Decomposing Predictive Information in Social Dynamics

Akira Kawano[a], Liam O'Shaughnessy[b], Radmila Neiman[a], Kosmas Deligkaris[a], Luis Carretero Rodriguez[c], Ichiro Masai[c], Greg J. Stephens[a,b]
[a]*Biological Physics Theory Unit, OIST Graduate University, Okinawa 904-0495, Japan*
[b]*Department of Physics and Astronomy, VU University Amsterdam, 1081HV Amsterdam, The Netherlands*
[c]*Developmental Neurobiology Unit, OIST Graduate University, Okinawa 904-0495, Japan*

Social behaviors include some of the most interesting interactions in living systems yet their principled characterization remains unsolved. Here we suggest that at the core of social interactions is the notion of mutual prediction, which we analyze in the context of two male zebrafish engaged in a dominance contest. Using 3D velocity trajectories, we construct the mutual information between a two-animal past and one-animal future, and we quantify the redundant, unique, and synergistic components using partial information decomposition across time windows. We find that predictive information decomposition naturally aligns with important social concepts, such as mirroring and dominance. At contest end, we find asymmetries in self-unique and redundant information that reflect the emergent dominance relationship. Applied to *mecp2* zebrafish mutants, an autism model, we find that predictive information is reduced overall, but especially for synergistic flows, which is indicative of difficulties in more complex social dynamics.

Significance Statement

Social interactions are rich and diverse, ranging from mirroring to complementary actions. A unifying framework for defining and analyzing such interaction types has long been needed. Here, based on modern information theory, we formulate how the past state of interacting organisms encodes the future state of an individual. This framework provides a natural decomposition of pairwise social dynamics into mirroring, independent action, directed influence, and joint action. Applied to dominance contests in zebrafish, these modes of interaction capture distinct phases of conflict, their assessment strategies, and the resulting dominance relationships. Moreover, our analysis reveals a specific disruption in the social behavior of mutant zebrafish linked to autism, shedding new light on impairments in communication and social learning.

Introduction

Resulting from interactions between individuals, social behavior occurs both among and across species, from microorganisms to humans. These interactions are broadly important; cooperation is thought to have played a crucial role in the evolution of multicellularity [1], and human sociality is a primary determinant of mental and physical well-being (see e.g. [2]). Despite such universal importance and in part due to it's complexity, social behavior has been predominantly analyzed either qualitatively or with specifically crafted measures and experiments. These include the frequency and context of human-derived action patterns, as in an ethogram, and discipline-defined experimental protocols, as in social neuroscience (see e.g. [3]). Another approach is to account directly for the interactions through modeling, for example, by viewing the social system as a collection of interacting particles [4]. While this later approach is promisingly general, interactions can change abruptly in time, making estimation difficult. Also, the form of the interactions varies with the choice of behavioral coordinates, which can hinder interpretation.

An alternative view is to focus on social dynamics as a reciprocal flow of influence, with intertwined actions and reactions [5, 6]. Such an approach naturally aligns with the broad perspective of social behavior as communication; we interact to change and be changed by our interaction partners [7]. But how do we quantify the volleys of dynamics within this reciprocal perspective? Here we introduce an information-theoretic framework, which combines partial information decomposition (PID) [8] with prediction, to characterize the flows of information between participants in a social system. Our approach leverages the representation invariance of information theory and is thus less dependent both on human-derived measures and on particular movement trajectories, which can vary in presentation across organisms even in similar social contexts. As PID provides a fundamental decomposition of information flow, we also generalize previous approaches that used transfer entropy to determine dominance [9]. We compute and interpret

these information flows using both previous work on the pairwise dominance behavior of adult zebrafish [10], as well as new behavior measurements of *mecp2* zebrafish mutants, which have been linked to Autism Spectrum Disorder [11].

Decomposing predictive information in social dynamics

As a fundamental framework for understanding social dynamics, we focus on the concept of mutual prediction in the simplest social case of two individuals, Fig. 1(A). Our perspective is motivated by growing empirical evidence of social coupling, wherein the brain encodes information about an interacting partner [12, 13]. Accordingly, we aim to describe how the joint past state of the two individuals s_i^{past} and s_j^{past} predicts the future state of either s_i^{future}. This predictive relationship is naturally quantified using the joint mutual information,

$$I\left(\{s_i^{\text{past}}, s_j^{\text{past}}\}; s_i^{\text{future}}\right) = \left\langle \log \frac{P(s_i^{\text{future}}|s_i^{\text{past}}, s_j^{\text{past}})}{P(s_i^{\text{future}})} \right\rangle, \tag{1}$$

where $\langle \cdots \rangle$ denotes an average over the joint distribution, $P(s_i^{\text{past}}, s_j^{\text{past}}, s_i^{\text{future}})$. We note that this predictive information is a function of three variables and thus generally includes higher-order interactions [14–16]. The mutual information of Eq. 1 is a principled, scalar measure of the average amount of instantaneous predictive information. Yet, this overall information can result from a variety of processes. For example, how much of the future depends on the state of both individuals together? How much of the future of one individual is determined by itself or by the other? Previous research has examined directed influence using both Granger causality and it's nonlinear generalization, transfer entropy [17]. However, the framework of partial information decomposition (PID) [8, 18, 19], from which transfer entropy can also be computed, provides a more elemental picture. We argue here that PID of predictive information is a natural and quantitative architecture for capturing information flow resulting from social interactions, and thus for achieving deeper understanding of social behavior.

In PID we decompose the total or multi-information of a collection of variables into a sum of non-negative "information atoms" that capture redundant, synergistic and unique information present within the system. In an example with three variables $\{A, B, C\}$ the redundant information that $\{A, B\}$ carry about C is information this is contained in either A *or* B, while synergistic information is information that $\{A, B\}$ carry about C that requires both A *and* B. The unique information that A carries about C is information about C that is not contained in B and similarly for the unique information that B carries about C. Synergy and redundancy have been used to characterize neural activity [20] and large-scale human brain dynamics [21], and PID has also been used to understand gene expression [22].

We apply PID to two-organism social dynamics, with the future state of the focal organism s_i^{future} as the target and the past states of both organisms, s_i^{past} and s_j^{past} as predictors Fig. 1(B),

$$\begin{aligned} I\left(\{s_i^{\text{past}}, s_j^{\text{past}}\}; s_i^{\text{future}}\right) &= R_{i,j\to i} + U_{i\to i} + U_{j\to i} + S_{i,j\to i} \\ I\left(s_i^{\text{past}}; s_i^{\text{future}}\right) &= R_{i,j\to i} + U_{i\to i} \\ I\left(s_j^{\text{past}}; s_i^{\text{future}}\right) &= R_{i,j\to i} + U_{j\to i} \end{aligned} \tag{2}$$

where $R_{i,j\to i}$ denotes the redundant information about s_i^{future} shared by s_i^{past} and s_j^{past}, $U_{i\to i}$ and $U_{j\to i}$ are unique information from s_i^{past} and s_j^{past} respectively about s_i^{future}, and $S_{i,j\to i}$ represents the synergistic information about s_i^{future} available only in the joint state, $\{s_i^{\text{past}}, s_j^{\text{past}}\}$. The redundant information $R_{i,j\to i}$ is the degree to which predictive information is shared [23] and captures mirroring or synchronization between organisms. The self-unique information $U_{i\to i}$ is the information isolated within the focal organism, thus providing a measure of independence: organisms with larger self-unique information are more autonomous. In contrast, the other-unique information $U_{j\to i}$ is the information transferred from the other to the focal organism and thus characterizes directional influence. The synergistic information $S_{i,j\to i}$ is the predictive information modified by the opponent, capturing joint actions or coordination. An example of such coordination is the division of labor, as seen in leader–follower roles. Leaders determine the future group state, while followers adapt to the leader's actions, where the followers' predictive information is modified by the behavior of the leaders. While we introduce information flow in the context of two organisms, our approach can also be applied to larger groups, a point we discuss further in the discussion. Finally, we note that the transfer entropy is determined explicitly in terms of information atoms [24], $TE_{j\to i} = I\left(\{s_i^{\text{past}}, s_j^{\text{past}}\}; s_i^{\text{future}}\right) - I\left(s_i^{\text{past}}; s_i^{\text{future}}\right) = U_{j\to i} + S_{i,j\to i}$.

Joint coordinates and interaction changes across behavioral epochs

We explore our framework of predictive information flow in the context of the social behavior of zebrafish, an important vertebrate genetic model system. We focus on a pair social setting where two previously isolated adult male zebrafish are introduced into a novel environment, in which they then seek to establish a dominance hierarchy [25, 26]. We track their movements at high spatiotemporal resolution using a single 3D coordinate for each fish, the "pectoral point" located approximately at the base of the pectoral fins. Both the experiments and tracking have been previously described [10].

We quantify the movement trajectory of each fish using a suitably determined 3D velocity, Fig. 2(A). For each individual, we introduce a co-moving coordinate system which is spanned by the axis from the focal to the other organism $\hat{\boldsymbol{r}}^i$, the lab-frame z axis $\hat{\boldsymbol{z}}$, and a 3rd orthogonal axis $\hat{\boldsymbol{\theta}}^i = \hat{\boldsymbol{r}}^i \times \hat{\boldsymbol{z}}$. We compute the pectoral point velocity $\boldsymbol{V}^i$ using finite differences and then project onto the co-moving frame so that $v_r^i = \boldsymbol{V}^i \cdot \hat{\boldsymbol{r}}^i$, $v_z^i = \boldsymbol{V}^i \cdot \hat{\boldsymbol{z}}$, and $v_\theta^i = \boldsymbol{V}^i \cdot \hat{\boldsymbol{\theta}}^i$. We note that our coordinate system allows for ready interpretation. For example, the approach velocity v_r captures important agonistic interactions such as attack and chasing behavior, where v_r is predominantly positive for the chaser and negative for the fish being chased, Fig. 2(B). We focus our following analysis on v_r, with the corresponding results for v_θ and v_z provided in SI.

An important feature of social interactions is that they can change abruptly as information is transferred between participants; just imagine the reaction when walking into a sports bar and loudly rooting for the non-local team! In our two-fish setting of a dominance contest, these changes can be coarsely characterized as behavioral epochs that are quantitatively distinguished by the nature of the movements [10]: a pre-fight epoch with relatively weak interactions and symmetry between individuals, a fight epoch with strong interactions, and a post-fight epoch with a clear between individual symmetry breaking, Fig. 2(C, top). We use behavior in the post-fight epoch to designate the individual fish as "winner" or "loser". Among other attributes, the winner does most of the chasing.

We illustrate the interactions within each epoch though the cross-correlation function $C\left(v_r^{\mathrm{W}}(\tau), v_r^{\mathrm{L}}\right)$, Fig. 2(C) (Methods: Cross-correlation function). In the pre-fight epoch, the overall magnitude of the correlation, the linear coupling between individuals, is small. In the fight epoch, the larger coupling with an oscillation reflects cyclic attacks, where a fish being attacked will attack back after approximately 1 s. The symmetry of the interactions disappears after the fight with larger correlations for a negative time shift, indicating a larger influence from winners to losers. To systematically understand such changes of interaction, we next apply our predictive information decomposition to the dynamics of v_r.

Predictive information flows within behavioral epochs

We define a one-step future and a τ-step past for each organism, Fig. 3(A),

$$v_i^{\mathrm{past}}(t) = [v_i(t-\tau), v_i(t-\tau+\mathrm{d}t), ..., v_i(t)]$$
$$v_i^{\mathrm{future}}(t) = v_i(t+\mathrm{d}t),$$

and we illustrate our approach using the approach velocity v_r^i. We choose $\tau = 5\,\mathrm{s}$ to account for the significant cross-correlations, and we estimate PID using sliding time windows across the recording and a Gaussian approximation for the full distribution. Sliding time windows allow us to account for changing interactions Fig. 3(B-left) and we choose the window size to maximize the total information flow, Fig. 3(B-right). Note that this window size $\Delta t^* = 30\,\mathrm{s}$ is substantially longer than the past length τ. In the Gaussian approximation both the full predictive information of Eq. 1 and the partial information decomposition of Eq. 2 are analytically computable (Methods: Partial information decomposition for Gaussian variables). We provide movies of example social behaviors showing characteristic information flows in SI.

