

RESEARCH ARTICLE

Breakdown of local information processing may underlie isoflurane anesthesia effects

Patricia Wollstadt^{1*}, Kristin K. Sellers^{2,3}, Lucas Rudelt⁴, Viola Priesemann^{4,5*}, Axel Hutt^{6,7}, Flavio Fröhlich^{2,3,8,9,10,11}, Michael Wibral¹

1 MEG Unit, Brain Imaging Center, Goethe University, Frankfurt/Main, Germany, **2** Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA, **3** Neurobiology Curriculum, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA, **4** Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany, **5** Bernstein Center for Computational Neuroscience, BCCN, Göttingen, Germany, **6** Deutscher Wetterdienst, Section FE 12 - Data Assimilation, Offenbach/Main, Germany, **7** Department of Mathematics and Statistics, University of Reading, Reading, United Kingdom, **8** Department of Cell Biology and Physiology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA, **9** Department of Biomedical Engineering, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA, **10** Neuroscience Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA, **11** Department of Neurology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

* patricia.wollstadt@gmx.de (PW); viola@nld.ds.mpg.de (VP)



OPEN ACCESS

Citation: Wollstadt P, Sellers KK, Rudelt L, Priesemann V, Hutt A, Fröhlich F, et al. (2017) Breakdown of local information processing may underlie isoflurane anesthesia effects. PLoS Comput Biol 13(6): e1005511. <https://doi.org/10.1371/journal.pcbi.1005511>

Editor: Claus C. Hilgetag, Hamburg University, GERMANY

Received: August 25, 2016

Accepted: April 11, 2017

Published: June 1, 2017

Copyright: © 2017 Wollstadt et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data are within the paper and its Supporting Information files. Raw data available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.kk40s>.

Funding: Research reported in this publication was partially supported by the National Institute of Mental Health of the National Institutes of Health (NIH) under Award no. R01-MH-101547. We also received funding from the University of North Carolina Department of Psychiatry, University of North Carolina School of Medicine, and Foundation

Abstract

The disruption of coupling between brain areas has been suggested as the mechanism underlying loss of consciousness in anesthesia. This hypothesis has been tested previously by measuring the information transfer between brain areas, and by taking reduced information transfer as a proxy for decoupling. Yet, information transfer is a function of the amount of information available in the information source—such that transfer decreases even for unchanged coupling when less source information is available. Therefore, we reconsidered past interpretations of reduced information transfer as a sign of decoupling, and asked whether impaired local information processing leads to a loss of information transfer. An important prediction of this alternative hypothesis is that changes in locally available information (signal entropy) should be at least as pronounced as changes in information transfer. We tested this prediction by recording local field potentials in two ferrets after administration of isoflurane in concentrations of 0.0%, 0.5%, and 1.0%. We found strong decreases in the source entropy under isoflurane in area V1 and the prefrontal cortex (PFC)—as predicted by our alternative hypothesis. The decrease in source entropy was stronger in PFC compared to V1. Information transfer between V1 and PFC was reduced bidirectionally, but with a stronger decrease from PFC to V1. This links the stronger decrease in information transfer to the stronger decrease in source entropy—suggesting reduced source entropy reduces information transfer. This conclusion fits the observation that the synaptic targets of isoflurane are located in local cortical circuits rather than on the synapses formed by interareal axonal projections. Thus, changes in information transfer under isoflurane seem to be a consequence of changes in local processing more than of decoupling between brain areas. We suggest that source entropy changes must be considered whenever interpreting changes in information transfer as decoupling.

of Hope. VP received financial support from the German Ministry for Education and Research (BMBF) via the Bernstein Center for Computational Neuroscience (BCCN) Göttingen under Grant No. 01GQ1005B. The authors acknowledge support by the Human Brain Project (EU Grant No. 604102). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Author summary

Currently we do not understand how anesthesia leads to loss of consciousness (LOC). One popular idea is that we lose consciousness when brain areas lose their ability to communicate with each other—as anesthetics might interrupt transmission on nerve fibers coupling them. This idea has been tested by measuring the amount of information transferred between brain areas, and taking this transfer to reflect the coupling itself. Yet, information that isn't available in the source area can't be transferred to a target. Hence, the decreases in information transfer could be related to less information being available in the source, rather than to a decoupling. We tested this possibility measuring the information available in source brain areas and found that it decreased under isoflurane anesthesia. In addition, a stronger decrease in source information lead to a stronger decrease of the information transferred. Thus, the input to the connection between brain areas determined the communicated information, not the strength of the coupling (which would result in a stronger decrease in the target). We suggest that interrupted information processing within brain areas has an important contribution to LOC, and should be focused on more in attempts to understand loss of consciousness under anesthesia.

Introduction

To this day it is an open question in anesthesia research how general anesthesia leads to loss of consciousness (LOC). Several recent theories agree in proposing that anesthesia-induced LOC may be caused by the disruption of long range inter-areal information transfer in cortex [1–5]—a hypothesis supported by a series of recent studies [3, 6–9]. In all of these studies, information transfer is quantified using transfer entropy [10], an information theoretic measure, which has become a quasi-standard for the estimation of information transfer in anesthesia research, or by transfer entropy's linear implementation as a Granger causality. In many of these studies reduced information transfer has been interpreted as a sign of inter-areal long range connectivity being disrupted by anesthesia.

Yet, information transfer between a source of information and a target depends on information (entropy) being available at the source in the first place. Considering constraints of this kind we can easily conceive of cases where a decrease in information transfer under anesthesia is observed despite unchanged long range coupling, e.g., when the available information at the source decreases due to an anesthesia-related change in *local* information processing. Ultimately, this dissociation between information transfer and causal coupling just reflects that information transfer is *one* possible consequence of physical coupling, but not identical to it [11–13].

Therefore, we consider it necessary to evaluate the hypothesis that the reduced inter-areal information transfer observed under anesthesia possibly originates from disrupted information processing in local circuits rather than from disrupted long range connectivity. This alternative hypothesis receives additional support for the case of isoflurane, which potentiates agonist actions at *GABA_A*-receptors and inhibits nicotinic acetylcholine (nAChR) receptors. Conversely, evidence on direct inhibitory effects of isoflurane on AMPA and NMDA synapses, which are the dominant mediators of long-range cortico-cortical interactions, is sparse at best (see table 2 in [14]). Under the alternative hypothesis of changed local information processing, a decrease in transfer entropy under anesthesia must be accompanied by:

1. a reduction in locally available information per brain area, i.e. in the sources of information transfer,
2. and the fact that the strongest decrease in locally available information is found at the source of the link with the strongest decrease in information transfer, rather than at its target (i.e. the end point),

Here, we perform tests of these predictions by estimating local information processing in and information transfer between local field potentials (LFPs) simultaneously recorded from primary visual cortex (V1) and prefrontal cortex (PFC) of two ferrets under different levels of isoflurane. We quantify local information processing by estimating the signal entropy (measuring available information) and quantified information transfer between recording sites by estimating transfer entropy. Additionally, to demonstrate the effect of reduced source entropy on transfer entropy, we estimated transfer entropy on simulated data with a constant coupling between processes, but a varying source entropy.

To better understand potential changes in local information processing we also quantified the active information storage, a measure of the information available at a recording site that can be predicted from past signals at that site, i.e. the information stored from past to present.

Because the estimation of such information theoretic quantities from finite data is difficult in general, we employ two complementary strategies: (i) probability density estimation based on nearest-neighbor searches in continuous data, and (ii) Bayesian estimation based on discretized data.

We also test whether the previously reported decrease of transfer entropy under anesthesia can indeed be replicated when avoiding some recently identified pitfalls in estimation of information transfer related to the use of symbolic time series, suboptimal embedding of the time series, and the use of net transfer entropy without identification of the individual information transfer delays (for problems related to these approaches see [15]).

Our results provide first evidence for the alternative hypothesis of altered local information processing causing reduced information transfer, as the above predictions were indeed met. We suggest to consider the alternative hypothesis as a serious candidate mechanism for LOC, and to use causal interventions to gather further experimental evidence.

Preliminary results for this study were published in abstract form in [16].

Results

Existence of information transfer between recording sites

Before analyzing differences in information transfer induced by isoflurane, we tested for the existence of significant information transfer between the recording sites (Table 1). For both animals, TE_{SPO} was significant in the top-down direction, while the bottom-up direction was significant for animal 1 only. A non-significant information transfer in bottom-up direction for animal 2 may be explained by the dark experimental environment, i.e., the lack of visual input.

Changes in information theoretic measures under anesthesia

Overall, for higher isoflurane levels both locally available information and information transfer were decreased, while information storage in local activity increased.

As the estimation of information theoretic measures from finite length neural recordings poses a considerable challenge we present detailed, converging results from two complementary strategies to deal with this challenge—nearest-neighbor based estimators, and a Bayesian approach to entropy estimation suggested by Nemenman, Shafe, and Bialek (NSB-estimator)

Table 1. Results significance test of TE_{SPO} estimates in both animals and for both directions of interaction.

animal	direction of interaction	p-value
Ferret 1	PFC → V1	<0.01**
	V1 → PFC	<0.05*
Ferret 2	PFC → V1	<0.05*
	V1 → PFC	0.2262 n.s.

* $p < 0.05$;

** $p < 0.01$;

*** $p < 0.001$;

Bonferroni-corrected

<https://doi.org/10.1371/journal.pcbi.1005511.t001>

[17, 18]. This latter approach required a discretization of the continuous-valued LFP data, but yields principled control of bias, while the first approach allows the estimation of information-theoretic measures directly from continuous data, and thus conserves the information originally present in those data. Statistical testing was performed using a nonparametric permutation ANOVA (pANOVA), and a linear mixed model (LMM). The LMM approach was used in addition to the main pANOVA for the purpose of comparison to older studies using parametric statistics.

Results based on next neighbor-based estimation from continuous data. For higher isoflurane levels, we found an overall reduction in the locally available information (H), and in the information transfer (TE_{SPO}). We found an increase in the locally predictable (stored) information (AIS) (Table 2 and Fig 1).

In general, H decreased in both animals under higher isoflurane levels (main effect *isoflurane level*, $p < 0.01^{**}$ for ferret 1 and $p < 0.001^{***}$ for ferret 2), indicating a reduction of locally available information for higher isoflurane concentrations. Yet, an isoflurane concentration of 0.5% (abbreviated as *iso* 0.5%, with other concentrations abbreviated accordingly) led to a slight increase in H for ferret 1, followed by a decrease for concentration *iso* 1.0%, which was below initial entropy values. This rise in H in condition *iso* 0.5% was present only in V1 of ferret 1, while H decreased monotonically in PFC. In ferret 2, H increased monotonically in both

Table 2. Results of permutation analysis of variance for information theoretic measures (p-values).

measure	effect	Ferret 1	Ferret 2
TE_{SPO}	<i>isoflurane level</i>	<0.0001***	<0.0001***
	<i>direction</i>	0.0003***	0.9345
	<i>interaction</i>	0.0052**	0.2486 ^a
AIS	<i>isoflurane level</i>	0.0017**	<0.0001***
	<i>recording site</i>	0.6774	<0.0001***
	<i>interaction</i>	<0.0001***	0.0029**
H	<i>isoflurane level</i>	0.0010**	<0.0001***
	<i>recording site</i>	0.9326	<0.0001***
	<i>interaction</i>	<0.0001***	0.0243*

* $p < 0.05$;

** $p < 0.01$;

*** $p < 0.001$;

^aThis effect was significant when using LMM for statistical testing

<https://doi.org/10.1371/journal.pcbi.1005511.t002>

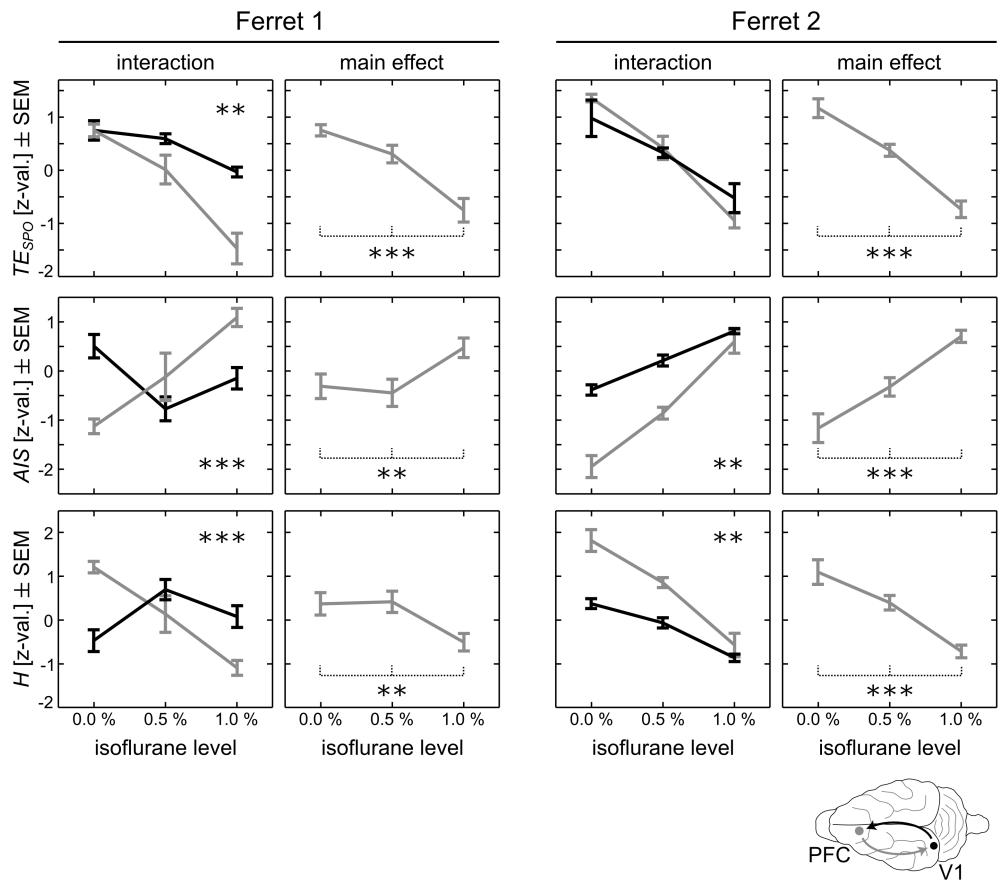


Fig 1. pANOVA results for nearest-neighbor based estimates of transfer entropy (TE_{SPO}), active information storage (AIS), and entropy (H). Left columns show interactions $isoflurane\ level \times direction$ and $isoflurane\ level \times recording\ site$ for both animals; right columns show main effects $isoflurane\ level$. Grey lines in interaction plots indicate TE_{SPO} from prefrontal cortex (PFC) to primary visual areas (V1), or H and AIS in PFC; black lines indicate TE_{SPO} from V1 to PFC, or H and AIS in V1. Error bars indicate the standard error of the mean (SEM); stars indicate significant interactions or main effects (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$). Axis units for all information theoretic measures based on continuous variables are z-normalized values across conditions.

