuracy; IIS, integrated information structure.

[b] β from regressing *z*-scored Φ values or classification accuracies; values are $\times 10^{-2}$

[c] degrees of freedom always for likelihood ratio test = 1 (see Methods).