For each of the behavioral epochs (pre-fight, fight, post-fight) from each experiment we compute the predictive partial information decomposition of the approach velocities averaged across windows and normalized by the total information flow. We then average across experiments and show the results in Fig. 3(C). We find that self-unique information is high with the pre-fight epoch, implying that the fish are more independent, while redundant information increases during fighting, reflecting more synchronized movements. Self-unique information rises again in the post-fight epoch and includes a notable between-organism asymmetry: the self-unique information of the dominant fish is higher. Thus one aspect of the emergent hierarchy is that the movements of the dominant fish are more independent. Redundant predictive information is higher for losers, reflecting their tendency for more mirroring. Interestingly, this is in agreement with results from neural recordings of fighting mice, where subordinate's neurons predict opponent behavior more than dominant's neurons [13]. Synergy follows the same trend of redundancy, high information during

the fighting epoch and a post-fight asymmetry with higher information in losers than in winners, Fig. S2. This indicates that, fish coordinate their movement with an opponent more in fight, and in post-fight losers respond more to opponent movements compared to winners. Finally, we note that opponent-unique information is approximately zero across all epochs, Fig. S2, indicating that components of future movement uniquely determined by the opponent are negligible. This is essentially due to the larger auto-correlation compared to cross-correlation (Methods: Relationship between unique information and correlation functions). We note that a multidimensional extension of the target, beyond a one-step future, could reveal non-zero opponent-unique information, as further explored in Discussion.

Predictive information flows across behavioral epochs

To further understand how predictive information flows change during the span of our recordings, we align the window average to the start and stop of each contest, which was determined previously [10], and find that redundant, unique, and synergistic information flows display distinct dynamics across epoch boundaries, Fig. 4. Redundancy ramps up from the start of fight until the end, while self-unique information continuously decreases until the end of fight, suggesting that fish are increasingly engaged in interaction until the winner-loser decision. Interestingly, synergy shows a different trend than redundancy, remaining roughly constant for the duration of the fight. This constant synergy implies that a specific type of information exchange through their joint actions occurs at a constant frequency. In animal contest studies, the information exchange at constant frequency is considered a constraint from the mutual assessment process in which animals assess their opponent's ability relative to their own [27].

Predictive information flows are disrupted in social mutants

The *mecp2* mutant zebrafish is a model of Autism Spectrum Disorder [28] and displays clear behavioral abnormalities including motor anomalies and defective thigmotaxis [29]. We capture disruptions in mutants' social interaction by comparing predictive information flows between wildtype zebrafish pairs and *mecp2* mutant pairs from 30 s windows across a 5-hour recording. The total information flow $I\left(\{s_i^{\text{past}}, s_j^{\text{past}}\}; s_i^{\text{future}}\right)$ is lower in mutants than in wildtype, as shown in Fig. 5(A). This result suggests that the social behavior of mutants is less structured and less predictable compared to that of wild-type, potentially reflecting reduced motor activity [29], impaired sociability, or a combination of both factors. We then compute the average of the normalized partial informations, Fig. 5(B). Mutants show higher unique information and lower synergy, suggesting more independence and less ability to coordinate in social situations, which aligns with expectations for an autism model [28]. Interestingly, redundant flow is similar to wild type, suggesting that behavioral mirroring, which may require less cognitive demand than complementary actions, remains intact in this mutant.

Discussion

Organisms couple to their environment through an action-perception loop [30, 31], and social interactions are distinguished by the addition of feedback between these loops: the actions of one participant influence the perceptions and actions of another (see e.g. [6]). Here we account for this feedback using predictive information, the mutual information between the past and future of our social system. This predictive information can be expressed in redundant, unique and synergistic components using partial information decomposition. We find that these components are directly interpretable in the example of two zebrafish engaged in dominance interactions, in both wild-type and mutant examples. This interpretability is a key advantage of our information-theoretic framework, distinguishing it from neural network based behavioral clustering methods (e.g., [32]), which systematically classify behavior in an unsupervised manner but can yield clusters that are difficult to interpret.

The generality of our information-theoretic approach, which in principle is agnostic to behavioral coordinates, is ideal for comparative studies, where a key challenge lies in making meaningful cross-species comparisons and moving beyond case studies toward general principles underlying behavioral variation [33]. For example, comparing information flows arising from dominance interactions across species could provide quantitative and complementary insights into the interplay among contest strategies, cognition, and ecology, domains that have traditionally been analyzed through game theory [34–36].

We chose a one-step future $s^{\text{future}}(t) = x(t + \mathrm{d}t)$ to take advantage of a closed-form solution for PID that is only available for a univariate target. More generally, we can define the future state as the trajectory $s^{\text{future}}(t) = [x(t + \mathrm{d}t), x(t + 2\mathrm{d}t), \ldots, x(t + \tau)]$, enabling the characterization of higher-dimensional temporal dependencies. In Gaussian

variables for example, our partial information decomposition is constructed from all possible pairwise correlations and thus a longer future includes new correlations such as $\langle s_i^{\text{past}} x(t+2dt)\rangle$. Whether the use of higher-dimensional trajectory future can uncover non-zero opponent-unique information in zebrafish social dynamics is an important direction for future investigation. For a higher-dimensional Gaussian target, a numerical estimator of PID based on the BROJA unique information formulation has been proposed [37]. For generic distributions, a neural network approach based on the BROJA unique information [38] can accommodate high-dimensional targets. As an alternative, we can discretize the time series of $s^{\text{past}}(t)$ and $s^{\text{future}}(t)$ in any dimension and estimate PID using the original redundancy measure for discrete distributions [8].

Our Gaussian approximation for the joint distribution within each time window enabled the use of a closed-form solution for PID. For general continuous distributions, there are numerical PID estimators based on the Cone Program [39] and neural network optimization [38]. Both are based on BROJA unique information formulation. By discretizing the time-series, we can also estimate PID in discrete distributions based on the redundancy measure using the specific information [8, 22]. By relaxing the Gaussian approximation, we expect to extend the length of time windows and more accurately capture the complexity of social dynamics.

We can naturally extend our approach to more organisms by considering information flow from an N-organism past to a focal-organism future. For example, with a 3-organism past (3 predictors), redundant information about a target will take many forms such as the redundancy between a single predictor X and a joint of predictors $\{Y, Z\}$ and PID offers $m = 18$ modes of information flow. For arbitrary N, the decomposition can be determined by the Möbius transform, for which an efficient estimator has been developed [40]. We note however that the number of PID modes grows super-exponentially with N, determined by the Dedekind numbers minus two [18], and the interpretation of the resulting very large number PID modes for more than 3-organisms is a significant challenge. We note that it might be possible to coarse grain the information flows into macroscopic flows (see an attempt in ΦID [41]). Another possibility is to focus on the information about the focal individual's state that is carried by its own past and by the state of all other individuals, s_{other}, $I\left(\{s_i^{\text{past}}, s_{\text{other}}^{\text{past}}\}; s_i^{\text{future}}\right)$, or influence of a focal individual on the dynamics of other, $I\left(\{s_{\text{other}}^{\text{past}}, s_i^{\text{past}}\}; s_{\text{other}}^{\text{future}}\right)$. In such formulations, the number of PID modes grows linearly with the number of organisms, making PID analysis more tractable for larger groups, including those exhibiting collective behavior.

For simplicity we focused here on information flow from a two-organism past to a one-organism future. In dyadic settings, we can more generally investigate the information between a two-organism past and two-organism future, using mutual information $I\left(\{s_i^{\text{past}}, s_j^{\text{past}}\}; \{s_i^{\text{future}}, s_j^{\text{future}}\}\right)$. A natural decomposition of such information needs to describe how information encoded in a specific form (unique, redundant, and synergistic) evolves over time and be stored in any of each form in the future. This will introduce a picture of the "product" of two PIDs, one decomposing the information flow from past to future (predictors to targets) and another decomposing the flow from future to past (targets to predictors). Such an idea has been formalized as integrated information decomposition, ΦID [41]. By decomposing information between two predictors and two targets, ΦID provides a larger $N = 16$ set of modes of information corresponding to pairs of original PID modes evolving from past to future. For example, one of the ΦID modes represents information that is uniquely encoded in one of the variables in the past and as redundant information in the future, this can be interpreted as information duplication along the dynamics. Exploring ΦID for social dynamics is a promising future direction.

Zebrafish have emerged as a promising model organism for investigating social dysfunctions, complementary to traditional rodent models [28, 42]. Accordingly, substantial efforts are underway to advance behavioral assays and analysis methods for studying their social deficits using mutants [32, 42]. In this study, we identified behavioral anomalies in *mecp2* mutant zebrafish pairs, characterized by reduced predictive information, particularly in synergistic components, suggesting impairments in more complex social dynamics. We believe that our framework can be a useful tool for detecting subtle behavioral defects in other zebrafish models of neurological disorders, thereby contributing to the systematic investigation of these diseases. In addition, since neuropsychiatric disorders in human are often linked to social dysfunctions [43], our findings highlight the potential of using our framework for identifying biomarkers relevant to psychiatric diagnosis. Such biomarkers related to social dysfunctions may define a new transdiagnostic domain, offering targets for innovative treatments.

Acknowledgments

This work immediately follows previous research [10] enabled by grant RGP0055 from the Human Frontiers Science Program, and was supported by funds from OIST Graduate University (AK, RN, KD, LCR, IM, GJS) and Vrije Universiteit Amsterdam (LOS,GJS). We thank Tatsuo Izawa from the Biological Physics Theory Unit at OIST Graduate University for useful discussions and zebrafish behavior insight. We are also grateful for the help and support provided by the Engineering and Scientific Computing & Data Analysis sections of Research Support Division

at OIST.

Methods

Data: We use body point trajectories obtained from a recent advance in the measurement and analysis of zebrafish dominance contests [10], as well as from new experiments for the mutant and wildtype comparison of Fig. 5. A link to the data from Ref. [10] is available in the README file of the following Github project: https://github.com/liamshock/Dynamics of dominance. Briefly, the data consists of the 3D location of three body points (approximately situated near the head, pectoral fins and base of the tail), for each individual of a pair of adult male zebrafish. These body point locations are sampled at 100 Hz across a measurement period which covers pre-contest, contest and post-contest epochs. The dataset includes precise temporal boundaries for fight epochs, and we use $N = 18$ recordings in which a clear dominance structure emerges. For the mutant comparison, social interactions of $N = 8$ wild-type male pairs and $N = 6$ *mecp2* mutant male pairs were recorded for 5 hours at a sampling rate of 140 Hz. The fish were 13–17 months old and the recordings followed 24 to 72 hours of isolation. We analyzed the recordings as described above.

Zebrafish strains: Zebrafish were maintained at 28.5 °C on a 14-hour light/10-hour dark cycle, following standard procedures [44]. Okinawa wild-type (oki) was used as a wild-type strain. We used a *mecp2* mutant (sa21196, ZIRC ID ZL11092.02) allele with one point mutation that introduced a premature stop codon, provided by the Zebrafish International Resource Center (ZIRC, Eugene, OR, USA).

Ethics statement: All zebrafish experiments were carried out under the guidelines of the OIST Animal Care and Use Program, following the principles outlined in the Guide for the Care and Use of Laboratory Animals by the National Research Council of the Nation Academies. These procedures received approval from the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC). The OIST Institutional Animal Care and Use Committee approved all related experimental protocols (Protocol: ACUP-2023-018, ACUP-2023-022-2).

Behavioral epochs: The identification of the temporal boundaries of fight behavior was described previously [10]. Based on the boundaries of the fight epoch, we define the pre-fight epoch as the 10-minute period preceding the first fight in each recording, and the post-fight epoch as the 10-minute period following the final fight. If the first fight begins within 10 minutes of the start of the recording, the pre-fight epoch becomes accordingly shorter. Similarly, if the recording ends before 10 minutes have passed since the last fight, the post-fight epoch becomes shorter.