<https://doi.org/10.1371/journal.pcbi.1005511.g001>

recording sites, with a stronger decrease in PFC. The interaction effect ($isoflurane\ level \times brain\ region$) was significant for both animals ($p < 0.001^{***}$ for ferret 1 and $p < 0.01^{**}$ for ferret 2).

The information transfer as measured by the self prediction optimal transfer entropy (TE_{SPO} , [15]) decreased significantly with higher levels of isoflurane in both animals (main effect $isoflurane\ level$, $p < 0.001^{***}$), indicating an overall reduction in information transfer. This reduction was stronger in the top-down direction $PFC \rightarrow V1$ (significant interaction effect $isoflurane\ level \times direction$ in ferret 1, $p < 0.01^{**}$). In ferret 2 this interaction was not significant in the permutation ANOVA (pANOVA) on aggregated data, but was highly significant using the LMM approach (see next section).

The stored information as measured by the active information storage (AIS , [19]) increased in both animals under higher isoflurane levels (main effect $isoflurane\ level$, $p < 0.01^{**}$ for ferret 1 and $p < 0.001^{***}$ for ferret 2), indicating more predictable information in LFP signals under higher levels of isoflurane. In ferret 1 the concentration $iso\ 0.5\%$ led to a slight decrease in AIS , followed by an increase compared to initial levels for concentration $iso\ 1.0\%$. This initial

decrease in *AIS* in condition *iso* 0.5% was present only in V1 of this animal, while in its PFC *AIS* increased monotonically. In ferret 2, *AIS* increased monotonically in both recording sites, with a stronger increase in PFC. The interaction effect was significant for both animals ($p < 0.001^{***}$ for ferret 1 and $p < 0.01^{**}$ for ferret 2). Overall, *AIS* behaved complementary to *H* for all animals and isoflurane levels, despite the fact that *AIS* is one component of *H* [19].

Alternative statistical testing using linear mixed models. We additionally performed a parametric test, using linear mixed models (LMM) on non-aggregated data from individual epochs of recording sessions (adding *recording* as additional random factor to encode the recording session) to enable a comparison to results from earlier studies using parametric testing.

For both animals, model comparison showed a significant main effect of factor *isoflurane level* on TE_{SPO} , *AIS*, and *H* as well as a significant interaction of factors *isoflurane level* and *direction* (see Table 3). The only exception to this was the factor direction in the evaluation of TE_{SPO} in Ferret 2. Thus, the results of this alternative statistical analysis were in agreement

Table 3. Results of parametric statistical testing using model comparison between linear mixed models. Simple effects of factors *isoflurane level* and *direction* or *recording site* were tested against the null model; interaction effects *isoflurane level times direction* and *isoflurane level times recording site* were tested against the additive models *isoflurane level + direction* and *isoflurane level + recording site*, respectively. *** as defined in Table 2.

animal	measure	effect	BIC	deviance	χ^2	χ^2 df	p
Ferret 1	TE_{SPO}	<i>null model</i>	345430.8	345401.7	NA	NA	NA
		<i>isoflurane level</i>	345421.5	345373.1	28.58	2	<< 0.000***
		<i>direction</i>	342828.4	342789.7	2612.04	1	<< 0.000***
		<i>isoflurane l. + direction</i>	342819.3	342761.3	NA	NA	NA
		<i>isoflurane l. × direction</i>	342139.9	342062.4	698.84	2	<< 0.000***
	<i>AIS</i>	<i>null model</i>	37581.2	37551.8	NA	NA	NA
		<i>isoflurane level</i>	37555.8	37506.8	46.00	2	<< 0.000***
		<i>recording site</i>	33296.2	33256.9	4294.82	1	<< 0.000***
		<i>isoflurane l. + recording site</i>	33270.9	33212.1	NA	NA	NA
		<i>isoflurane l. × recording site</i>	30754.3	30675.8	2536.29	2	<< 0.000***
	<i>H</i>	<i>null model</i>	353650.8	353621.8	NA	NA	NA
		<i>isoflurane level</i>	353653.1	353604.7	17.03	2	<0.000***
		<i>recording site</i>	353446.8	353408.1	213.64	1	<< 0.000***
		<i>isoflurane l. + recording site</i>	353449.1	353391.1	NA	NA	NA
		<i>isoflurane l. × recording site</i>	347204.7	347127.2	6263.85	2	<< 0.000***
Ferret 2	TE_{SPO}	<i>null model</i>	131164.7	131135.3	NA	NA	NA
		<i>isoflurane level</i>	131152.3	131103.2	32.06	2	<< 0.000***
		<i>direction</i>	131168.9	131129.6	5.66	1	0.020
		<i>isoflurane l. + direction</i>	131109.5	131097.5	NA	NA	NA
		<i>isoflurane l. × direction</i>	130871.5	130793.0	304.56	2	<< 0.000***
	<i>AIS</i>	<i>null model</i>	109203.1	109173.7	NA	NA	NA
		<i>isoflurane level</i>	109183.1	109134.1	39.63	2	<< 0.000***
		<i>recording site</i>	105827.2	105788.0	3385.68	1	<< 0.000***
		<i>isoflurane l. + recording site</i>	105807.3	105748.4	NA	NA	NA
		<i>isoflurane l. × recording site</i>	104173.5	104095.0	1653.46	2	<< 0.000***
	<i>H</i>	<i>null model</i>	-14074.8	-14104.2	NA	NA	NA
		<i>isoflurane level</i>	-14088.6	-14137.7	33.43	2	<0.000***
		<i>recording site</i>	-18423.0	-18462.3	4358.04	1	<< 0.000***
		<i>isoflurane l. + recording site</i>	-18436.7	-18495.6	NA	NA	NA
		<i>isoflurane l. × recording site</i>	-19867.2	-19945.7	1450.09	2	<< 0.000***

<https://doi.org/10.1371/journal.pcbi.1005511.t003>

with those from the pANOVA. The detailed [Table 3](#) reports the Bayesian Information Criterion with $BIC = -2 \log p(\mathbf{x}|M, \hat{a}) + k(a)\log(N)$, where \mathbf{x} are the observed realizations of the data, \hat{a} are the parameters that optimize the likelihood for a given model M , $k(a)$ is the number of parameters and N the number of data points. The BIC becomes smaller with better model fit. In addition, [Table 3](#) gives the deviance, $-2 \log p(\mathbf{x}|M, \hat{a})$, which is higher for better model fit, and the χ^2 with the corresponding p -value, describing the likelihood ratio, which follows a χ^2 -distribution.

Bayesian estimation on discretized data. In addition to the neighbor-distance based estimators for TE_{SPO} , AIS , and H used above, we also applied Bayesian estimators recently proposed by Nemenman, Shafe, and Bialek (NSB) [17, 18] to our data. The Bayesian approach promises to yield unbiased estimators when priors are chosen appropriately. However, one has to keep in mind that these estimators currently require the discretization of continuous data, and therefore may lose important information.

When applying the NSB estimator to the discretized LFP time series with N_{bins} discretization steps, we observed that for $N_{bins} \geq 8$ the results were qualitatively consistent for different choices of numbers of bins. We present results for $N_{bins} = 12$ ([Fig 2](#)), which provides a

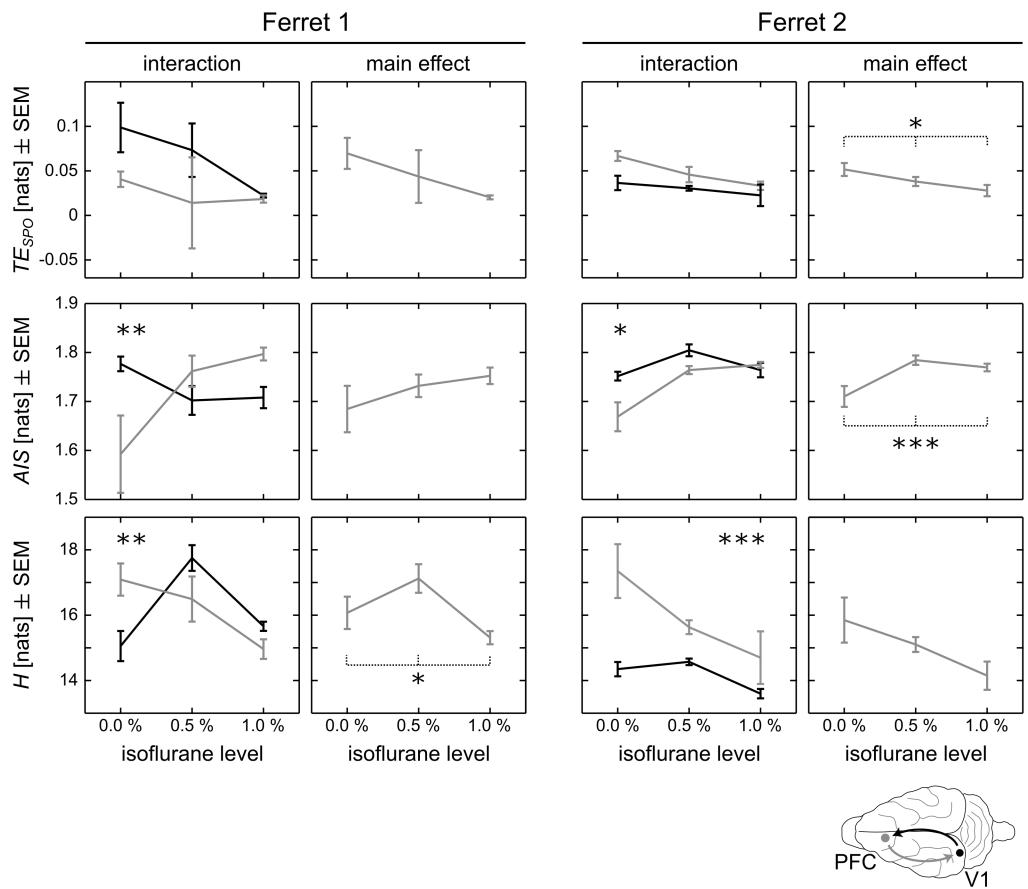


Fig 2. pANOVA results for Bayesian estimates of transfer entropy (TE_{SPO}), active information storage (AIS), and entropy (H). Left columns show interactions $isoflurane\ level \times direction$ and $isoflurane\ level \times recording\ site$ for both animals; right columns show main effects $isoflurane\ level$. Grey lines in interaction plots indicate TE_{SPO} from prefrontal cortex (PFC) to primary visual areas (V1), or H and AIS in PFC; black lines indicate TE_{SPO} from V1 to PFC, or H and AIS in V1. Error bars indicate the standard error of the mean (SEM); stars indicate significant interactions or main effects (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$). 