Pectoral point velocity: We restrict our attention to the pectoral point of each organism $\boldsymbol{x}^i_{\text{pec}}$ and construct the pectoral point velocity via the centered difference approximation:

$$\boldsymbol{V}^i(t) = \frac{\boldsymbol{x}^i_{\text{pec}}(t + \mathrm{d}t) - \boldsymbol{x}^i_{\text{pec}}(t - \mathrm{d}t)}{2\mathrm{d}t}$$

where $\mathrm{d}t = 10^{-2}$ s is the time step of the recording.

Projection of velocity onto co-moving coordinates: We introduce a pair coordinate system by defining an axis from the focal individual i to another j, $\hat{\boldsymbol{r}}^i = \frac{\boldsymbol{r}_{i\to j}}{|\boldsymbol{r}_{i\to j}|}$, where $\boldsymbol{r}_{i\to j} = \boldsymbol{x}^j_{\text{pec}} - \boldsymbol{x}^i_{\text{pec}}$, axis of z direction $\hat{\boldsymbol{z}}$, and another orthogonal axis $\hat{\boldsymbol{\theta}}^i = \hat{\boldsymbol{r}}^i \times \hat{\boldsymbol{z}}$. We then project the pectoral point velocity onto these axes:

$$v^i_r = \boldsymbol{V}^i \cdot \hat{\boldsymbol{r}}^i$$
$$v^i_z = \boldsymbol{V}^i \cdot \hat{\boldsymbol{z}}$$
$$v^i_\theta = \boldsymbol{V}^i \cdot \hat{\boldsymbol{\theta}}^i$$

Cross-correlation function: We compute the time-lagged cross-correlation function from the velocity components of each individual:

$$C\left(v^1(\tau), v^2\right) = \frac{\left\langle (v^1(t+\tau) - \langle v^1\rangle)(v^2(t) - \langle v^2\rangle)\right\rangle}{\sqrt{\langle (v^1(t) - \langle v^1\rangle)^2\rangle \langle (v^2(t) - \langle v^2\rangle)^2\rangle}}$$

For plots in Fig. 2(C), we compute the cross-correlation function for each epoch and then compute the function's average, solid line, and standard deviation, shaded area, across recordings.

Partial information decomposition for two predictor variables and one target variable: We consider the case where we predict a target variable Y from two predictor variables X_1 and X_2. Any nonempty subset of predictors can become a *source* of prediction, forming $\mathcal{S} = \{\{X_1\}, \{X_2\}, \{X_1, X_2\}\}$. The joint variable is considered another independent source because of the synergistic nature of information (e.g., the XOR operation). Any combination of sources may have non-negative redundant information about the target Y, denoted as $I_\cap(\boldsymbol{X}; Y)$, such that $\{X_1\}$ and $\{X_2\}$ have some common information about Y, $I_\cap^{\{X_1\}\{X_2\}}(\boldsymbol{X}; Y)$, and $\{X_1\}$ and $\{X_1\}$ have self-redundancy, which is equal to the mutual information: $I_\cap^{\{X_1\}}(\boldsymbol{X}; Y) = I(X_1; Y)$. However, if a source A is a subset of another source B, then $I_\cap^{AB}(\boldsymbol{X}; Y) = I_\cap^{A}(\boldsymbol{X}; Y)$, and thus redundant information in any combination of sources from $\mathcal{S}$ are simplified to four fundamental redundancies $I_\cap^{\{X_1\}\{X_2\}}(\boldsymbol{X}; Y)$, $I_\cap^{\{X_1\}}(\boldsymbol{X}; Y)$, $I_\cap^{\{X_2\}}(\boldsymbol{X}; Y)$, or $I_\cap^{\{X_1, X_2\}}(\boldsymbol{X}; Y)$. They are ordered such that $I_\cap^{\{X_1\}\{X_2\}}(\boldsymbol{X}; Y) \leq I_\cap^{\{X_1\}}(\boldsymbol{X}; Y) \leq I_\cap^{\{X_1, X_2\}}(\boldsymbol{X}; Y)$ and $I_\cap^{\{X_1\}\{X_2\}}(\boldsymbol{X}; Y) \leq I_\cap^{\{X_2\}}(\boldsymbol{X}; Y) \leq I_\cap^{\{X_1, X_2\}}(\boldsymbol{X}; Y)$. Therefore,

$$
\begin{aligned}
I_\cap^{\{X_1\}}(\boldsymbol{X}; Y) &= I_\cap^{\{X_1\}\{X_2\}}(\boldsymbol{X}; Y) + U_1 \\
I_\cap^{\{X_2\}}(\boldsymbol{X}; Y) &= I_\cap^{\{X_1\}\{X_2\}}(\boldsymbol{X}; Y) + U_2 \\
I_\cap^{\{X_1, X_2\}}(\boldsymbol{X}; Y) &= I_\cap^{\{X_1\}\{X_2\}}(\boldsymbol{X}; Y) + U_1 + U_2 + S
\end{aligned}
$$

where the unique information in individual variables U_1, U_2, and the synergistic contribution from the joint variable S are non-negative. By rewriting $I_\cap^{\{X_1\}\{X_2\}}(\boldsymbol{X}; Y)$ with R and self-redundancies with the mutual-information, we obtain the PID for two predictors:

$$
\begin{aligned}
I(X_1; Y) &= R + U_1 \\
I(X_2; Y) &= R + U_2 \\
I(\{X_1, X_2\}; Y) &= R + U_1 + U_2 + S
\end{aligned}
$$

These fundamental redundancies and their ordering can be systematically obtained as *antichains* of sources and the resulting redundancy lattice, respectively, for two or more predictors [8, 18, 19].

The three equations of PID contain four unknowns, and thus one more constrain is required to measure partial informations. In the original PID study from Williams and Beer [8], the redundancy measure is proposed by introducing the *specific information* of an outcome of target $Y = y$ given the random variable X:

$$
I(X; Y = y) = \int dx p(x|y) \left(\log \frac{1}{p(y)} - \log \frac{1}{p(y|x)} \right)
$$

Redundancy is the expected value over all y of the minimum specific information that two predictors X_1 and X_2 provide:

$$
R_{X_1, X_2 \to Y} = \int dy p(y) \min_{i \in \{1,2\}} I(X_i; Y = y)
$$

Alternatively, Bertschinger *et al.* [45] defined the unique information, called BROJA unique information, by minimizing the conditional mutual information over a set of hypothetical probability distributions q with the constraints that their marginals of each predictor X_i and target Y are fixed by the true marginals p:

$$
U_{X_1 \to Y} = \min_{q \in Q} I_q(X_1; Y | X_2)
$$

where $Q = \{q(x_1, x_2, y) | q(x_1, y) = p(x_1, y), q(x_2, y) = p(x_2, y)\}$. PID with BROJA unique information is equivalent to another definition from Griffith *et al.* [46].

Estimating dynamic information flows from data: To capture changing interactions, we introduce sliding time-windows, and we assume that the underlying joint distribution is Gaussian in each window, Fig. 3(B). The Gaussian predictive information is

$$
I\left(s^{\text{past}}; s^{\text{future}}\right) = \frac{1}{2} \log \frac{|\Sigma_{s^{\text{future}}}|}{\left|\Sigma_{s^{\text{future}}|s^{\text{past}}}\right|}, \tag{3}
$$

where $|\Sigma|$ is the determinant of a covariance matrix Σ, called the generalized variance, measuring the spread of samples [47]. The determinant of the partial covariance $\Sigma_{Y|X} = \Sigma_Y - \Sigma_{X,Y} \Sigma_X^{-1} \Sigma_{X,Y}^{\mathrm{T}}$ can be simply computed by

$\left|\Sigma_{Y|X}\right| = \frac{|\Sigma_{X,Y}|}{|\Sigma_X|}$ [17]. We estimate the covariance matrix with an adaptive regularization using the OAS estimator [48]. By maximizing the joint information flow $I\left(\{s_i^{\text{past}}, s_j^{\text{past}}\}; s_i^{\text{future}}\right)$, we determine $\Delta t^* = 30$ s as the optimal time-window size for the set of our recordings, Fig. 3(B). For a short time-window, the number of samples is small, and the covariance is highly regularized, resulting small information flow. On the other hand, within long time-windows, distinct linear dynamics appear, correlations decrease, and information flow becomes small.

Partial information decomposition for Gaussian variables: We use a closed-form solution of redundant information in two Gaussian predictors and a scaler Gaussian target, called "minimum mutual information" [49]:

$$\begin{aligned} R_{X_1,X_2\to Y} &= \min_{i\in\{1,2\}} I(X_i;Y) \\ &= \frac{1}{2}\min\left(\log\frac{|\Sigma_Y|}{|\Sigma_{Y|X_1}|}, \log\frac{|\Sigma_Y|}{|\Sigma_{Y|X_2}|}\right) \end{aligned}$$

This provides the solution to both PIDs from Williams and Beer [8] and from the BROJA unique information [45] for this Gaussian case. Using the definition of PID provided in Eq. 2 we obtain the other partial informations:

$$\begin{aligned} U_{X_1\to Y} &= \frac{1}{2}\max\left(\log\frac{|\Sigma_{Y|X_2}|}{|\Sigma_{Y|X_1}|}, 0\right) \\ U_{X_2\to Y} &= \frac{1}{2}\max\left(\log\frac{|\Sigma_{Y|X_1}|}{|\Sigma_{Y|X_2}|}, 0\right) \\ S_{X_1,X_2\to Y} &= \frac{1}{2}\min\left(\log\frac{|\Sigma_{Y|X_1}|}{|\Sigma_{Y|X_1,X_2}|}, \log\frac{|\Sigma_{Y|X_2}|}{|\Sigma_{Y|X_1,X_2}|}\right) \end{aligned}$$

Relationship between unique-information and correlation fuctions: For PID of predictive information, the opponent-unique information can be expressed as

$$U_{s_j\to s_i} = \frac{1}{2}\max\left(\log\frac{\left|\Sigma_{s_i^{\text{future}}|s_i^{\text{past}}}\right|}{\left|\Sigma_{s_i^{\text{future}}|s_j^{\text{past}}}\right|}, 0\right)$$

Here, the partial covariance $\left|\Sigma_{s_i^{\text{future}}|s_i^{\text{past}}}\right|$ quantifies the conditional variance of s_i^{future} given s_i^{past}, and becomes smaller when s_i^{future} is more predictable from its own past i.e. when autocorrelation is higher. Similarly, $\left|\Sigma_{s_i^{\text{future}}|s_j^{\text{past}}}\right|$ becomes smaller when there is high cross-correlation with s_j^{past}. A value of $U_{s_j\to s_i} = 0$ implies

$$\left|\Sigma_{s_i^{\text{future}}|s_i^{\text{past}}}\right| < \left|\Sigma_{s_i^{\text{future}}|s_j^{\text{past}}}\right|$$

indicating that s_i^{future} is better predicted from its own past than from the past of s_j, and auto-correlation is higher than cross-correlation. The condition of zero opponent-unique information, observed in our analysis of zebrafish contests, may reflect a generic feature of behavior in which self-predictability exceeds prediction from its environment. We note that this formulation of unique information is valid only under the assumption of a scalar Gaussian target. Alternative formulations, such as incorporating multi-step futures, may reveal non-zero opponent-unique information, as discussed further in the Discussion section.

Choosing past and future variables: An important step in constructing predictive information is the definition of the past and future variables. Most generally, we can choose the past state as a long history $s^{\text{past}}(t) = [x(t-\tau), x(t-\tau+\mathrm{d}t), ..., x(t)]$ and the future state as a long future trajectory $s^{\text{future}}(t) = [x(t+\mathrm{d}t), x(t+2\mathrm{d}t), ..., x(t+\tau)]$, where τ is the decay time (i.e., $I(x(t); x(t+\tau)) \sim 0$). In this study, we focus on the one-step future $s^{\text{future}}(t) = [x(t+\mathrm{d}t)]$ as the closed-form solution of PID for continuous variables only available for a scaler variable [49]. This solution is robust to the dimension of predictor variables, and thus, we adopt $s^{\text{past}}(t) = [x(t-\tau^*), x(t-\tau^*+\mathrm{d}t), ..., x(t)]$, where the cross-correlation function decays at τ^* (i.e., $C(\tau) \sim 0$ for $|\tau| > \tau^*$). We use $\tau^* = 5$ s as the length of the past state, thus incorporating the significant cross-correlations, Fig 2(C).

Covariance estimation: Our target covariance matrix Σ is high-dimensional $p \times p$, but the number of samples n is limited. In such a case, the sample covariance $S = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^T$ does not minimize the mean squared error

$E\{\|S-\Sigma\|^2\}$ [48], and the estimation requires a correct regularization. We adopt the oracle approximating shrinkage (OAS) estimator [48]:

$$\Sigma_{\text{OAS}} = (1-\rho)S + \rho F$$

where $F = \frac{\text{Tr}(S)}{p} I$ and

$$\rho = \min\left(\frac{(1-2/p)\text{Tr}(S^2) + \text{Tr}^2(S)}{(n+1-2/p)[\text{Tr}(S^2) - \text{Tr}^2(S)/p]}, 1\right)$$

Computing the average of normalized partial information: We focus on changes in the relative amount of partial informations from each time-window, and thus normalize each partial information by the total information flow, $I\left(\{s_i^{\text{past}}, s_j^{\text{past}}\}; s_i^{\text{future}}\right)$. For plots in Fig. 3(C), we first compute the mean of the normalized partial informations within epochs for each recording and obtain the average and SEM across recordings. For plots in Fig. 4, we first align the time by the start and end of the fight epoch and compute the average and SEM of normalized partial informations across recordings at every time point.

Bibliography

[1] Karl J Niklas. The evolutionary-developmental origins of multicellularity. *American journal of botany*, 101(1):6–25, 2014.

[2] Andrew Haddon Kemp, Juan Antonio Arias, and Zoe Fisher. Social ties, health and wellbeing: a literature review and model. *Neuroscience and social science: The missing link*, pages 397–427, 2017.

[3] Christian Keysers and Valeria Gazzola. Neurobiology of social interactions across species, 2025.

[4] Amir Haluts, Alex Jordan, and Nir S Gov. Modelling animal contests based on spatio-temporal dynamics. *Journal of the Royal Society Interface*, 20(202):20220866, 2023.

[5] Uri Hasson and Chris D. Frith. Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1693):20150366, May 2016.

[6] Conor Heins, Beren Millidge, Lancelot Da Costa, Richard P Mann, Karl J Friston, and Iain D Couzin. Collective behavior from surprise minimization. *Proceedings of the National Academy of Sciences*, 121(17):e2320239121, 2024.

[7] Yuvalal Liron, Noa Raindel, and Uri Alon. Dramatic action: A theater-based paradigm for analyzing human interactions. *Plos one*, 13(3):e0193404, 2018.

[8] Paul L. Williams and Randall D. Beer. Nonnegative Decomposition of Multivariate Information, April 2010. arXiv:1004.2515 [math-ph, physics:physics, q-bio].

[9] Rishita Das and Maurizio Porfiri. A controlled transfer entropy approach to detect asymmetric interactions in heterogeneous systems. *Journal of Physics: Complexity*, 4(2):025020, 2023.

[10] Liam O'Shaughnessy, Tatsuo Izawa, Ichiro Masai, Joshua W. Shaevitz, and Greg J. Stephens. Dynamics of Dominance in Interacting Zebrafish. *PRX Life*, 2(4):043006, October 2024.

[11] Akram Tayanloo-Beik, Shayesteh Kokabi Hamidpour, Mina Abedi, Hamide Shojaei, Mostafa Rezaei Tavirani, Nazli Namazi, Bagher Larijani, and Babak Arjmand. Zebrafish modeling of autism spectrum disorders, current status and future prospective. *Frontiers in Psychiatry*, 13:911770, 2022.

[12] Greg J. Stephens, Lauren J. Silbert, and Uri Hasson. Speaker–listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences*, 107(32):14425–14430, August 2010.

[13] Lyle Kingsbury, Shan Huang, Jun Wang, Ken Gu, Peyman Golshani, Ye Emily Wu, and Weizhe Hong. Correlated Neural Activity and Encoding of Behavior across Brains of Socially Interacting Animals. *Cell*, 178(2):429–446.e16, July 2019.

[14] Satosi Watanabe. Information theoretical analysis of multivariate correlation. *IBM Journal of research and development*, 4(1):66–82, 1960.

[15] S-I Amari. Information geometry on hierarchy of probability distributions. *IEEE transactions on information theory*, 47(5):1701–1711, 2001.

[16] Elad Schneidman, Susanne Still, Michael J. Berry, and William Bialek. Network Information and Connected Correlations. *Physical Review Letters*, 91(23):238701, December 2003. Publisher: American Physical Society.

[17] Lionel Barnett, Adam B. Barrett, and Anil K. Seth. Granger Causality and Transfer Entropy Are Equivalent for Gaussian Variables. *Physical Review Letters*, 103(23):238701, December 2009.

[18] A. J. Gutknecht, M. Wibral, and A. Makkeh. Bits and pieces: understanding information decomposition from part-whole relationships and formal logic. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 477(2251):20210110, July 2021. Publisher: Royal Society.

[19] Abel Jansma. Mereological approach to higher-order structure in complex systems: From macro to micro with Möbius. *Physical Review Research*, 7(2):023016, April 2025.

[20] Elad Schneidman, William Bialek, and Michael J. Berry. Synergy, redundancy, and independence in population codes. *Journal of Neuroscience*, 23(37):11539–11553, 2003.

[21] Andrea I Luppi, Pedro AM Mediano, Fernando E Rosas, Judith Allanson, John Pickard, Robin L Carhart-Harris, Guy B Williams, Michael M Craig, Paola Finoia, Adrian M Owen, Lorina Naci, David K Menon, Daniel Bor, and Emmanuel A Stamatakis. A synergistic workspace for human consciousness revealed by integrated information decomposition. *eLife*, 12:RP88173, jul 2024.

[22] Thalia E. Chan, Michael P.H. Stumpf, and Ann C. Babtie. Gene Regulatory Network Inference from Single-Cell Data Using Multivariate Information Measures. *Cell Systems*, 5(3):251–267.e3, September 2017.

[23] Ryan G James, Blanca Daniella Mansante Ayala, Bahti Zakirov, and James P Crutchfield. Modes of information flow. *arXiv preprint arXiv:1808.06723*, 2018.

[24] Ryan G. James, Nix Barnett, and James P. Crutchfield. Information Flows? A Critique of Transfer Entropies. *Physical Review Letters*, 116(23):238701, June 2016.

[25] Rui F. Oliveira, Joana F. Silva, and José M. Simões. Fighting Zebrafish: Characterization of Aggressive Behavior and Winner–Loser Effects. *Zebrafish*, 8(2):73–81, June 2011.

[26] Andres Laan, Marta Iglesias-Julios, and Gonzalo G. De Polavieja. Zebrafish aggression on the sub-second time scale: evidence for mutual motor coordination and multi-functional attack manoeuvres. *Royal Society Open Science*, 5(8):180679, August 2018.

[27] Robert J.H. Payne. Gradually escalating fights and displays: the cumulative assessment model. *Animal Behaviour*, 56(3):651–662, September 1998.

[28] Victoria Rea and Terence J. Van Raay. Using Zebrafish to Model Autism Spectrum Disorder: A Comparison of ASD Risk Genes Between Zebrafish and Their Mammalian Counterparts. *Frontiers in Molecular Neuroscience*, 13:575575, November 2020.

[29] Thomas Pietri, Angel-Carlos Roman, Nicolas Guyon, Sebastián A. Romano, Philip Washbourne, Cecilia B. Moens, Gonzalo G. De Polavieja, and Germán Sumbre. The first mecp2-null zebrafish model shows altered motor behaviors. *Frontiers in Neural Circuits*, 7, 2013.

[30] William T. Powers. Feedback: Beyond Behaviorism. *Science*, 179(4071):351–356, January 1973. Publisher: American Association for the Advancement of Science.

[31] Alex Gomez-Marin and Asif A. Ghazanfar. The Life of Behavior. *Neuron*, 104(1):25–36, October 2019.

[32] Yijie Geng, Christopher Yates, and Randall T. Peterson. Social behavioral profiling by unsupervised deep learning reveals a stimulative effect of dopamine D3 agonists on zebrafish sociality. *Cell Reports Methods*, 3(1):100381, January 2023.

[33] Nicholas Jourjine and Hopi E. Hoekstra. Expanding evolutionary neuroscience: insights from comparing variation in behavior. *Neuron*, 109(7):1084–1099, April 2021.

[34] J Maynard Smith and George R Price. The logic of animal conflict. *Nature*, 246(5427):15–18, 1973.

[35] Gareth Arnott and Robert W. Elwood. Assessment of fighting ability in animal contests. *Animal Behaviour*, 77(5):991–1004, May 2009.

[36] Michael S. Reichert and John L. Quinn. Cognition in Contests: Mechanisms, Ecology, and Evolution. *Trends in Ecology & Evolution*, 32(10):773–785, October 2017.

[37] Praveen Venkatesh, Corbett Bennett, Sam Gale, Tamina K. Ramirez, Greggory Heller, Severine Durand, Shawn Olsen, and Stefan Mihalas. Gaussian Partial Information Decomposition: Bias Correction and Application to High-dimensional Data, July 2023. arXiv:2307.10515 [cs].

[38] Ari Pakman, Amin Nejatbakhsh, Dar Gilboa, Abdullah Makkeh, Luca Mazzucato, Michael Wibral, and Elad Schneidman. Estimating the Unique Information of Continuous Variables, October 2021. arXiv:2102.00218 [cs, math, stat].

[39] Abdullah Makkeh, Dirk Oliver Theis, and Raul Vicente. BROJA-2PID: A Robust Estimator for Bivariate Partial Information Decomposition. *Entropy*, 20(4):271, April 2018. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[40] Abel Jansma, Pedro A. M. Mediano, and Fernando E. Rosas. The Fast Möbius Transform: An algebraic approach to information decomposition, October 2024. arXiv:2410.06224.

[41] Pedro A. M. Mediano, Fernando E. Rosas, Andrea I. Luppi, Robin L. Carhart-Harris, Daniel Bor, Anil K. Seth, and Adam B. Barrett. Towards an extended taxonomy of information dynamics via Integrated Information Decomposition, September 2021. arXiv:2109.13186 [physics, q-bio].

[42] Yijie Geng and Randall T. Peterson. The zebrafish subcortical social brain as a model for studying social behavior disorders. *Disease Models & Mechanisms*, 12(8):dmm039446, August 2019.

[43] Stefano Porcelli, Nic Van Der Wee, Steven van der Werff, Moji Aghajani, Jeffrey C. Glennon, Sabrina van Heukelum, Floriana Mogavero, Antonio Lobo, Francisco Javier Olivera, Elena Lobo, Mar Posadas, Juergen Dukart, Rouba Kozak, Estibaliz Arce, Arfan Ikram, Jacob Vorstman, Amy Bilderbeck, Ilja Saris, Martien J. Kas, and Alessandro Serretti. Social brain, social dysfunction and social withdrawal. *Neuroscience & Biobehavioral Reviews*, 97:10–33, February 2019.

[44] Monte Westerfield. *The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio)*. Univ. of Oregon Press, Eugene, 4th edition, 2000.

[45] Nils Bertschinger, Johannes Rauh, Eckehard Olbrich, Jürgen Jost, and Nihat Ay. Quantifying Unique Information. *Entropy*, 16(4):2161–2183, April 2014. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[46] Virgil Griffith, Edwin K. P. Chong, Ryan G. James, Christopher J. Ellison, and James P. Crutchfield. Intersection Information Based on Common Randomness. *Entropy*, 16(4):1985–2000, April 2014. Number: 4 Publisher: Multidisciplinary

Digital Publishing Institute.

[47] T. W. Anderson. *An introduction to multivariate statistical analysis*. Wiley series in probability and statistics. Wiley-Interscience, Hoboken, N.J, 3rd ed edition, 2003.

[48] Yilun Chen, Ami Wiesel, Yonina C. Eldar, and Alfred O. Hero. Shrinkage Algorithms for MMSE Covariance Estimation. *IEEE Transactions on Signal Processing*, 58(10):5016–5029, October 2010.