<https://doi.org/10.1371/journal.pcbi.1005511.g002>

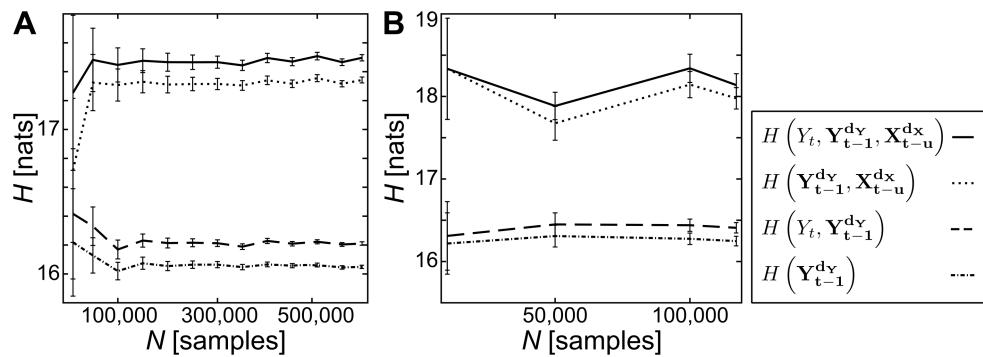


Fig 3. Examples for estimates of entropy (H) terms for transfer entropy calculation (Eq 12) by number of data points N , using the Nemenman-Shafee-Bialek-estimator (NSB). (A) entropies for ferret 1, V1, *iso* 0.5%: estimates are stable for $N \geq 100,000$; (B) entropies for ferret 2, PFC, *iso* 0.0%: an insufficient number of data points does not allow for verification of the estimate's robustness (recording was excluded from further analysis). Variables labeled X reflect data from the source variable, Y from the target variable. t is an integer time index, and bold typeface indicates the state of a system (see Methods).

<https://doi.org/10.1371/journal.pcbi.1005511.g003>

reasonable resolution of the signal while still allowing for a reliable estimation of entropies within the scope of available data.

We confirmed the reliability of the estimator by systematically reducing the sample size and found no substantial impact on our estimates (Fig 3).

The estimation of H , AIS , and TE_{SPO} by Bayesian techniques for the binned signal representations provided results that were qualitatively consistent with results from the neighbor-distance based estimation techniques (compare Figs 1 and 2, and Tables 2 and 4). While the Bayesian estimates showed larger variances across different recordings and sample sizes, on average, we saw a decrease of TE_{SPO} and H for higher concentrations of isoflurane (main effect *isoflurane level*), while AIS increased for higher concentrations. For ferret 1 we also found an interaction effect for TE_{SPO} , with a stronger reduction in information transfer in top-down direction.

Note that we also applied an alternative Bayesian estimation scheme based on Pitman-Yor process priors [20]. However, for this estimation procedure, we observed that the data were

Table 4. Results of permutation analysis of variance for information theoretic measures obtained through Bayesian estimation (p-values).

measure	effect	Ferret 1	Ferret 2
TE_{SPO}	<i>isoflurane level</i>	0.0549	0.0013*
	<i>direction</i>	0.0272*	0.0002***
	<i>interaction</i>	0.6627	0.0616
AIS	<i>isoflurane level</i>	0.2148	<0.0001***
	<i>direction</i>	0.0788	0.0032**
	<i>interaction</i>	0.0026**	0.0153*
H	<i>isoflurane level</i>	0.0184*	<0.0001***
	<i>recording site</i>	0.2738	<0.0001***
	<i>interaction</i>	0.0017**	0.0053

* $p < 0.05$;

** $p < 0.01$;

*** $p < 0.001$

<https://doi.org/10.1371/journal.pcbi.1005511.t004>

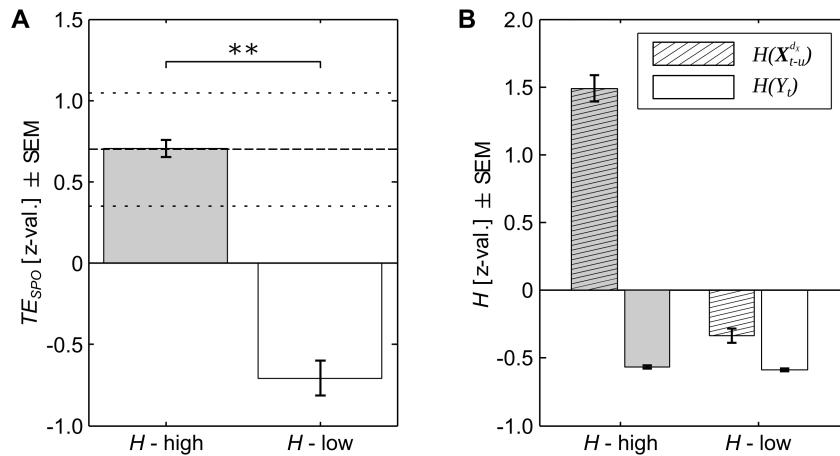


Fig 4. Simulated effects of changed source entropy on transfer entropy (TE_{SPO}). (A) TE_{SPO} for two simulated cases of high (*H*-high) and low (*H*-low) source entropy (** $p < 0.01$, error bars indicate the standard error of the mean, SEM, over data epochs). Dashed lines indicate the average TE_{SPO} estimated from the original data under 0.0% isoflurane \pm SEM; (B) source entropy $H(X_{t-u}^{d_X})$ (dashed bars) and target entropy $H(Y_t)$ (empty bars) for the two simulated test cases of high (gray bars) and low entropy (white bars), error bars indicate the SEM over data epochs. Source entropy was higher in the high-entropy simulation, while target entropy was approximately the same for both cases. Results are given as z-values across estimates for all epochs from both simulations.

<https://doi.org/10.1371/journal.pcbi.1005511.g004>

insufficient to allow for a robust estimation of the tailing behavior of the distribution as indicated by large variances and unreasonably high estimates across different sample sizes.

Simulated effects of changed source entropy on transfer entropy. To test the effect of reduced source entropy on TE_{SPO} between a source and target process, we simulated two test cases with high and low source entropy respectively. In both test cases signals were based on the original recordings, and the coupling between source and target was held constant (see [Methods](#)).

We found significantly higher TE_{SPO} in the case with high source entropy than in the case with low source entropy (Fig 4A). The simulation thus demonstrated that a lower source entropy does indeed lead to a reduction in TE_{SPO} despite an unchanged coupling. Information transfer in the high-entropy case was similar to the average information transfer found in recordings under an isoflurane concentration of 0.0%, indicating that the simulation scheme mirrored information transfer found in the experimental recordings.

Optimized embedding parameters for TE_{SPO} and AIS estimation

As noted in the introduction, the estimation of information theoretic measures from finite data is challenging. For the measures that describe distributed computation in complex systems, such as transfer entropy, estimation is further complicated because the available data are typically only scalar observations of a process with multidimensional dynamics. This necessitates the approximate reconstructions of the process' states via a form of embedding [21], where a certain number of past values of the scalar observations spaced by an embedding delay are taken into account (e.g. for a pendulum swinging through its zero position one additional past position value will clarify whether it's going left or right). An important part of proper transfer entropy estimation is thus optimization of this number of past values (embedding dimension), and of the embedding delay. These two embedding parameters then approximately define past states, whose correct identification is crucial for the estimation of transfer

entropy [22], but also for the estimation of active information storage. Without it information storage may be underestimated and, erroneous values of the information transfer will be obtained; even a detection of information transfer in the wrong direction is likely (see [15, 22, 23] and [Methods](#) section). Existing studies using transfer entropy often omitted the optimization of embedding parameters, and instead used ad-hoc choices, which may have had a detrimental effect on the estimation of transfer entropy—hence the need for a confirmation of previous results in this study.

In the present study, we therefore used a formal criterion proposed by Ragwitz [21], to find an optimal embedding defined by the embedding dimension d (the number of values collected) and delay τ (the temporal spacing between them), to find embeddings for the TE_{SPO} and AIS past states. We used the implementation of this criterion provided by the TRENTOOL toolbox. Since the bias of the estimators used depends on d , we used the maximum d over all conditions and directions of interaction as the common embedding dimension for estimation from each epoch to make the estimated values statistically comparable. The resulting dimension used was 15 samples, which is considerably higher than the value commonly used in the literature [6, 7], when choosing the embedding dimension ad-hoc or using other criteria as for example the sample size [9]. The embedding delay τ was optimized individually for single epochs in each condition and animal as it has no influence on the estimator bias.

Relevance of individual information transfer delay estimation

Several previous studies on information transfer under anesthesia reported the sign of the so called net transfer entropy (TE_{net}) as a measure of the predominant direction of information transfer between two sources. TE_{net} is essentially just the normalized difference between the transfer entropies measured in the two directions connecting a pair of recording sites (see [Methods](#)). When calculating TE_{net} it is particularly important to individually account for physical delays δ in the two potential directions of information transfer, because otherwise the sign of TE_{net} may become meaningless (see examples in [15]). As this delay is unknown a priori it has to be found prior to the actual estimation of information transfer. We have recently shown that this can be done by using a variable delay parameter u in a delay sensitive transfer entropy estimator TE_{SPO} (see [Methods](#)); here, the u that maximizes the transfer entropy over a range of assumed values for u reflects the physical delay [15].

The necessity to individually optimize u for each interaction is not a mere theoretical concern but was clearly visible in the present study: [Fig 5C](#) shows representative results from ferret 1 under 0.0% isoflurane, where apparent TE_{SPO} values strongly varied as a function of the delay u . As a consequence, TE_{net} values also varied if a *common* delay u was chosen for both directions. In other words, the sign of the TE_{net} varied as a function of individual choices for $u_{PFC \rightarrow V1}$ and $u_{V1 \rightarrow PFC}$ for each direction of information transfer and hence became uninterpretable ([Fig 5D](#)).

As a consequence of the above, we here individually optimized u for each direction of information transfer in each condition and animal to estimate the true delay of information transfer following the mathematical proof in [15]. We individually optimized u to obtain estimates of transfer entropy that were not biased by a non-optimal choice for u . We used the implementation in TRENTOOL [23] and scanned values for u ranging from 0 to 20 ms. Averages for optimized delays ranged from 4–7 ms across animals and isoflurane levels ([Fig 6](#)).

Note that in [Fig 5C](#), TE_{SPO} as a function of u shows multiple peaks, especially for the direction $V1 \rightarrow PFC$. Since these peaks resembled an oscillatory pattern close to the low-pass filter frequency used for preprocessing the data, we investigated the influence of filtering on delay reconstruction by simulating two coupled time series for which we reconstructed the delay

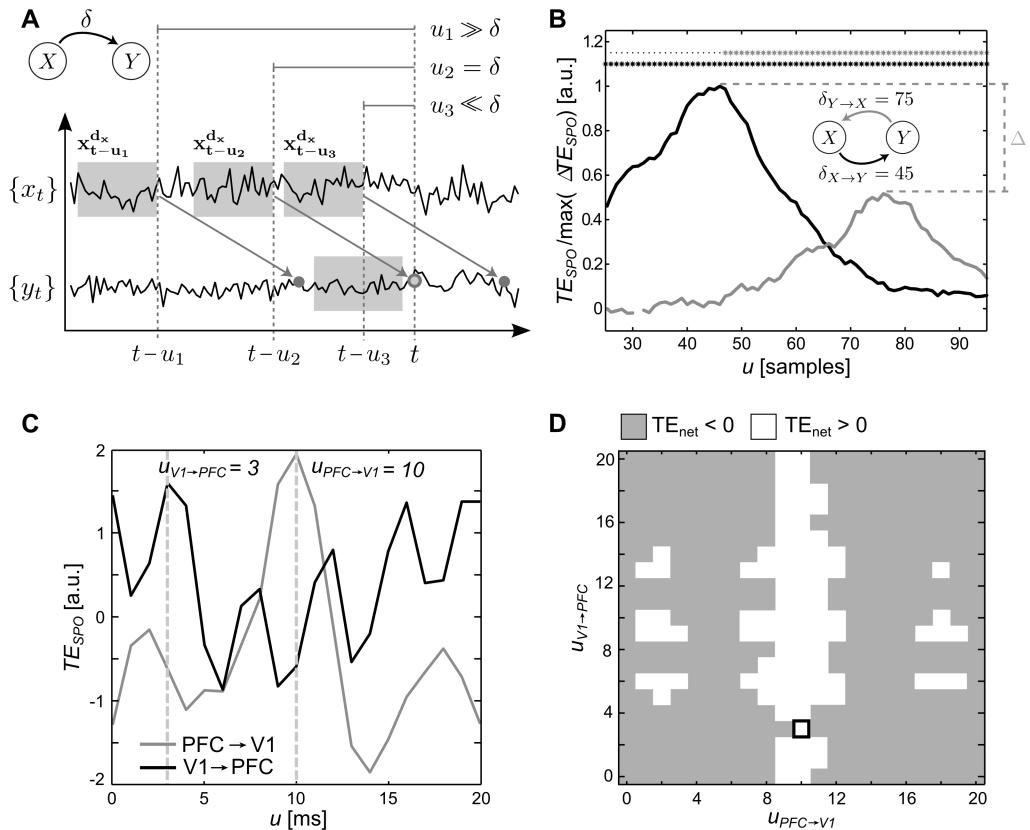


Fig 5. Estimation of information transfer delays. (A, modified from [15]) estimation of transfer entropy (TE_{SPO}) depends on the choice of the delay parameter u , if u is much smaller or bigger than the true delay δ , information arrives too late or too early in the target time series and information transfer is not correctly measured; (B, modified from [15]) TE_{SPO} values estimated from two simulated, bidirectionally coupled Lorenz systems (see [15] for details) as a function of u for both directions of analysis, $TE_{SPO}(X \rightarrow Y, t, u)$ (black line) and $TE_{SPO}(Y \rightarrow X, t, u)$ (gray line); TE_{SPO} values vary with the choice of u and so does the absolute difference between values; using individually optimized transfer delays for both directions of analysis yields the meaningful difference Δ (dashed lines), where $TE_{SPO}(X \rightarrow Y, t, u_{opt}) > TE_{SPO}(Y \rightarrow X, t, u_{opt})$; (C) example transfer entropy analysis for one recorded epoch: TE_{SPO} values vary greatly as a function of u , optimal choices of u are marked by dashed lines; (D) sign (gray: negative, white: positive) of TE_{net} for different values of $u_{PFC \rightarrow V1}$ (x-axis) and $u_{V1 \rightarrow PFC}$ (y-axis), calculated from TE_{SPO} values shown in panel C: the sign varies with individual choices of u ; the black frame marks the combination of individually optimal choices for both parameters that yields the correct result.

<https://doi.org/10.1371/journal.pcbi.1005511.g005>

with and without prior band-pass filtering (Fig 7). For filtered, simulated data TE_{SPO} as a function of u indeed showed an oscillatory pattern for certain filter settings. However, broadband filtering—as used here for the original data—did not lead to the reconstruction of an incorrect value for the delay δ (Fig 7C). Yet, when filtering using narrower bandwidths, the reconstruction of the correct information transfer delay failed (Fig 7D–7F). This finding supports the general notion that filtering, especially narrow-band filtering, may only be applied with caution before estimating connectivity measures [24, 25].

Yet, since broad-band filtering did not lead to the observed oscillatory patterns in the simulated data, alternative explanations for multiple peaks in TE_{SPO} are more likely: one alternative cause of multiple peaks in the TE_{SPO} are multiple information transfer channels with various delays between source and target (see [15], especially Fig 6). Each information transfer channel with its individual delay will be detected when estimating TE_{SPO} as a function of

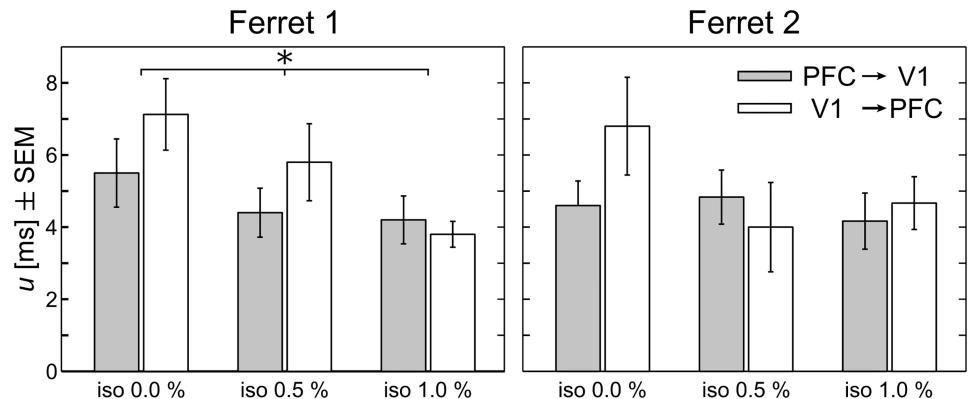


Fig 6. Optimized information transfer delays u for both directions of interaction and three levels of isoflurane, by animal. Bars denote averages over recordings per condition; error bars indicate the standard error of the mean (SEM). There was a significant main effect of isoflurane level for ferret 1 ($p < 0.05$).

<https://doi.org/10.1371/journal.pcbi.1005511.g006>

u (see simulations in [15]). A further possible cause for multiple peaks in the TE_{SPO} is the existence of a feedback loop between the source of TE_{SPO} and a third source of neural activity [15]. In such a feedback loop, information in the source is fed back to the source such that it occurs again at a later point in time, leading to the repeated transfer of identical bits of information. Both causes are potential explanations for the occurrence of multiple peaks in the TE_{SPO} in the present study—yet, deciding between these potential causes requires additional research, specifically interventional approaches to the causal structure underlying the observed information transfer.

In sum, our results clearly indicate the necessity to individually optimize information transfer delays and that often employed ad-hoc choices for u may result in spurious results in information transfer. The additional simulations of filtering effects showed that narrow-band filtering may have detrimental effects on this optimization procedure and should thus be avoided.

Relation of information-theoretic measures with other time-series properties

The information-theoretic measures used in this study are relatively novel and have been applied in neuroscience research rarely. Thus, it is conceivable that these measures quantify signal properties that are more easily captured by traditional time-series measures, namely, the autocorrelation decay time (ACT), the signal variance, and the power in individual frequency bands. To investigate the overlap between information-theoretic and traditional measures, we calculated correlations between AIS and ACT, between H and signal variance, between TE_{SPO} and ACT, between TE_{SPO} and signal variance, and between AIS, H , and TE_{SPO} and the power in various frequency bands, respectively. Fig 8 shows average ACT, signal variance, and power over recordings, for individual isoflurane levels and recording sites, and for both animals.

Detailed tables showing correlation coefficients and explained variances are provided as supporting information S1 to S6 Tables. In summary, for AIS and ACT we found significant correlations in both animals, however the median of the variance explained, \hat{R}^2 , over individual recordings was below 0.1 for each significant correlation in an animal, recording site and isoflurane level. Hence, even though there was some shared variance between ACT and AIS, there remained a substantial amount of unexplained variance in AIS, indicating that AIS

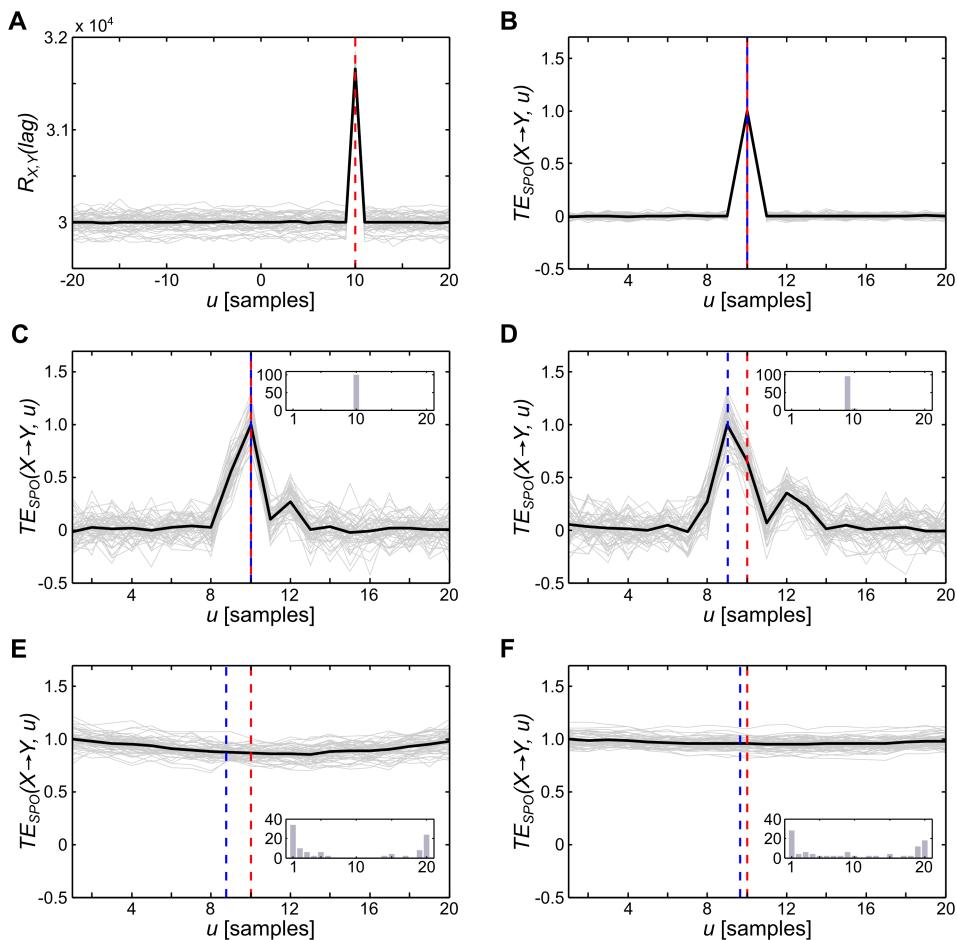


Fig 7. Simulated effects of filtering on the reconstruction of information transfer delays (u). (A) Cross-correlation $R_{X,Y}$ between two simulated, coupled time series ($N = 100000$, drawn from a uniform random-distribution over the open interval $(0, 1)$, true coupling delay 10 samples, indicated by the red dashed line), the simulation was run 50 times, black lines indicate the mean over simulation runs, gray lines indicate individual runs; (B) TE_{SPO} as a function of information transfer delay u before filtering the data; (C–D) TE_{SPO} as a function of u after filtering the data with a bandpass filter (fourth order, causal Butterworth filter, implemented in the MATLAB toolbox FieldTrip [26]); red dashed lines indicate the simulated delay δ , blue lines indicate the average reconstructed delay u over simulation runs, histograms (inserts) show the distribution of reconstructed values for u over simulation runs in percent; (C) broadband filtering (0.1–300 Hz) introduced additional peaks in TE_{SPO} values, however, the maximum peak indicating the optimal u was still at the simulated delay for all simulated runs; (D) filtering within a narrower band (0.1–200 Hz) led to an imprecise reconstruction of the correct δ in each run with an error of 1 sample, i.e. δ tended to be underestimated; (E–F) narrow-band filtering in the beta range (12–30 Hz) and theta range (4–8 Hz) led to a wide distribution of u with a large absolute error of up to 10 samples.

<https://doi.org/10.1371/journal.pcbi.1005511.g007>

quantified properties other than the signal's ACT—in line with results from our earlier studies [27, 28]. This is expected because the autocorrelation (decay) time ACT measures how long information persists in a time series in a linear encoding, while AIS, in contrast, measures how much information is stored (at any moment), and also reflects nonlinear transformations of this information. The fact that there is still some shared variance between the two measures here may be explained by the construction of the AIS's past state embedding, where the time steps between the samples in the embedding vector are defined as a fractions of the ACT.

Between H and signal variance, we found no significant correlation; moreover, correlations for individual recordings were predominantly negative.

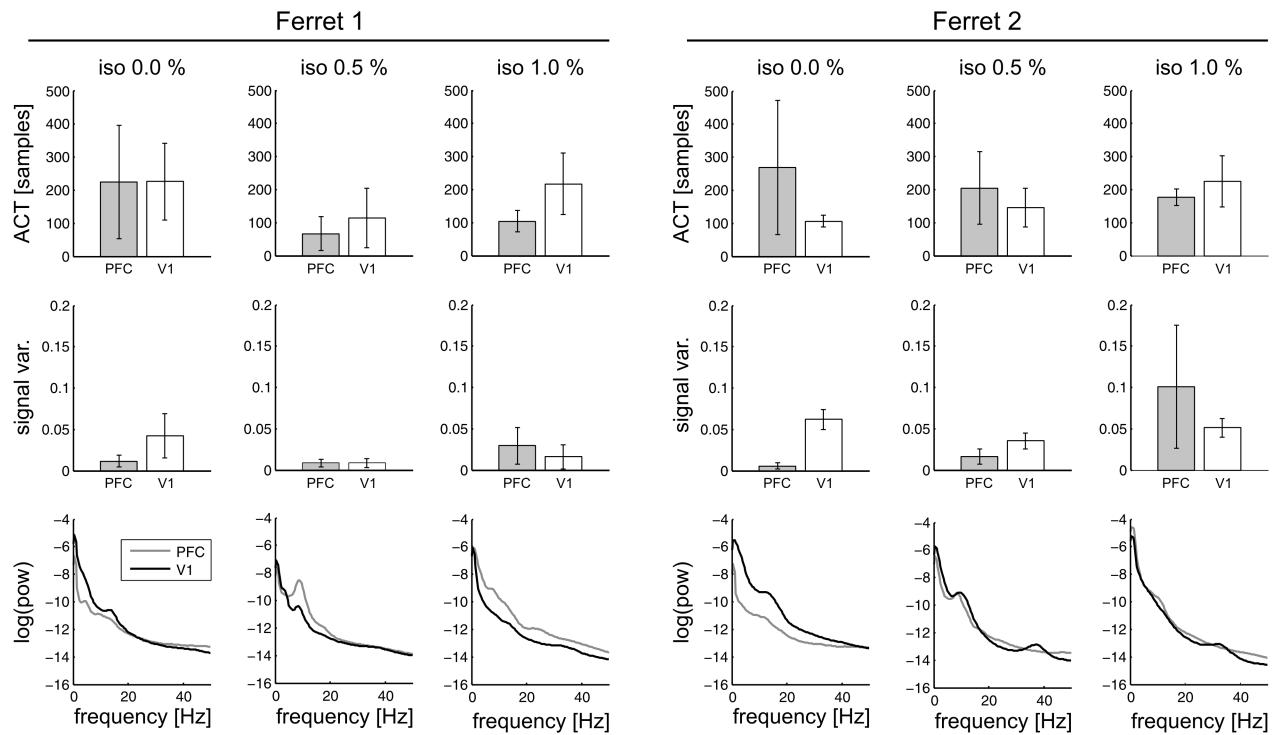


Fig 8. Autocorrelation decay time (ACT), signal variance, and power by isoflurane level and recording site for both animals.
Averages over recordings per isoflurane level and recording site; error bars indicate one standard deviation.