[49] Adam B. Barrett. Exploration of synergistic and redundant information sharing in static and dynamical Gaussian systems. *Physical Review E*, 91(5):052802, May 2015. Publisher: American Physical Society.

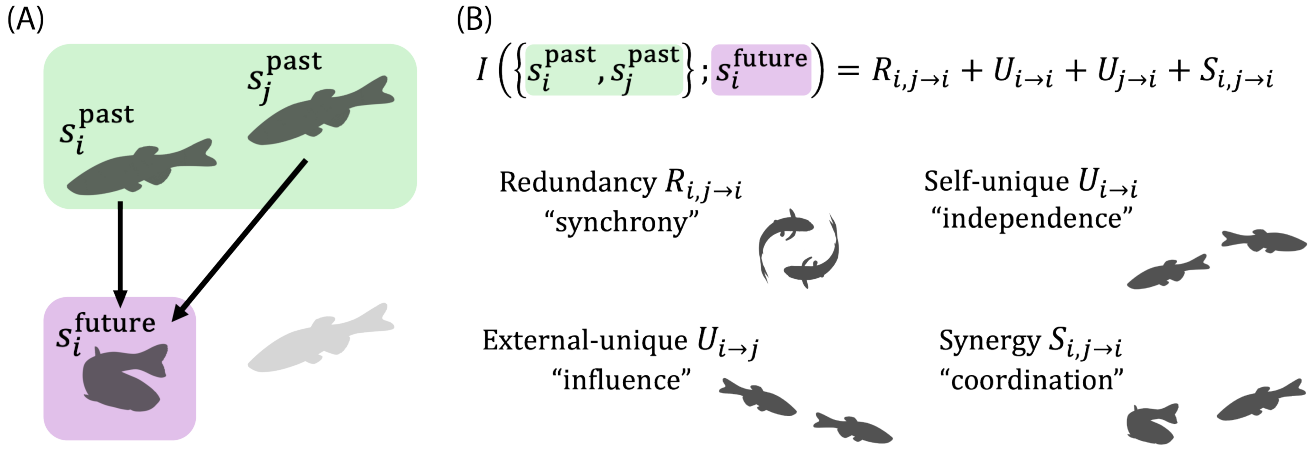

FIG. 1: **Decomposing predictive information in social dynamics.** (A) In social behavior, the future state of a focal individual s_i^{future} depends on both the past state of itself s_i^{past} and the past state of others s_j^{past}. Even the minimal social system of two agents thus involves three variables that support multiple modes of interactions, such as mirroring and complementary actions between individuals. (B) We introduce partial information decomposition (PID) of predictive information for social dynamics. In PID we decompose the total predictive information $I(\{s_i^{\text{past}}, s_j^{\text{past}}\}; s_i^{\text{future}})$ into four information terms, where redundant information $R_{i,j\to i}$, self-unique information $U_{i\to i}$, external-unique information $U_{j\to i}$, and synergistic information $S_{i,j\to i}$ quantifies the degree of synchrony, independence, directional influence, and joint coordination, respectively.

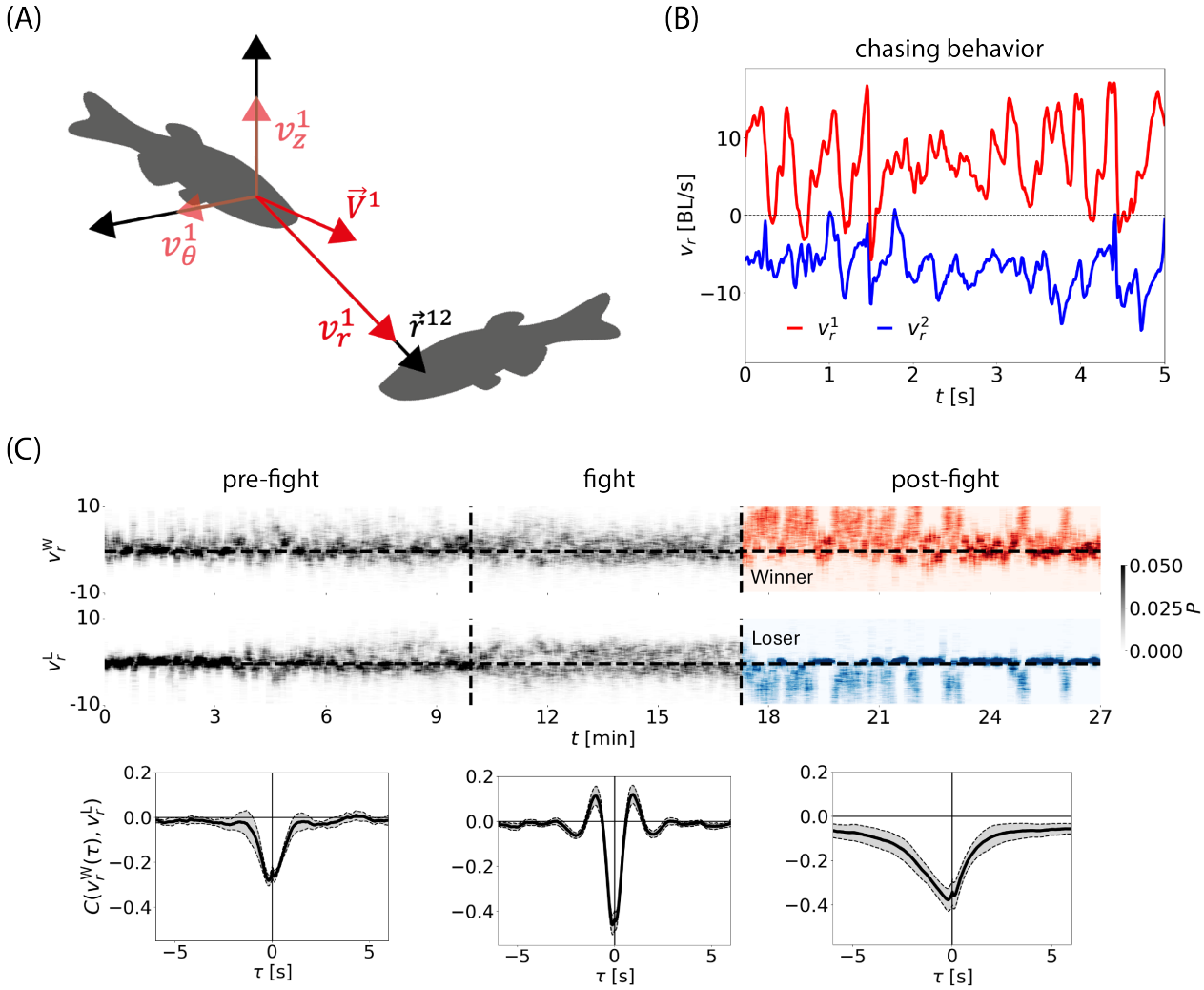

FIG. 2: **We introduce co-moving coordinates and show that interactions change dramatically across behavioral epochs.** (A) We capture interactions through movements using projections of the velocity vector $\boldsymbol{V}^1$ onto the axis from the focal individual to another $\boldsymbol{r}^{12}$ (v_r^1), onto the z-axis (v_z^1), and the axis orthogonal to the other two axes (v_θ^1). (B) v_r reflects approaching movements, such as chasing behavior, and we show an example time series where fish 1 chases fish 2. (C) Social dynamics of two male zebrafish show sequence structure, i.e., epochs, consisting of "pre-fight", "fight", and "post-fight". The approach velocity v_r is peaked around $v_r = 0$ in pre-fight epochs while it is more varied during the fight. In post-fight epochs, the approach velocity of the winner v_r^{W} tends to be positive while the approach velocity of the loser v_r^{L} is generally negative, reflecting the predominance of winners chasing losers. We compute the average cross-correlation function of v_r between the (eventual) winner and loser fish $C\left(v_r^{\mathrm{W}}(\tau), v_r^{\mathrm{L}}\right)$ (Methods: Cross-correlation function) in each epoch across experiments. In pre-fight epochs, the magnitude of the correlation, i.e., coupling, is small. In fight epochs the coupling is larger and displays oscillations reflecting cyclic attacks. In post-fight epochs, the time-symmetry is broken, demonstrating larger influences from winner to loser. We use predictive information decomposition to systematically describe such changes in the interactions.

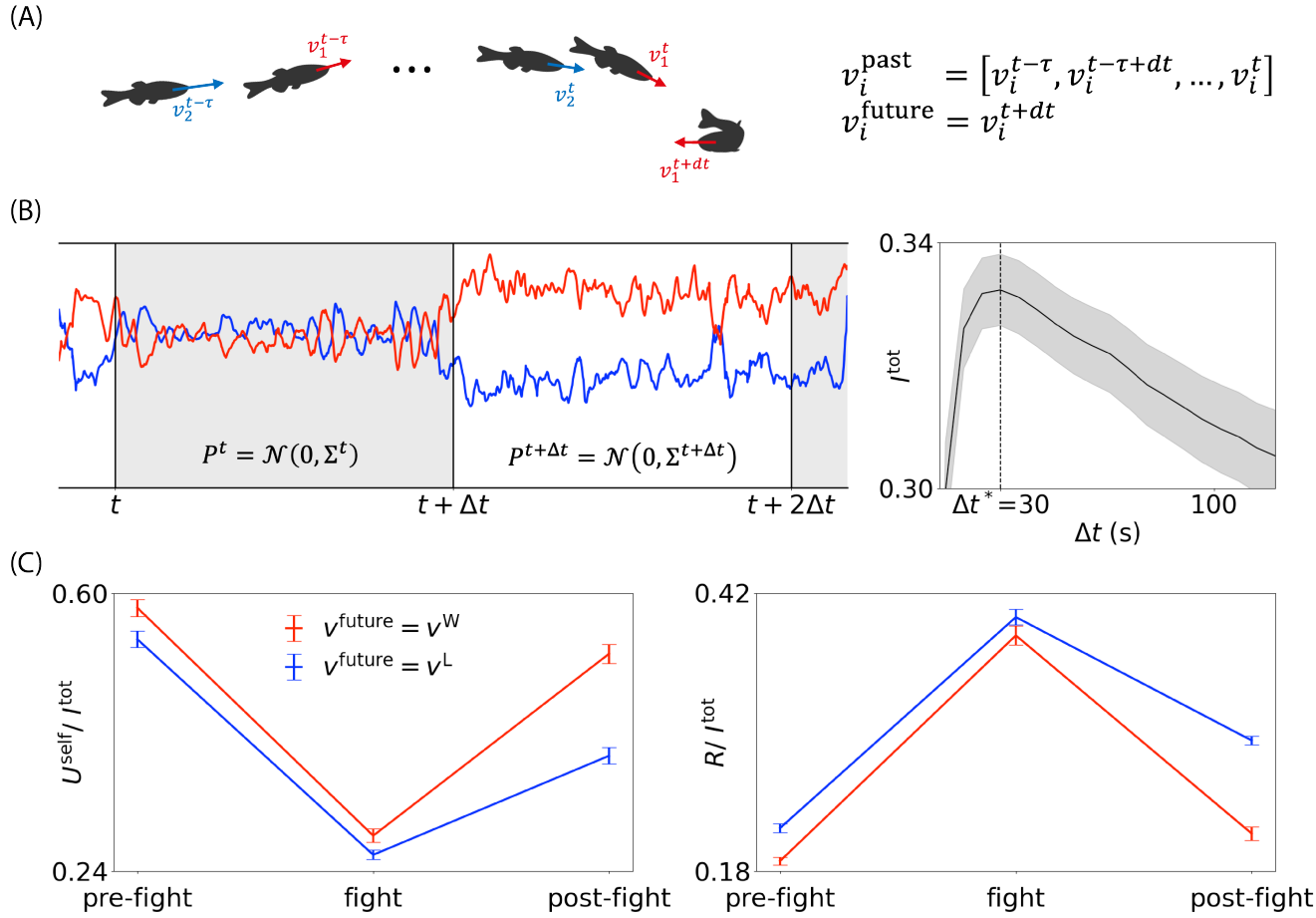

FIG. 3: **Predictive information decomposition within epochs.** (A) We define a one-step future and a τ-step past for each organism, and set $\tau = 5\,\text{s}$ to include significant temporal correlations, see Fig. 2(C). (B) We estimate predictive information decomposition using sliding time windows across the recording, which captures changing interactions, and a Gaussian approximation for the full distribution, for which an analytic PID expression is available. We choose the window size $\Delta t^* = 30\,\text{s}$ to maximize the total information flow $I^{\text{tot}} = I\left(\{s_i^{\text{past}}, s_j^{\text{past}}\}; s_i^{\text{future}}\right)$ in the Gaussian limit, Eq. 3. (C) We normalize the partial informations by I^{tot} in each window and compute the mean information within each epoch. Plots show the average across experiments of the mean information, and error bars show SEM. Self-unique information is high and symmetric in the pre-fight epoch, reflecting more independent movements. Redundancy increases during the fight, suggesting more coupled dynamics. Self-unique information rises again in the post-fight epoch, with an asymmetry reflecting the emergent dominance relationship, demonstrating that the winning organism is more independent.