<https://doi.org/10.1371/journal.pcbi.1005511.g008>

The TE_{SPO} was significantly correlated with source ACT for an isoflurane level of 1.0%. Also for the correlation of TE_{SPO} and the source's variance, we found significant correlations for the bottom-up direction under 0.5% and 1.0% isoflurane in animal 1, and for both directions under 1.0% isoflurane in animal 2. For all significant correlations we found $\tilde{R}^2 \leq 0.015$, again indicating a substantial amount of variance in TE_{SPO} that was not explained by ACT or signal variance.

When correlating band power with information-theoretic measures, we found significant correlations between AIS and all bands in both animals (see supporting information S4 Table, $\tilde{R}^2 < 0.16$); for H and band power, we found correlations predominantly in higher frequency bands (beta and gamma, see supporting information S5 Table, $\tilde{R}^2 < 0.1$); for TE_{SPO} and band power, we found significant correlations predominantly in the gamma band and one significant correlation in the delta band for animal 2 and in the beta band for animal 1, respectively (see supporting information S6 Table, $\tilde{R}^2 < 0.024$). In sum, we found relationships between the power in individual frequency bands and all three information-theoretic measures. For all measures, the variance explained was below 0.2, indicating again that band power did not fully capture the properties measured by TE_{SPO} , AIS , and H .

Discussion

We analyzed long-range information transfer between areas V1 and PFC in two ferrets under different levels of isoflurane. We found that transfer entropy was indeed reduced under isoflurane and that this reduction was more pronounced in top-down directions. These results validate earlier findings made using different estimation procedures [3, 6–9]. As far as information transfer alone was concerned our results are compatible with an interpretation of reduced long-range information transfer due to reduced inter-areal coupling. Yet, this

interpretation provides no direct explanation for the findings of reduced locally available information as explained below. In contrast, the alternative hypothesis that the reduced long range information transfer is a secondary effect of changes in local information processing provides a concise explanation for our findings both with respect to locally available information, and information transfer.

Reduction in transfer entropy may be caused by changes in local information processing

To test our alternative hypothesis we evaluated two of its predictions about changes in locally available information, as measured by signal entropy, under administration of different isoflurane concentrations. First, entropy should be reduced; second, the strongest decrease in information transfer should originate from the source node with the strongest decrease in entropy, rather than end in this node. Indeed, we found that signal entropy decreased; the most pronounced decrease in signal entropy was found in PFC. In accordance with our prediction, we found that PFC—the node with the larger decrease in entropy—was also at the source, not the target, of the most pronounced decrease in transfer entropy. This is strong evidence against the theoretical possibility that in a recording site, entropy decreased due to a reduced influx of information—because in this latter scenario the strongest reduction in entropy should have been found at the target (end point) of the most pronounced decrease in information transfer. Hence, our hypothesis that long-range cortico-cortical information transfer is reduced because of changes in local processing must be taken as a serious alternative to the currently prevailing theories of anesthetic action based on disruptions of long range interactions. We suggest that a renewed focus on local information processing in anesthesia research will be pivotal to advance our understanding of how consciousness is lost.

Our predictions that reductions in entropy should potentially be reflected in reduced information transfer derives from the simple principle that information that is not available at the source cannot be transferred. We may thus in principle reduce TE_{SPO} to arbitrarily low values by reducing the entropy of one of the involved processes, without changing the physical coupling between the two systems, just by changing their internal information processing. This trivial but important fact has been neglected in previous studies when interpreting changes in information transfer as changes in coupling strength. Even when this bound is not attained, e.g. because only a certain fraction of the local information is transferred even under 0.0% isoflurane, it seems highly plausible that reductions in the locally available information affect the amount of information transferred. This line of reasoning generalizes to applications beyond anesthesia research, i.e., every application that observes changes in information transfer between two or more experimental conditions should consider changes in local information processing as a potential alternative cause.

A possible indication of how exactly local information processing is changed is given by the observation of increased active information storage in PFC and V1 (also see the next paragraph). This means that more ‘old’ information is kept stored in a source’s activity under anesthesia, rather than being dynamically generated. Such stored source information will not contribute to a measurable transfer entropy under most circumstances because it is already known at the target (see [29], section 5.2.3, for an illustrative example).

Relation between changed locally available information and information storage

In general, AIS increases if a signal becomes more predictable when knowing its past, but is unpredictable otherwise. This also means that the absolute AIS is upper-bounded by the

system's entropy H (see [Methods, Eq 5](#)). Thus, a decrease in H can in principle lead to a decrease in AIS , i.e., fewer possible system states may lead to a decrease in absolute AIS . However, in the present study, we observed an *increase* in AIS while H decreased—this indicates an increase in predictability that more than compensates for the decrease in locally available information. In other words, the system visited fewer states in total but the next state visited became more predictable from the system's past. Thus, a reduction in entropy and increase in predictability points at highly regular neural activity for higher isoflurane concentrations. Such a behavior in activity is in line with existing electrophysiological findings: under anesthesia signals have been reported to become more uniform, exhibiting repetitive patterns, interrupted by bursting activity (see [5] for a review). For example, Purdon and colleagues observed a reduction of the median frequency and an overall shift towards high-power, low-frequency activity during LOC [30]. In particular, slow-wave oscillatory power was more pronounced during anesthesia-induced LOC. LOC was also accompanied by a significant increase in power and a narrowing of oscillatory bands in the alpha frequency range in their study.

Limitations of the applied estimators and measures

Unfortunately a quantitative comparison between different information theoretic estimates that would directly relate H , AIS and TE_{SPO} is not possible using the continuous estimators applied in this study. Their estimates are not comparable because the bias properties of each estimator differ. For each estimator, the bias depends on the number of points used for estimation as well as on the dimensionality of involved variables—however, the exact functional relationship between these two quantities and the bias is unknown and may differ between estimators. (In our application, the dimensionality is determined mainly by the dimension of the past state vectors, and by how many different state variables enter the computation of a measure).

This lack of comparability makes it impossible to normalize estimates; for example, transfer entropy is often normalized by the conditional entropy of the present target state to compare the fraction of transferred information to the fraction of stored information. We forgo this possibility of comparison for a greater sensitivity and specificity in the detection of changes in the individual information theoretic measures here. New estimators, e.g. Bayesian estimators like the ones tested here, promise more comparable estimates by tightly controlling the biases. Yet, these estimators were not as reliable as expected in our study, displaying a relatively high variance.

A further important point to consider when estimating transfer entropy between recordings from neural sites, is the possibility of third unobserved sources influencing the information transfer. In the present study, third sources (e.g., in the thalamus) may influence the information transfer between PFC and V1—for example, if a third source drove the dynamics in both areas, the areas would become correlated, leading to non-zero estimates of TE_{SPO} . This TE_{SPO} is then attributable to the correlation between source and target, but does not measure an actual information transfer between sources. Hence, information transfer estimated by transfer entropy should in general not be directly equated with a causal connection or causal mechanism existing between the source and target process (see also [12] for a detailed discussion of the difference between transfer entropy and measures of causal interactions).

Potential physiological causes for altered information transfer under anesthesia

We here tested the possibility that changes in local information processing lead to the frequently observed reduction in information transfer between cortical areas under isoflurane

administration, instead of altered long-range coupling. Results on entropies and active information storage suggest that this is a definite possibility from a mathematical point of view.

This is supported by the neurophysiology related to the mode of action of isoflurane, because a dominant influence of altered long-range coupling on TE_{SPO} would mandate that synaptic terminals of the axons mediating long range connectivity should be targets of isoflurane. Such long range connectivity is thought to be dominated by glutamatergic AMPA receptors for inter-areal bottom-up connections, and glutamatergic NMDA receptors for inter-areal top-down connections, building upon findings in [31] and [32] (but see [33] for some evidence of GABAergic long range connectivity). Yet, evidence for isoflurane effects on AMPA and NMDA receptors is sparse to date (table 2 in [14]). In contrast, the receptors most strongly influenced by isoflurane seem to be $GABA_A$ and nicotinic acetylcholine (nAChR) receptors. More specifically, isoflurane potentiates agonist interactions at the former, while inhibiting the latter.

Thus, if one adopts the current state of knowledge on the synapses involved in long-range inter-areal connectivity, evidence speaks against a dominant effect of modulation of effective long-range connections by isoflurane. This, in turn, points at local information processing as a more likely reason for changed transfer entropy under isoflurane anesthesia. This interpretation is perfectly in line with our finding that decreases in source entropy seem to determine the transfer entropy decreases, instead of decreases in transfer entropy determining the target entropies.

Nevertheless, targeted local interventions by electrical or optogenetic activation of projection neurons, combined with the set of information theoretic analyses used here, will most likely be necessary to reach final conclusions on the causal role of local entropy changes in reductions of transferred information.

How may altered long-range information transfer lead to loss of consciousness?

Investigating long-range information transfer under anesthesia is motivated by the question how changed information transfer may cause LOC. To that effect, our findings—a dominant decrease in top-down information transfer under anesthesia, and a decrease in locally available information possibly driving it—may be interpreted in the framework of predictive coding theory [34–36]. Predictive coding proposes that the brain learns about the world by constructing and maintaining an internal model of the world, such that it is able to *predict* future sensory input at lower levels of the cortical hierarchy. Whether predictions match actual future input is then used to further refine the internal model. It is thus assumed that top-down information transfer serves the propagation of predictions [37]. Theories of conscious perception within this predictive coding framework propose that conscious perception is “determined” by the internal prediction (or ‘hypothesis’) that matches the actual input best [35, p. 201]. It may be conversely assumed that the absence of predictions leads to an absence of conscious perception.

In the framework of predictive coding theory two possible mechanisms for LOC can be inferred from our data: (1) the disruption of information transfer, predominantly in top-down direction, may indicate a failure to propagate predictions to hierarchically lower areas; (2) the decrease in locally available information and in entropy rates in PFC may indicate a failure to integrate information in an area central to the generation of a coherent model of the world and the generation of the corresponding predictions. These hypotheses are in line with findings reviewed in [4] and [38], which discuss activity in frontal areas and top-down modulatory activity as important to conscious perception.

Future research should investigate top-down information transfer more closely; for example, recent work suggests that neural activity in separate frequency bands may be responsible for the propagation of predictions and prediction errors respectively [37, 39]. Future experiments may target information transfer within a specific band to test if the disruption of top-down information transfer happens in the frequency band responsible for the propagation of predictions.

Last, it should be kept in mind that different anesthetics may lead to loss of consciousness by vastly different mechanisms. Ketamine, for example, seems to increase, rather than decrease, overall information transfer—at least in sub-anesthetic doses [40].

Comparison of the alternative approaches to the estimation of information theoretic measures and to statistical testing

The main analysis of this study was based on information theoretic estimators relying on distances between neighboring data points and on a permutation ANOVA. As each of these has its weaknesses we used additional alternative approaches, first, a Bayesian estimator for information theoretic measures, and second, statistical testing of non-aggregated data using LMMs. Both approaches returned results qualitatively very similar to those of the main analysis. Specifically, replacing neighbor-distance based estimators with Bayesian variants, we replicated the main finding of our study—a reduction in information transfer and locally available information, and an increase in information storage under anesthesia. However, using the Bayesian estimators we did not find a predominant reduction of top-down compared to bottom up information transfer. In contrast, using LMMs for statistical testing instead of a pANOVA additionally revealed a more pronounced reduction in top-down information transfer also in ferret 2 (in this animal, this effect was not significant when performing permutation tests based on the aggregated data).

The Bayesian estimators performed slightly worse than the next neighbor-based estimators in terms of their higher variance in estimates across recordings. Thus, even though Bayesian estimators are currently the best available estimators for discrete data, they may be a non-optimal choice for continuous data. A potential reason for this is the destruction of information on neighborhood relationships through data binning. This is, however, necessary to make the current Bayesian estimators applicable to continuous data.

In sum, we obtained similar results through three different approaches. This makes us confident that locally available information and information transfer indeed decrease under anesthesia, while the amount of predictable information increases.

On information theoretic measures obtained from continuous time processes via time-discrete sampling

When interpreting the information theoretic measures presented in this work, it must be kept in mind that they were obtained via time-discrete sampling of processes that unfold in continuous time (such as the LFP, or spike trains). This time-discrete sampling was taken into account in recent work by Spinney and colleagues [41], who could show that the classic transfer entropy as defined by Schreiber is indeed ambiguous when applied to sampled data from time-continuous processes—it can either be seen as the integral of a transfer entropy rate in continuous time over one sampling interval, and is then given in bits, or it can be seen as an approximation to the continuous time transfer entropy rate itself, and be given in bits/sec. In our study we stick with the first notion of transfer entropy, and note that our results will numerically change when using different sampling intervals. In contrast to the case of transfer entropy covered in [41], corresponding analytical results for AIS are not known at present, as

this is still a field of active research. To nevertheless elucidate the practical dependency of AIS on the sampling rate, when using our estimators, we have taken the original data and have up- and down-sampled them. Indeed the empirical AIS depended on the sampling rate (supporting information S1 Fig). Yet, all relations of AIS values between different isoflurane concentrations, i.e., the qualitative results, are independent of sampling. Thus, sampling effects do not affect the conclusions of the current study.

A cautionary note on the interpretation of information theoretic measures evaluated on local field potential data

When interpreting the results obtained in this study it should be kept in mind that LFP signals are not in themselves the immediate carriers of information relevant to individual neurons. This is because from a neuron's perspective information arrives predominantly in the form postsynaptic potentials generated by incoming spikes and chemical transmission in the synaptic cleft (but see [42] for a potential influence of LFPs on neural dynamics and computation). Thus, LFP signals merely reflect a coarse grained view of the underlying neural information processing. As a consequence, our results only hold in as far as at least some relevant information about the underlying information processing survives this coarse graining in the recording process, and little formal mathematical work has been carried out to estimate bounds on the amount of information available after coarse graining (but see [43]).

Yet, the enormous success that brain reading approaches had when based on local field potentials or on even more coarse grained magnetoencephalography (MEG) recordings (e.g. [44]) indicates that relevant information on neural information processing is indeed available at the level of these signals. However, successful attempts at decoding neural representations of stimuli or other features of the experimental setting should not lead us to misinterpret the information captured by information theoretic measures of neural processing as necessarily *being about something we can understand and link to the outside world*. Quite to the contrary, the larger part of information captured by these measures may be related to intrinsic properties of the unfolding neural computation.

Conclusion

Using two different methods for transfer entropy estimation, and two different statistical approaches, we found that locally available information and information transfer are reduced under isoflurane administration. The larger decrease in the locally available information was found at the source of the larger decrease of information transfer, not at its end point, or target. Therefore, previously reported reductions in information transfer under anesthesia may be caused by changes in local information processing rather than a disruption of long range connectivity. We suggest to put this hypothesis more into the focus of future research effort to understand the loss of consciousness under anesthesia. This suggestion receives further support from the fact that the synaptic targets of the anesthetic isoflurane, as used in this study, are most likely located in local circuits.

Methods

Electrophysiological recordings

We conducted simultaneous electrophysiological recordings of the local field potential (LFP) in primary visual cortex (V1) and prefrontal cortex (PFC) of two female ferrets (17 to 20 weeks of age at study onset) under different levels of isoflurane (Fig 9). Raw data are available from [45]. The choice of the animal model is discussed further in [46]. Recordings were made in a

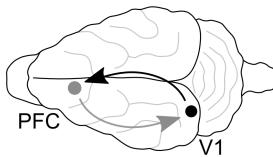


Fig 9. Recording sites in the ferret brain. Prefrontal cortex (PFC, gray dot): anterior sigmoid gyrus; primary visual area (V1, black dot): lateral gyrus. Arrows indicate analyzed directions of information transfer (gray: top-down; black: bottom-up).

<https://doi.org/10.1371/journal.pcbi.1005511.g009>

dark environment during multiple, individual sessions of max. 2 h length, during which the animals' heads were fixed. For recordings, we used single metal electrodes acutely inserted in putative layer IV, measured 0.3–0.6 mm from the surface of cortex (tungsten micro-electrode, 250- μ m shank diameter, 500-k Ω impedance, FHC, Bowdoin, ME). The hardware high pass filter was 0.1 Hz and the low pass filter was 5,000 Hz. A silver chloride wire placed between the skull and soft tissue was used as the reference electrode. The reference electrode was located approximately equidistant between the recording electrodes. This location was selected in order to have little shared activity with either recording electrode. The same reference was used for both recording locations; also the same electrode position was used for both animals and all isoflurane concentrations. To verify that electrode placement was indeed in V1, we mapped receptive fields by eliciting visually evoked potentials in a separate series of experiments. We confirmed electrode placement in PFC by lesioning through the recording electrode after completion of data collection and post-mortem histology (as described in [46]). Details on surgical procedures can be found in [47]. Unfiltered signals were amplified with gain 1,000 (model 1800, A-M Systems, Carlsborg, WA), digitized at 20 kHz (Power 1401, Cambridge Electronic Design, Cambridge, UK), and digitally stored using Spike2 software (Cambridge Electronic Design). For analysis, data were low pass filtered (300 Hz cutoff) and down-sampled to 1000 Hz.

All procedures were approved by the University of North Carolina-Chapel Hill Institutional Animal Care and Use Committee (UNC-CH IACUC) and exceed guidelines set forth by the National Institutes of Health and U.S. Department of Agriculture.

LFPs were recorded during wakefulness (condition *iso* 0.0%, number of recording sessions: 8 and 5 for ferret 1 and 2, respectively) and with different concentrations of anesthetic: 0.5% isoflurane with xylazine (condition *iso* 0.5%, number of sessions: 5 and 6), as well as 1.0% isoflurane with xylazine (condition *iso* 1.0%, number of sessions: 10 and 11). In the course of pilot experiments, both concentrations *iso* 0.5% and *iso* 1.0% lead to a loss of the righting reflex; however, a systematic assessment of this metric during recordings was technically not feasible. Additionally, animals were administered 4.25 ml/h 5% dextrose lactated Ringer and 0.015 ml/h xylazine via IV.

LFP recordings from each session were cut into epochs of 4.81 s length to be able to remove segments of data if they were contaminated by artifacts (e.g., due to movement). We chose a relatively short epoch length to avoid removing large chunks of data when there was only a short transient artifact. This resulted in 196 to 513 epochs per recording (mean: 428.6) for ferret 1, and 211 to 526 epochs (mean: 472.8) for ferret 2. Epochs with movement artifacts were manually rejected (determined by extreme values in the LFP raw traces). In the *iso* 0.0% condition, infrared videography was used to verify that animals were awake during the whole recording; additionally, *iso* 0.0% epochs with a relative delta power (0.5 to 4.0 Hz) of more than 30% of the total power from 0.5 to 50 Hz were rejected to ensure that only epochs during which the animal was truly awake entered further analysis.

Information theoretic measures

To measure information transfer between recording sites V1 and PFC, we estimated the transfer entropy [10] in both directions of possible interactions, $PFC \rightarrow V1$ and $V1 \rightarrow PFC$. To investigate local information processing within each recording site, we estimated active information storage (AIS) [19] as a measure of predictable information, and we estimated differential entropy (H) [48] as a measure of information available locally. We will now explain the applied measures and estimators in more detail, before we describe how these estimators were applied to data from electrophysiological recordings in the next section. To mathematically formalize the estimation procedure from these data, we assume that neural time series recorded from two systems \mathcal{X} and \mathcal{Y} (e.g. cortical sites) can be treated as collections of realizations x_t and y_t of random variables X_t , Y_t of two random processes $X = \{X_1, \dots, X_t, \dots, X_N\}$ and $Y = \{Y_1, \dots, Y_t, \dots, Y_N\}$. The index t here indicates samples in time, measured in units of the dwell time (inverse sampling rate) of the recording.

Transfer entropy. Transfer entropy [10, 22, 49] is defined as the mutual information between the future of a process Y and the past of a second process X , conditional on the past of Y . Transfer entropy thus quantifies the information we obtain about the future of Y from the past of X , taking into account information from the past of Y . Taking this past of Y into account here removes information redundantly available in the past of both X and Y , and reveals information provided synergistically by them [50]. In this study, we used an improved estimator of transfer entropy presented in [15], which accounts for arbitrary information transfer delays:

$$TE_{SPO}(X \rightarrow Y, t, u) = I(Y_t; \mathbf{X}_{t-u}^{d_X} | \mathbf{Y}_{t-1}^{d_Y}), \quad (1)$$

where I is the conditional mutual information (or the differential conditional mutual information for continuous valued variables) between Y_t and $\mathbf{X}_{t-u}^{d_X}$, conditional on $\mathbf{Y}_{t-1}^{d_Y}$; Y_t is the future value of random process Y , and $\mathbf{X}_{t-u}^{d_X}$, $\mathbf{Y}_{t-1}^{d_Y}$ are the past states of X and Y , respectively. Past states are collections of past random variables

$$\mathbf{Y}_{t-1}^{d_Y} = (Y_{t-1}, Y_{t-1-\tau}, \dots, Y_{t-1-(d_Y-1)\tau}), \quad (2)$$

that form a delay embedding of length d_Y [51], and that render the future of the random process conditionally independent of all variables of the random process that are further back in time than the variables forming the state. Parameters τ and d denote the embedding delay and embedding dimension and can be found through optimization of a local predictor as proposed in [21] (see next section on the estimation of information theoretic measures). Past states constructed in this manner are then maximally informative about the present variable of the target process, Y , which is an important prerequisite for the correct estimation of transfer entropy (see also [15]).

In our estimator (Eq 1), the variable u describes the assumed information transfer delay between the processes X and Y , which accounts for a physical delay $\delta_{X,Y} \geq 1$ [15]. The estimator thus accommodates arbitrary physical delays between processes. The true delay $\delta_{X,Y}$ must be recovered by ‘scanning’ various assumed delays and keeping the delay that maximizes TE_{SPO} [15]:

$$\hat{\delta}_{X,Y} = \arg \max_u (TE_{SPO}(X \rightarrow Y, t, u)). \quad (3)$$

Active information storage. AIS [19] is defined as the (differential) mutual information between the future of a signal and its immediate past state

$$AIS(Y_t) = I(Y_t; \mathbf{Y}_{t-1}^{d_Y}), \quad (4)$$

where Y again is a random process with present value Y_t and past state $\mathbf{Y}_{t-1}^{d_Y}$ (see Eq 2). AIS thus quantifies the amount of predictable information in a process or the information that is currently in use for the next state update [19]. AIS is low in processes that produce little information or are highly unpredictable, e.g., fully stochastic processes, whereas AIS is highest for processes that visit many equi-probable states in a predictable sequence, i.e., without branching. In other words, AIS is high for processes with “rich dynamics” that are predictable from the processes’ past [52]. A reference implementation of AIS can be found in the Java Information Dynamics Toolkit (JIDT) [53]. As for TE_{SPO} estimation, an optimal delay embedding $\mathbf{Y}_{t-1}^{d_Y}$ may be found through optimization of the local predictor proposed in [21].

Note, that AIS is upper bounded by the entropy as:

$$\begin{aligned} AIS(Y) &= I(Y_t; \mathbf{Y}_{t-1}^{d_Y}) \\ &= H(Y_t) - H(Y_t | \mathbf{Y}_{t-1}^{d_Y}). \end{aligned} \quad (5)$$

Differential entropy. The differential entropy H (see for example [48]) expands the classical concept of Shannon’s entropy for discrete variables to continuous variables:

$$H = - \int_{\mathbb{X}_{t-u}^{d_X}} f(\mathbf{X}_{t-u}^{d_X}) \log f(\mathbf{X}_{t-u}^{d_X}) d\mathbf{X}_{t-u}^{d_X}, \quad (6)$$

where $f(Y_t)$ is the probability density function of Y_t over the support \mathbb{Y} . Entropy quantifies the average information contained in a signal. Based on the differential entropy the corresponding measures for mutual and conditional mutual information and, thereby, active information storage and transfer entropy can be defined.

Entropy as an upper bound on information transfer

The transfer entropy from Eq 1 can be rewritten as:

$$\begin{aligned} TE_{SPO}(X \rightarrow Y, t, u) &= I(\mathbf{X}_{t-u}^{d_X} : Y_t | \mathbf{Y}_{t-1}^{d_Y}) \\ &= H(\mathbf{X}_{t-u}^{d_X} | \mathbf{Y}_{t-1}^{d_Y}) - H(\mathbf{X}_{t-u}^{d_X} | \mathbf{Y}_{t-1}^{d_Y}, Y_t), \end{aligned} \quad (7)$$

By dropping the negative term on the right hand side we obtain an upper bound (as already noticed by [54, P. 65]), and by realizing that a conditional entropy is always smaller than the corresponding *unconditional* one, we arrive at

$$TE_{SPO}(X \rightarrow Y, t, u) \leq H(\mathbf{X}_{t-u}^{d_X}) \quad (8)$$

This indicates that the overall entropy of the source states is an upper bound. Several interesting other bounds on information transfer exist as detailed in [54], yet these are considerably harder to interpret and were not the focus of the current presentation.

Estimation of information theoretic measures

In this section we will describe how the information theoretic measures presented in the last section may be estimated from neural data. In doing so, we will also describe the methodological pitfalls mentioned in the introduction in more detail and we will describe how these were

handled here. If not stated otherwise, we used implementations of all presented methods in the open source toolboxes TRENTOOL [23] and JIDT [53], called through custom MATLAB® scripts (MATLAB 8.0, The MathWorks® Inc., Natick, MA, 2012). Time series were normalized to zero mean and unit variance before estimation.