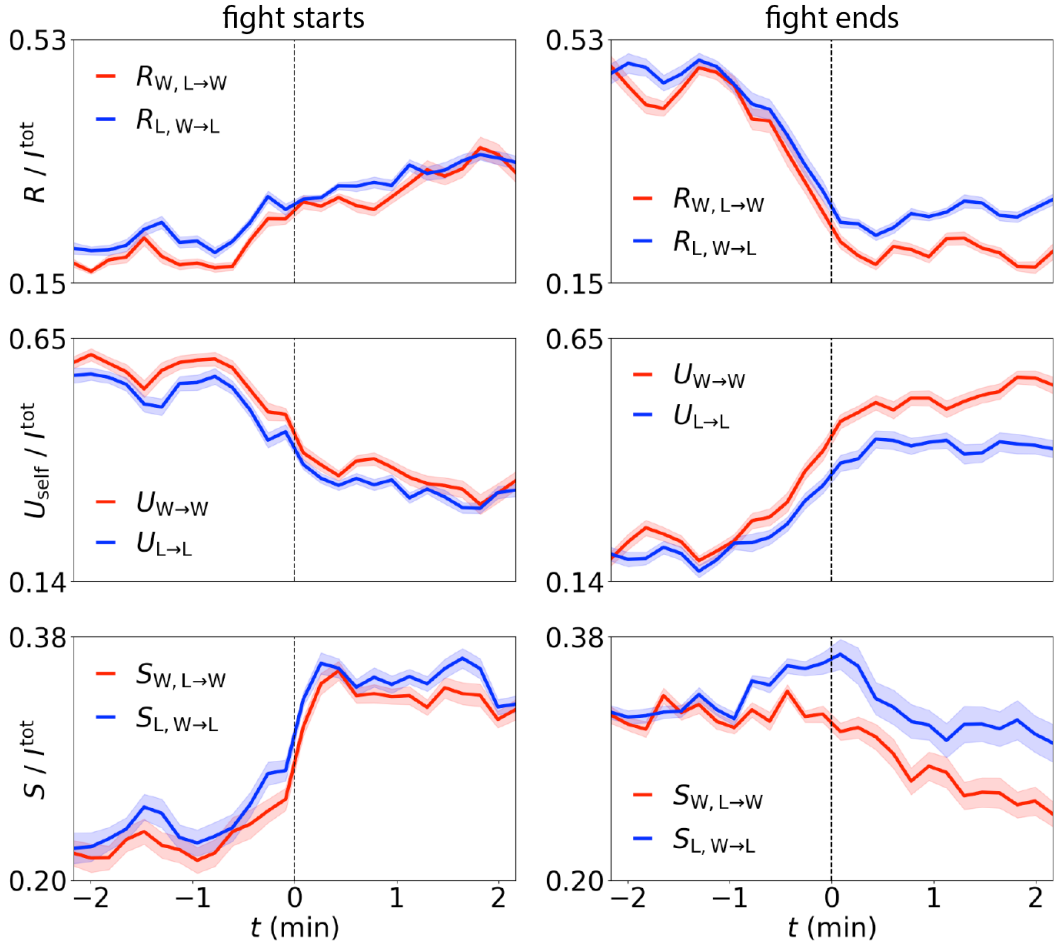

FIG. 4: **Redundant, unique, and synergistic information flows display different dynamics across contest boundaries.** (Left) We align by the start of the fight epoch and compute the average and SEM of the fraction of partial information across experiments. (Right) We align by the end of the fight epoch. Redundancy increases from the start of the fight until the end, while self-unique information continuously decreases until the end of the fight. Synergy shows a different trend than redundancy, remaining roughly constant for the duration of the fight. There are clear post-fight asymmetries across all flows.

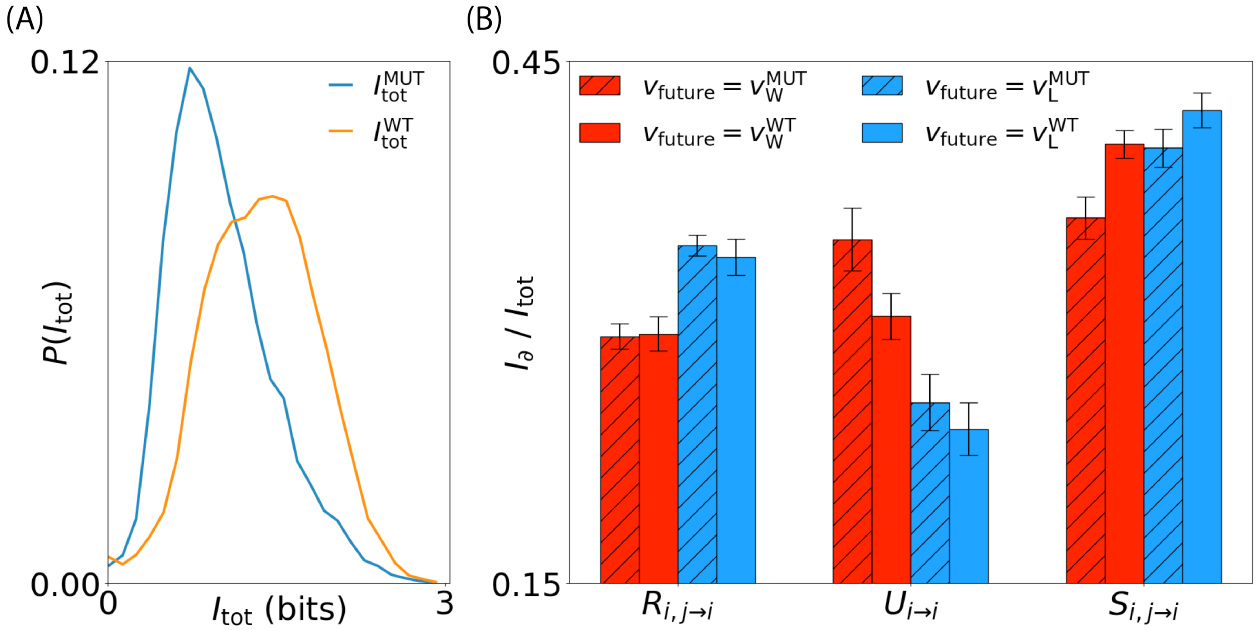

FIG. 5: **Predictive information flows are disrupted in *mecp2* mutants, a zebrafish model of Autism Spectrum Disorder.** (A) The total predictive information flow is lower in the mutants (blue curve) than in wildtype (red curve). We construct these distributions from 30 s windows across a 5-hour recording. (B) In both mutants and wildtype, the self-unique information flow towards the winner is larger than the flow towards the loser, demonstrating that this flow captures the dominance hierarchy in both cases. In contrast, mutant fish show higher unique information and lower synergy, suggesting more independence and less ability to coordinate in social situations. Interestingly, redundant flow is similar to wildtype, suggesting that behavioral mirroring remains intact in this mutant.

Supplementary Material

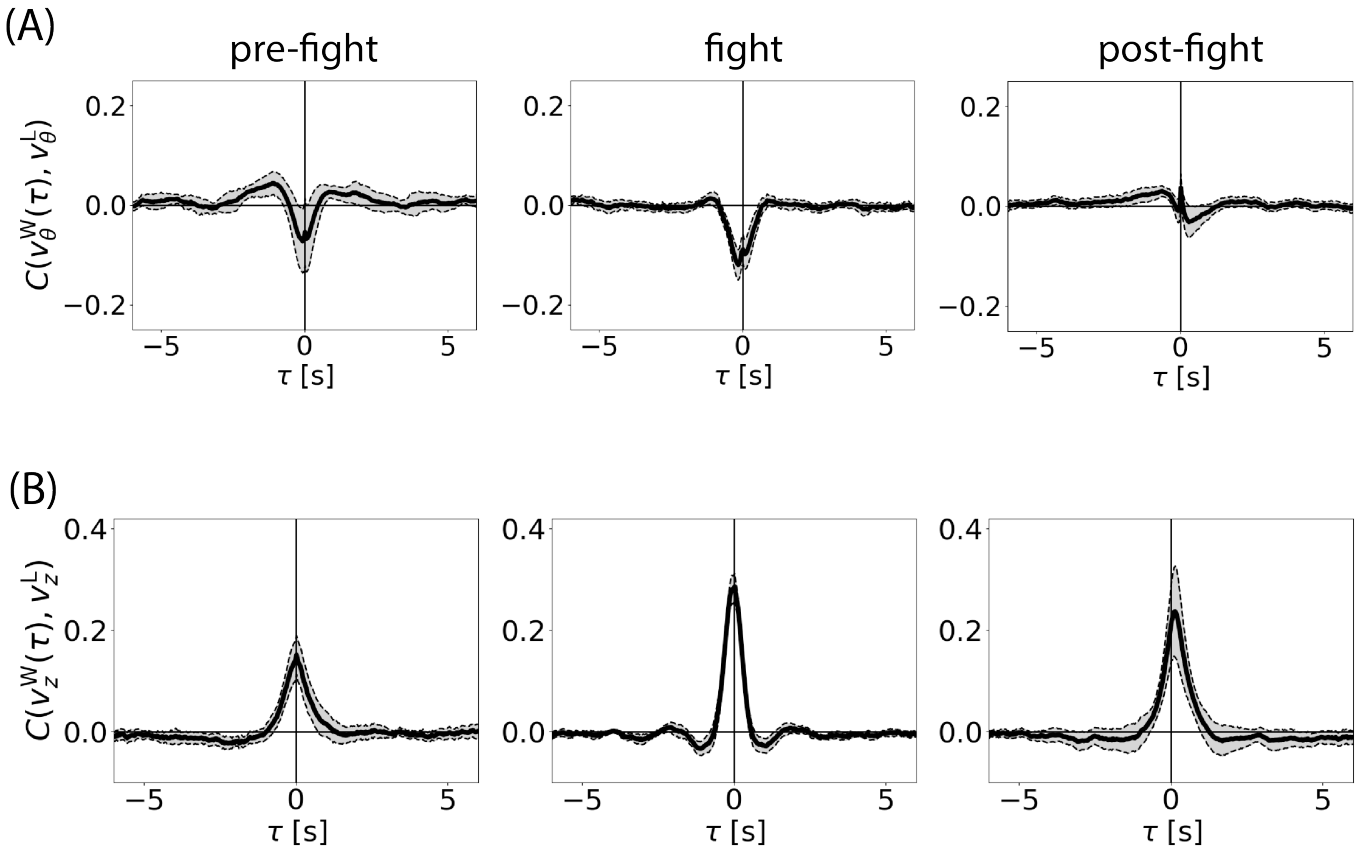

FIG. S1: **Cross-correlations in v_θ and v_z across epochs.** (A) We compute the cross-correlation of v_θ between individuals across behavioral epochs, and show the average across 18 experiments. Correlations are higher during the fight phase, reflecting simultaneous directional changes during chases. (B) The average cross-correlation of v_z increases from the pre-fight to fight epoch, indicating that behaviors increasingly involve movement in the z direction. The peak of the cross-correlation in the post-fight epoch is slightly shifted toward a positive time lag, suggesting that although the post-fight chase is dominated by winners, movement along the z-axis during the chase is led by losers.