Estimating information theoretic measures from continuous data. Estimation of information theoretic measures from continuous data is often handled by simply discretizing the data. This is done either by binning or the use of symbolic time series—mapping the continuous data onto a finite alphabet. Specifically, the use of symbolic time series for transfer entropy estimation was first introduced by [55] and maps the continuous values in past state vectors with length d (Eq 2) onto a set of rank vectors. Hence, the continuous-valued time series is mapped onto an alphabet of finite size $d!$. After binning or transformation to rank vectors transfer entropy and the other information theoretic measures can then be estimated using plug-in estimators for discrete data, which simply evaluate the relative frequency of occurrences of symbols in the alphabet. Discretizing the data therefore greatly simplifies the estimation of transfer entropy from neural data, and may even be necessary for very small data sets. Yet, binning ignores the neighborhood relations in the continuous data and the use of symbolic times series destroys important information on the absolute values in the data. An example where transfer entropy estimation fails due to the use of symbolic time series is reported in [56] and discussed in [15]: In this example, information transfer between two coupled logistic maps was not detected by symbolic transfer entropy [56]; only when estimating TE_{SPO} directly using an estimator for continuous data, the information transfer was identified correctly [15]. To circumvent the problems with binned or symbolic time series, we here used a nearest-neighbor based TE_{SPO} -estimator for continuous data, the Kraskov-Stögbauer-Grassberger (KSG) estimator for mutual information described in [57]. At present, this estimator has the most favorable bias properties compared to similar estimators for continuous data. The KSG-estimator leads to the following expression for the estimation of TE_{SPO} as introduced in Eq 1 [22, 58]:

$$\begin{aligned} TE_{SPO}(X \rightarrow Y, t, u) &= I(Y_t : \mathbf{X}_{t-u}^{d_X} | \mathbf{Y}_{t-1}^{d_Y}) \\ &= \psi(k) + \langle \psi(n_{y_{t-1}} + 1) - \psi(n_{y_t y_{t-1}^{d_Y}} + 1) - \psi(n_{y_{t-1} x_{t-u}^{d_X}} + 1) \rangle_r, \end{aligned} \quad (9)$$

where ψ denotes the digamma function, k is the number of neighbors in the highest-dimensional space spanned by variables Y_t , $\mathbf{Y}_{t-1}^{d_Y}$, $\mathbf{X}_{t-u}^{d_X}$, and is used to determine search radii for the lower dimensional subspaces; n_{\cdot} are the number of neighbors within these search radii for each point in the lower dimensional search spaces spanned by the variable indicated in the subscript. Angle brackets indicate the average over realizations r (e.g. observations made over an ensemble of copies of the systems or observations made over time in case of stationarity, which we assumed here). We used a k of 4 as recommended by Kraskov [59, p. 23], such as to balance the estimator's bias—which decreases for larger k —and variance—which increases for larger k (see also [60] for similar recommendations based on simulation studies). For a detailed derivation of TE_{SPO} -estimation using the KSG-estimator see [22, 57, 58].

The KSG-estimator comes with a bias that is not analytically tractable [59], hence, estimates can not be interpreted at face value, but have to be tested for their statistical significance against the null-hypothesis of no information transfer [22, 23]. A suitable test distribution for this null-hypothesis can be generated by repeatedly estimating TE_{SPO} from surrogate data. In these surrogate data, the quantity of interest—potential transfer entropy between source and target time series—should be destroyed while all other statistical properties of the data are preserved such that the estimation bias is constant between original and surrogate data. In the present study we thus generated 500 surrogate data sets by permuting whole epochs of the

target time series while leaving the order of source time course epochs unchanged (see [61] for a detailed account of surrogate testing).

We used the ensemble-method for transfer entropy estimation [61], implemented in TRENTOOL [23]. The ensemble method allows to pool data over epochs, which maximizes the amount of data entering the estimation, while providing an efficient implementation of this estimation procedure using graphics processing units (GPU). We tested the statistical significance of TE_{SPO} in both directions of interaction in the *iso* 0.0% condition. We only tested this condition, because TE_{SPO} was expected to be reduced for higher isoflurane levels based on the results of existing studies. Because the estimation of TE_{SPO} is computationally heavy, we used a random subset of 50 epochs to reduce the running time of this statistical test (see supporting information S1 Text and S7 Table for theoretical and practical running times of the used estimators). We first tested TE_{SPO} estimates for their significance within individual recording sessions against 500 surrogate data sets, and then used a binomial test to establish the statistical significance *over* recordings. We used a one-sided Binomial test under the null hypothesis of no significant TE_{SPO} estimates, where individual estimates l were assumed to be $B(l, p_0, n)$ -distributed, with $p_0 = 0.05$ and $n = 5$ for animal 1 and $p_0 = 0.05$ and $n = 8$ for animal 2.

As the KSG-estimator used for estimating TE_{SPO} (Eq 9) is an estimator of mutual information it can also be used for the estimation of AIS:

$$\begin{aligned} AIS(Y) &= I(Y_t : \mathbf{Y}_{t-1}^{d_Y}) \\ &= \psi(k) - 1/k + \psi(N) - \langle \psi(n_{y_t}) + \psi(n_{\mathbf{y}_{t-1}^{d_Y}}) \rangle_r, \end{aligned} \quad (10)$$

where again ψ denotes the digamma function, k is the number of neighbors in the highest-dimensional space, N is the number of realizations, and n . denotes the number of neighbors for each point in the respective search space. Again, we chose $k = 4$ for the estimation of AIS (see above). Note that the sampling rate has an effect on the estimated absolute values of AIS (a simulation of the effect of sampling on AIS estimates are shown as supporting information S1 Fig)—however, qualitative results are not influenced by the choice of sampling rate; in other words, relative differences between estimates are the same for different choices of sampling rates. As a consequence, for AIS estimation the sampling rate should be constant over data sets if the aim is to compare these estimates.

A conceptual predecessor of the KSG-estimator for mutual information is the Kozachenko-Leonenko (KL) estimator for differential entropies [62]. The KL-estimator also allows for the estimation of H from continuous data and reads

$$H(X) = -\psi(1) + \psi(N) + \sum_{i=1}^N \log(\epsilon(i)), \quad (11)$$

where $\epsilon(i)$ is twice the distance from data point i to its k -th nearest neighbor in the search space spanned by all points.

Bayesian estimators for discretized data. For Bayesian estimation we converted the continuous LFP time series to discrete data by applying voltage bins as follows: The voltages ± 3 standard deviations around the mean of the LFP were subdivided into N equally spaced bins. We added two additional bins containing all the values that were either smaller or larger than the 3 SD region, amounting to a total number of bins $N_{bins} = N + 2$. We then calculated H , AIS and TE_{SPO} for the discrete data. For AIS and TE_{SPO} estimation states were defined using the same dimension d and τ as for the KSG-estimator, optimized using the Ragwitz criterion. We decomposed TE_{SPO} into four entropies (Eq 9), and AIS into three entropies, which we then

estimated individually [63]. To reduce the bias introduced by the limited number of observed states, we used the NSB-estimator by Nemenman, Shafee, and Bialek [17], which is based on the construction of an almost uniform prior over the expected entropy using a mixture of symmetric Dirichlet priors \mathcal{P}_β . The estimator has been shown to be unbiased for a broad class of distributions that are typical in \mathcal{P}_β [18].

We further applied the recently proposed estimator by Archer et al. [20] that uses a prior over distributions with infinite support based on Pitman-Yor-processes. In contrast to the NSB prior, this prior also accounts for heavy-tailed distributions that one encounters frequently in neuronal systems and does not require knowledge of the support of the distribution.

When estimating entropies for the embedding dimensions used here, the number of possible states or “words” is between $K = 12^{14}$ and $K = 12^{31}$. This is much larger than the typical number of observed states per recording of around $N = 5 \cdot 10^5$. As a consequence, a precise estimation of entropies is only possible if the distribution is sparse, i.e. most words have vanishing probability. In this case, however, the estimates should be independent of the choice of support K as long as K is sufficiently large and does not omit states of finite probability. We chose $K' = 10^{12}$ for the results shown in this paper, which allowed a robust computation of the NSB estimator instead of the maximum support K that results from simple combinatorics.

Finding optimal embedding parameters. The second methodological problem raised in the introduction was the choice of embedding parameters for transfer entropy estimation. One important parameter here is the choice of the total signal history when constructing past states for source and target signal (see Eqs 1 and 2). Failure to properly account for signal histories may lead to a variety of errors, such as underestimating transfer entropy, failure to detect transfer entropy altogether, or the detection of spurious transfer entropy. Transfer entropy is underestimated or missed if the past state of the *source* time series does not cover all the relevant history, i.e., the source is under-embedded. In contrast, spurious transfer entropy may be detected if the past state of the *target* time series is under-embedded, such that spurious detection of transfer entropy is a false positive and therefore the most serious error. One scenario where spurious transfer entropy results from under-embedding is shown in Fig 10.

The choice of an optimal embedding is also relevant for the estimation of AIS, where under-embedding leads to underestimation of the true AIS. Note that on the other hand, we can not increase the embedding length to arbitrarily high values because this leads to computationally intractable problem sizes and requires exponentially more data for estimation.

Optimal embedding parameters d and τ may be found through the optimization of a local predictor proposed by Ragwitz [21]. Ragwitz’ criterion tests different combinations of a range of values for d and τ . The current combination is used to embed each point in a time series, then, the future state for each point is predicted from the future states of its neighbors. The parameter combination that leads to the best prediction on average is used as the optimal embedding. To determine the neighbors of a point, we used a k -nearest-neighbor search with $k = 4$, i.e., the same value for k as was used for k -nearest-neighbor searches when estimating information-theoretic measures. We minimized the mean squared error when optimizing Ragwitz’ local predictor. Further details on Ragwitz’ criterion can be found in the documentations of the TRENTOOL [23] and JIDT [53] toolbox.

Other approaches for embedding parameter optimization have been proposed, see for example non-uniform embedding using mutual information to determine all relevant past samples as proposed by [64].

Reconstruction of information transfer delays. The third methodological problem raised in the introduction was failure to account for a physical delay δ between neural sites

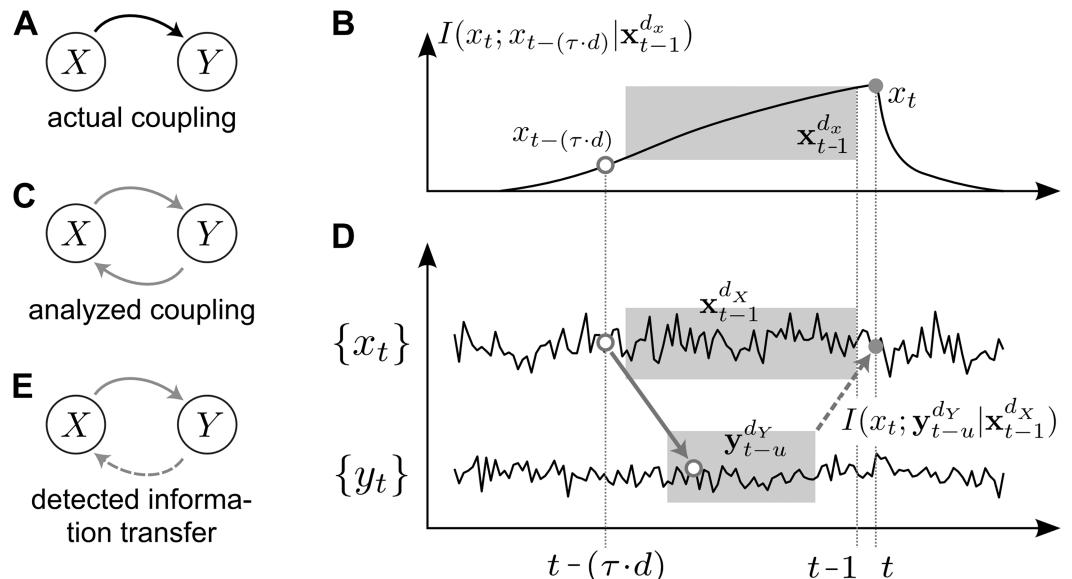


Fig 10. Spurious information transfer resulting from under-embedding (modified from [22]). (A) Actual coupling between processes X and Y . (B) Mutual information between the present value in X , x_t , and a value in the far past of X , $x_{t-(\tau \cdot d)}$, conditional on all intermediate values $x_{t-1}^{d_x}$ (shaded box), the mutual information is non-zero, i.e., $x_{t-(\tau \cdot d)}$ holds some information about x_t . (C) Both directions of interaction are analyzed; (D) Information in $x_{t-(\tau \cdot d)}$ (white sample point) is transferred to Y (solid arrow), because of the actual coupling $X \rightarrow Y$. The information in $x_{t-(\tau \cdot d)}$ about x_t is thus transferred to the past of Y , $y_{t-u}^{d_y}$, which thus becomes predictive of x_t as well. Assume now, we analyzed information transfer from Y to X , $I(x_t : y_{t-u}^{d_y} | x_{t-1}^{d_x})$, without a proper embedding of X , $x_{t-1}^{d_x}$: Because of the actually transferred information from $x_{t-(\tau \cdot d)}$ to $y_{t-u}^{d_y}$, the mutual information $I(x_t : x_{t-1}^{d_x})$ is non-zero. If we now under-embed X , such that the information in $x_{t-(\tau \cdot d)}$ is not contained in $x_{t-1}^{d_x}$ and is not conditioned out, $I(x_t : y_{t-u}^{d_y} | x_{t-1}^{d_x})$ will be non-zero as well. In this case, under-embedding of the target X will lead to the detection of spurious information transfer in the non-coupled direction $Y \rightarrow X$. (E) Information transfer is falsely detected for both directions of interaction, the link from Y to X is spurious (dashed arrow).