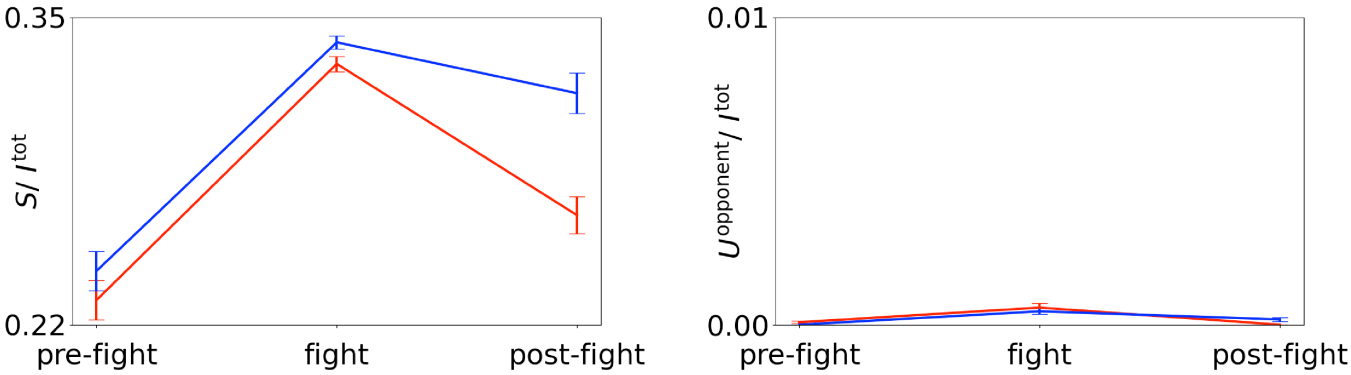

FIG. S2: **Normalized synergistic information and opponent-self information in v_r across epochs.** Synergy follows the same trend of redundancy: high information in the fight epoch and a post-fight asymmetry with higher information in losers than in winners. Opponent-unique information is approximately zero for all epochs as fish movements have inertia, and in their next move in 100 Hz, behavioral elements that uniquely determined by their opponent are negligible.

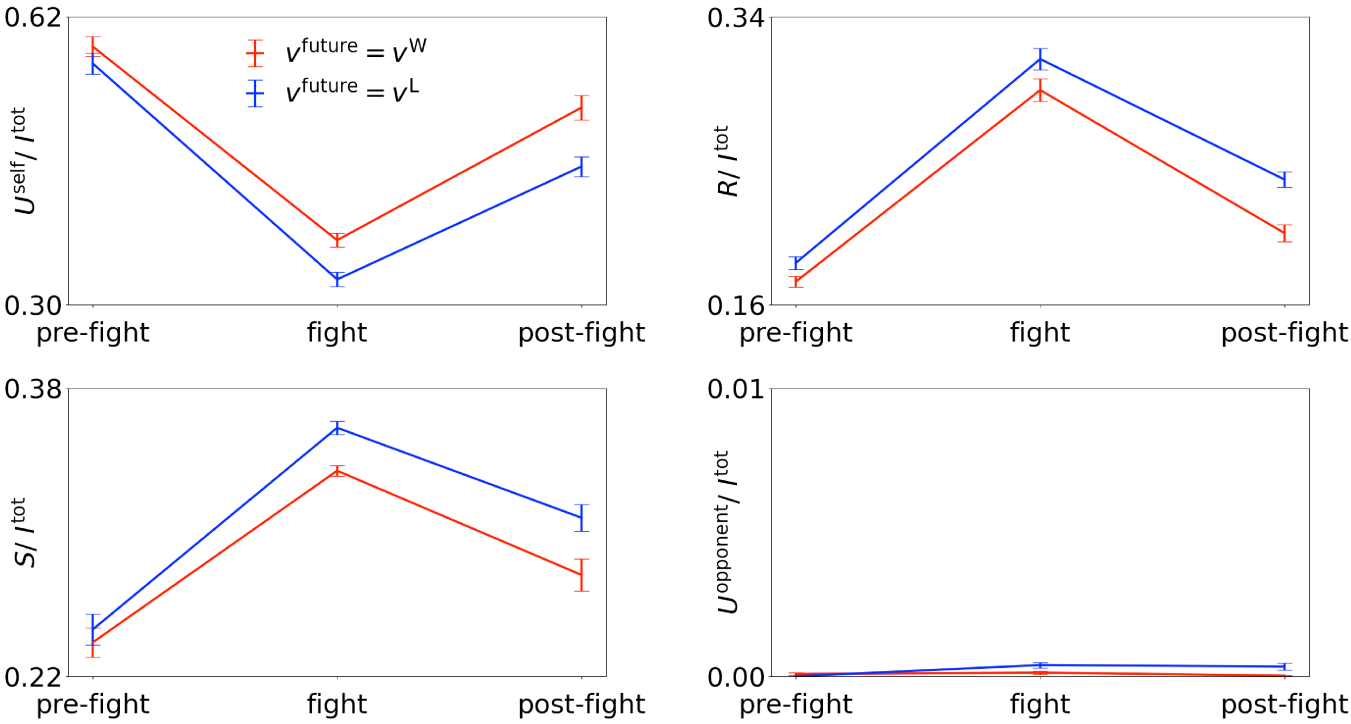

FIG. S3: **Normalized partial informations in v_θ across epochs.** We compute information flows as with v_r and show the average and SEM across experiments.

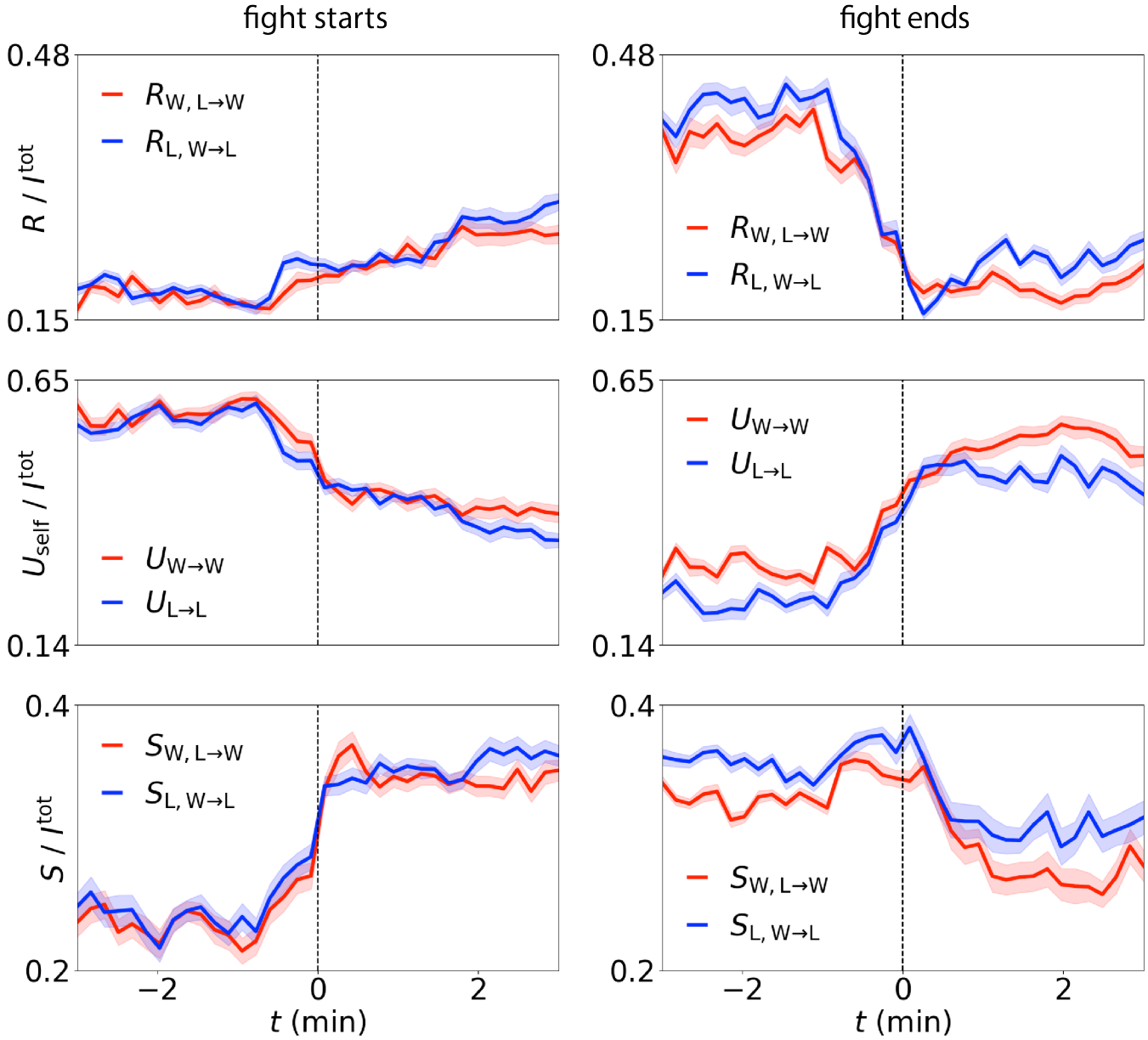

FIG. S4: **Dynamics of partial informations in v_θ across contest boundaries.** (Left) We align by the start of the fight epoch and compute the average and SEM of the fraction of partial information in v_θ across experiments. (Right) We align by the end of fight epoch.

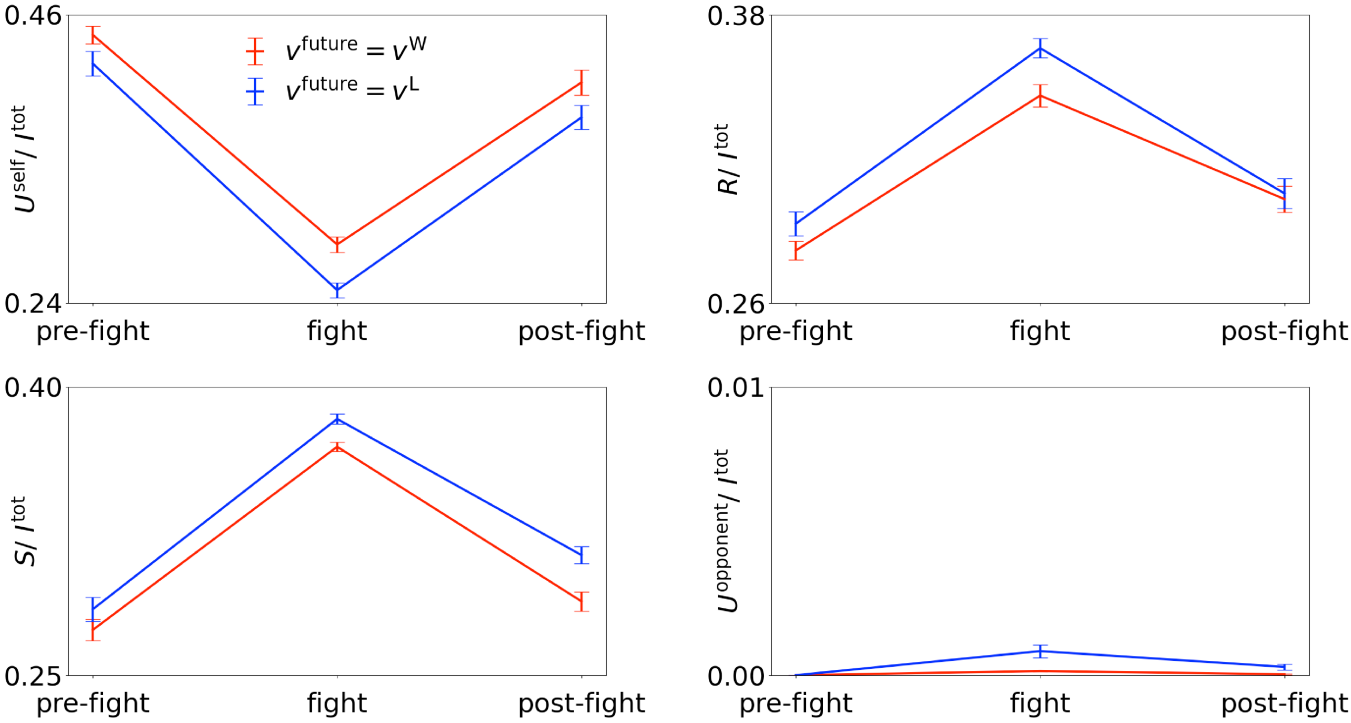

FIG. S5: **Normalized partial informations in v_z across epochs.** We compute information flows as with v_r and show the average and SEM across experiments. In contrast to the results from other coordinates, redundant information in the post-fight epoch is symmetric between winners and losers, implying that although the chase is dominated by winners, movements along the z-axis exhibit a similar degree of synchrony.

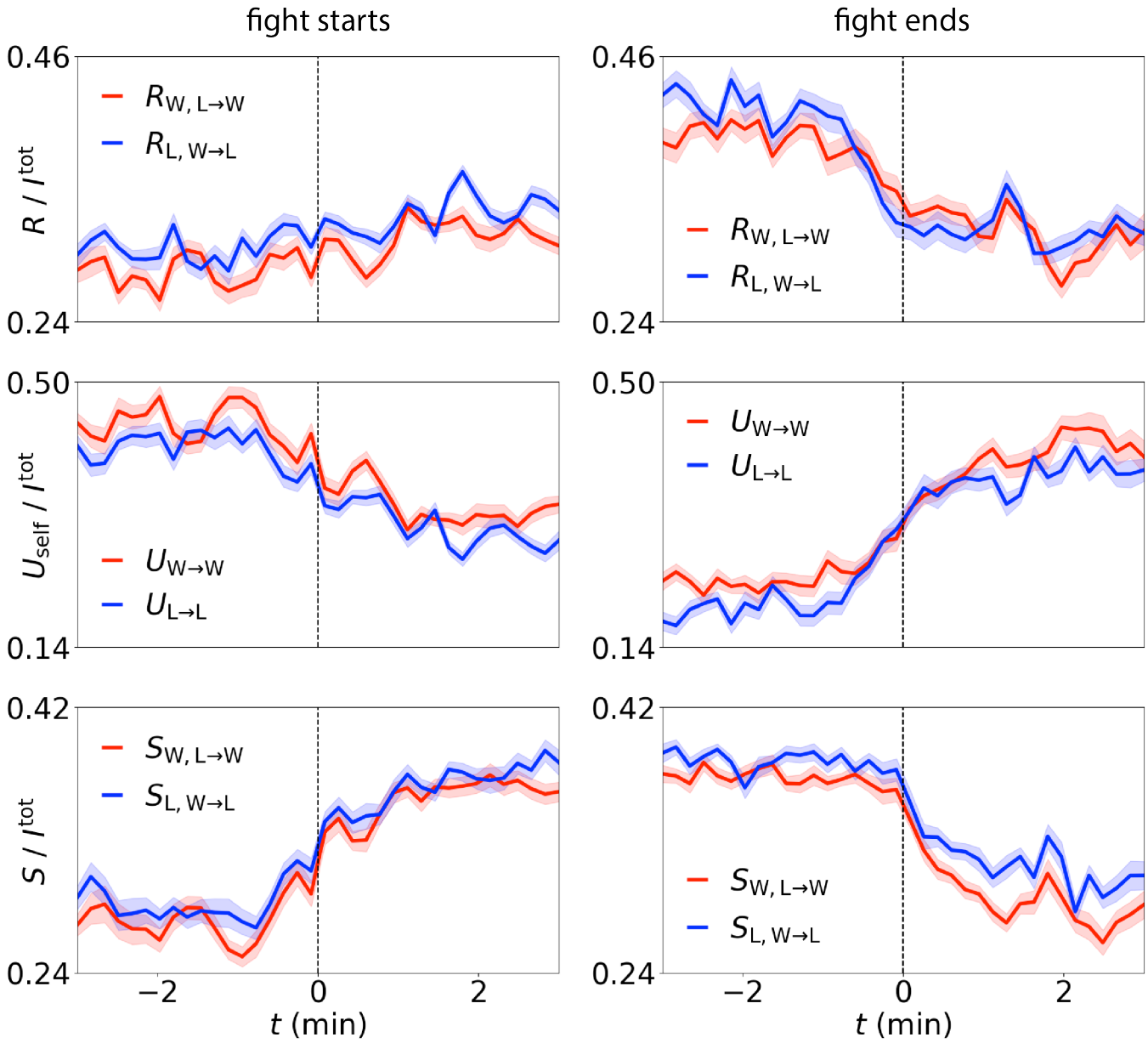

FIG. S6: **Dynamics of partial informations in v_z across contest boundaries.** (Left) We align by the start of the fight epoch and compute the average and SEM of the fraction of partial information in v_z across experiments. (Right) We align by the end of fight epoch.

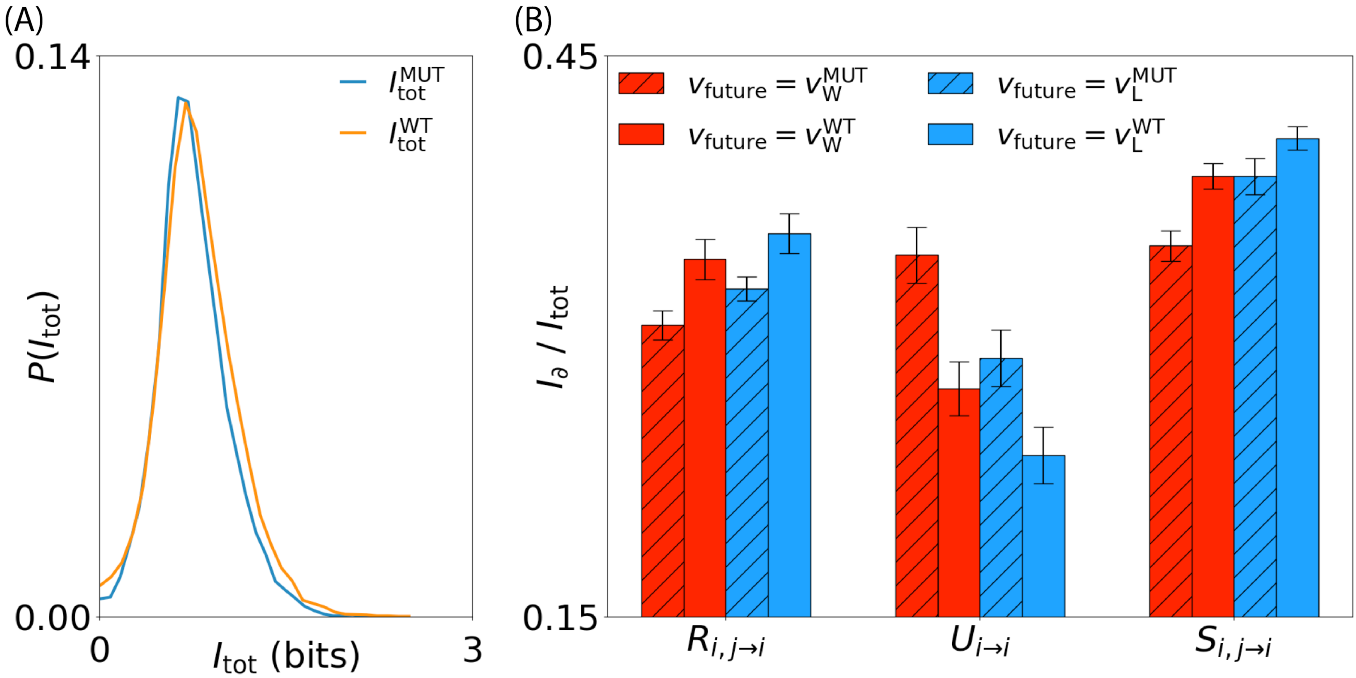

FIG. S7: **Predictive information in v_θ of *mecp2* mutant pairs and wild-type pairs.** We compute predictive information in v_θ from 30 s windows in 5 h recordings of mutant pairs and wild-type pairs. (A) We show the distributions of total information flow across windows (B) We show the normalized partial informations for winners and losers in both mutant and wild-type pairs.

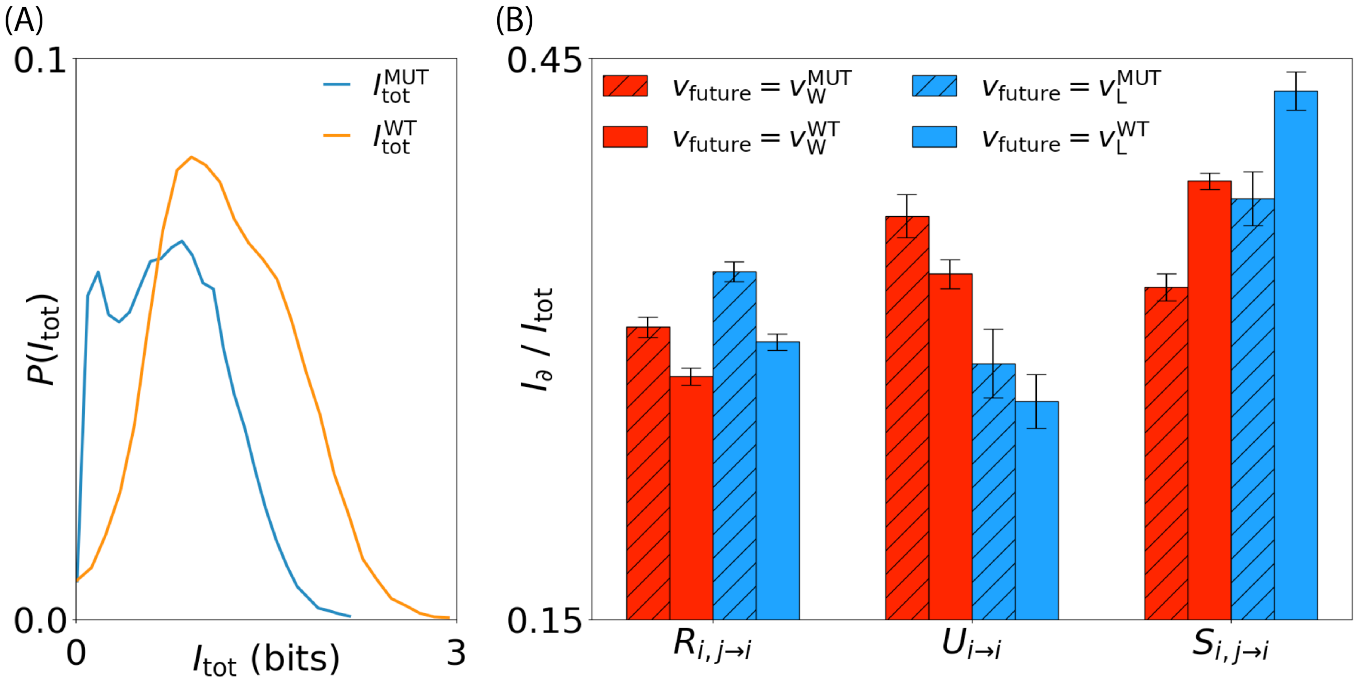

FIG. S8: **Predictive information in v_z of *mecp2* mutant pairs and wild-type pairs.** We compute predictive information in v_z from 30 s windows in 5 h recordings of mutant pairs and wild-type pairs. (A) We show the distribution of total information flow across windows. (B) We show the normalized partial informations for winners and losers in both mutant and wild-type pairs.

SI Movie 1: **Example social behaviors showing characteristic information flows.** We show reconstructed body-points movies with partial information calculated from the corresponding time-window. (Left) Pre-fight interactions with high self-unique information. (Middle) Cyclic attacks with high redundancy and synergy during fight. (Right) Post-fight chase dominated by the winner, showing asymmetric information flows. The movie is directly available here: https://tinyurl.com/r9ase6ep