<https://doi.org/10.1371/journal.pcbi.1005511.g010>

when estimating transfer entropy. In our estimator TE_{SPO} (Eq 9) we account for δ by introducing the parameter u . The delay u needs to be optimized to correctly estimate TE_{SPO} . If u is not optimal, i.e., u is not sufficiently close to δ (Fig 5 and [15]), information transfer may be underestimated or not measured at all. This is because choosing the parameter u too large ($u \gg \delta$) means that the information present in the evaluated samples of the source is also present in the history of the target already, and conditioned away. In contrast, choosing the parameter u too small means that the information of the evaluated samples of the source will only arrive in the future of the current target sample, and is useless for providing information about it (Fig 5A).

It can be proven for bivariate systems that TE_{SPO} becomes maximal when the true delay δ is chosen for u [15]. Therefore, the true delay δ —and thus an optimal choice for u —can be found by using the value for u that maximizes TE_{SPO} [15]. This optimal u can be found by scanning a range of assumed values. In the present study, we scanned values ranging from 0 to 20 ms. (Note that assumed values should be physiologically plausible to keep the computations practically feasible.)

Accounting for the information transfer δ by finding optimal parameters u for TE_{SPO} estimation has important consequences when calculating indices from estimated TE_{SPO} , such as TE_{net} :

$$TE_{net} = \frac{TE_{SPO}(X \rightarrow Y, t, u_{X \rightarrow Y}) - TE_{SPO}(Y \rightarrow X, t, u_{Y \rightarrow X})}{TE_{SPO}(X \rightarrow Y, t, u_{X \rightarrow Y}) + TE_{SPO}(Y \rightarrow X, t, u_{Y \rightarrow X})}, \quad (12)$$

or variations of this measure. The TE_{net} is popular in anesthesia research [7, 9] and indicates the predominant direction of information transfer between two bidirectionally coupled processes X and Y ($TE_{net} \geq 0$ if $TE_{SPO}(X \rightarrow Y, t, u_X \rightarrow Y) \geq TE_{SPO}(Y \rightarrow X, t, u_Y \rightarrow X)$ and $TE_{net} < 0$ if $TE_{SPO}(Y \rightarrow X, t, u_Y \rightarrow X) > TE_{SPO}(X \rightarrow Y, t, u_X \rightarrow Y)$). However, if values for $u_X \rightarrow Y$ and $u_Y \rightarrow X$ are not optimized individually, TE_{net} may take on arbitrary signs: In Fig 5B, we show a toy example of two coupled Lorenz systems [15], where the absolute difference between raw TE_{SPO} values changes as a function of a common u for both directions and where the difference even changes signs for values $u > 65$. To obtain a meaningful value from TE_{net} we thus need to find the individually optimal choices of u for both directions of transfer—in the example, these optima are found at $u_X \rightarrow Y = 46$ and $u_Y \rightarrow X = 76$, leading to the “true” difference.

Simulating the effect of filtering on information transfer delay reconstruction. To simulate the effects of filtering as a preprocessing technique on the ability of the TE_{SPO} estimator to reconstruct the correct information transfer delay, we simulated two coupled time series, for which we estimated transfer entropy before filtering and after band-pass filtering with different bandwidths. The simulation was repeated 50 times.

We simulated two time series with 100,000 samples each, which were drawn from a uniform random distribution over the open interval $(0, 1)$. We introduced a coupling between the time series by adding a scaled version (factor 0.2) of the first time series to the second with a delay of 10 samples. We estimated TE_{SPO} from the first to the second time series using the KSG-estimator implemented in the JIDT toolbox, with $k = 4$ and a history of one sample for both source and target.

We estimated TE_{SPO} with and without band-pass filtering the data for different values of u , ranging from 1 to 20 samples. We filtered the data using a fourth order, causal Butterworth filter, implemented in the MATLAB toolbox FieldTrip [26]. We filtered the data using four different bandwidths: from 0.1 to 300 Hz (corresponding to the filtering done in this study), from 0.1 to 200 Hz, from 12 to 30 Hz (corresponding to the beta frequency range), and from 4 to 8 Hz (corresponding to the theta frequency range).

Statistical testing using permutation testing

To test for statistically relevant effects of isoflurane levels and direction of interaction or recording site on estimated measures, we performed two-factorial permutation analyses of variance (pANOVA) for each animal and estimated measure [39, 65–67]. We used a MATLAB® implementation of the test described in [67], which is compatible with the FieldTrip toolbox data format [26] and is available from [68].

The permutation ANOVA can be used if the normality assumptions of parametric ANOVA are violated or—as in the present study—if assumptions are not testable due to too few data points per factor level. The permutation ANOVA evaluates the significance of the main effect of individual factors or their interaction effect under the null-hypothesis of no experimental effect at all. The significance is evaluated by calculating a F-ratio for the effect from the original data; this original F-ratio is then compared against a distribution of F-ratios obtained from permuted data. The F-ratio’s p-value is calculated as the fraction of ratios obtained from permuted data that is bigger than the original F-ratio. When permuting data, it is crucial to only permute data in such a way that the currently investigated effect is destroyed while all other effects are kept intact [65]: for example, consider a two-factorial design with factors A and B —if the main effect of factor A is to be tested, the assignment of levels of A to data points has to be permuted; yet, the permutation of levels of A can only happen *within* levels of factor B such as to not simultaneously destroy the effect of factor B . Thus, when testing for the effect of one factor, the effect of all other factors is preserved, making sure that variability due to one factor is tested for while

the variability due to the other factor is held constant. However, this permutation scheme is not applicable to the interaction effect because it leaves no possible permutations—destroying the interaction effect through permutation always also destroys the main effects of individual factors. Here, both factors have to be permuted, which yields an approximative test of the interaction effect (see [67] for a discussion of permutation strategies and simulations).

The factors in the permutation ANOVA were *isoflurane level* and *direction* for TE_{SPO} estimates, and *isoflurane level* and *recording site* for AIS and H estimates. The number of permutations was set to 10,000. In the present study, data was recorded in different sessions and segmented into epochs. As a result, estimates for individual epochs should not be pooled over recordings for statistical analysis (see for example [69], and also the next paragraph for a discussion). We therefore aggregated estimates over epochs to obtain one estimate per recording session. We used the median to aggregate the estimated values for each recording session over individual epochs, because the distribution of measures over epochs was skewed and we considered the median a more exact representation of the distributions' central tendencies. The aggregation of data resulted in relatively few observations per ANOVA cell and also in unequal number of observations between cells, thus violating two basic assumptions of parametric ANOVA (too few observations make it impossible to test for parametric assumptions like homogeneity of variances). Therefore, we used the non-parametric permutation approach over a parametric one, because the permutation ANOVA does not make any assumptions on data structure.

Statistical testing using linear mixed models

As described above, LFP recordings were conducted in epochs over multiple recording sessions. This introduces a so-called “nested design” [69], i.e., a hierarchical structure in the data, where data epochs are nested within recordings, which are nested in animals. Such structures lead to systematic errors or dependence within the data. This violates the assumption of uncorrelated errors made by the most common tests derived from the general linear model and leads to an inflation of the type I error [69]. A measure of the degree of dependence is the intraclass correlation coefficient (ICC), which was 0.35 for ferret 1 and 0.19 for ferret 2, indicating a significant dependency within the data (see [69] for a discussion of this measure). Thus, we performed an additional statistical tests where we again tested for a significant effect of the two factors *isoflurane level* and *direction* as well as their interaction, but we additionally modeled the nested structure in our data by a random factor *recordings*. Such a model is called a linear mixed effects model [70], and may yield higher statistical power than aggregating data within one level of the nested design (as was done for the permutation ANOVA). We used the following model for both animals separately:

$$y_{ij} = \beta_0 + \gamma_j + \beta_1 D_i + \beta_2 A_i^{(0.5)} + \beta_3 A_i^{(1.0)} + \epsilon_{ij},$$

where y_{ij} is the TE_{SPO} value from the i -th epoch in recording j , modeled as a function of isoflurane level and direction of interaction, where β_0 describes the model intercept, β_k are the regression coefficients and describe fixed effects, γ_j is the random deviation of recording j from the intercept β_0 , and ϵ_{ij} describes random noise. D_i , $A_i^{(0.5)}$, and $A_i^{(1.0)}$ are predictor variables, encoding factors *direction* as

$$D_i := \begin{cases} 1 & \text{if } \textit{direction} \text{ is PFC} \rightarrow V1, \\ -1 & \text{if } \textit{direction} \text{ is V1} \rightarrow PFC, \end{cases} \quad (13)$$

and factor *isoflurane level* as

$$A_i^{(0.5)} := \begin{cases} 1 & \text{if } \textit{isoflurane level} \text{ is } \textit{iso } 0.5\%, \\ 0 & \text{else,} \end{cases} \quad (14)$$

$$A_i^{(1.0)} := \begin{cases} 1 & \text{if } \textit{isoflurane level} \text{ is } \textit{iso } 1.0\%, \\ 0 & \text{else.} \end{cases} \quad (15)$$

Note that we used dummy coding for factor *direction* so that the estimated effect β_1 can be interpreted like the simple or main effect in a standard ANOVA framework. We used contrast coding for factor *isoflurane level* which allows to interpret estimated effects β_2 and β_3 as deviations from a reference group (in this case the condition *iso 0.0%*). We further assume that noise was i.i.d. and $\epsilon_{ij} \sim \mathcal{N}(0, \sigma_e^2)$ and $\gamma_j \sim \mathcal{N}(0, \sigma_\gamma^2)$.

We used the R language [71] and the function `lmer` from the `lme4`-package [72] for model fitting. We assessed statistical significance of individual factors by means of model comparison using the maximum likelihood ratio between models [73]. To allow for this model comparison, we used maximum likelihood estimation, instead of restricted maximum likelihood estimation, of random and fixed effects. To test for main effects, we compared models including individual factors *isoflurane level* (*fm_a*) and *direction* (*fm_d*) to a Null model (*fm_0*) including only the random effect; to furthermore test for an interaction effect, we compared the model including an interaction term (*fm_axb*) to a model where both factors only entered additively (*fm_ab*).

The models were fitted to 15,973 TE_{SPO} values from ferret 1 and 18,202 TE_{SPO} values from ferret 2, respectively.

Simulating the effect of reduced source entropy on transfer entropy

To test whether changes in source entropy influenced the transfer entropy despite unchanged coupling, we simulated two test cases, with high and low source entropy, respectively, while the coupling between source and target process were held constant. To simulate the two test cases, we randomly selected two recordings—one from the *iso 0.0%* condition, which on average showed higher source entropy, and one from the *iso 1.0%* condition, which on average showed lower source entropy (Figs 1 and 2). In the recording from the *iso 0.0%* condition, we permuted epochs in the target time series to destroy all information transfer present in the original data. From the permuted data, we simulated two cases of artificial coupling, first, using the high-entropy source time course from the *iso 0.0%* condition; and second, using the low-entropy source time course from the *iso 1.0%* condition. The coupling was simulated by adding a filtered, scaled and delayed version of the respective source time course to the target time course for each epoch. For filtering, we used a Gaussian filter with a smoothing of 10 samples; for the scaling factor, we used a value of 0.2, which resulted in a TE_{SPO} value for the high-entropy test case that was close to the TE_{SPO} in the *iso 0.0%* condition. By replacing the original coupling in both recordings with a simulated coupling, we made sure that the coupling was constant for both test cases.

For both test cases, we estimated TE_{SPO} and $H(\mathbf{X}_{t-u}^{dx})$ following the estimation procedures described above. We tested differences in TE_{SPO} using a permutation independent samples t-test with 10,000 permutations. To make sure that our simulation reflected information transfer found in the original data, we further tested for a significant difference between TE_{SPO} in the original data from the *iso 0.0%* recording and TE_{SPO} in the high-entropy test case (using the source time *iso 0.0%* condition). Here, we found no significant difference, indicating that

the transfer entropy in our test case did not differ significantly from the transfer entropy found in the original data.

Correlating information-theoretic measures with other time-series properties

We correlated TE_{SPO} , AIS, and H with more conventional measures from time-series analysis, namely, the autocorrelation decay time (ACT), signal variance, and power in individual frequency bands. This correlation was performed to investigate whether information-theoretic measures captured signal properties that could be described equally well by more simple measures.

The ACT was calculated by finding the lag at which the autocorrelation coefficient decayed below e^{-1} . The signal variance was calculated as the variance over time for each recording after subtracting the mean of the time series. The band power was calculated for individual frequency bands (delta = 0.5–4 Hz, theta = 4–8 Hz, alpha = 8–12 Hz, beta = 12–30 Hz, gamma = 30–40 Hz, following [47]), using the multitaper method with discrete prolate spheroidal sequences (Slepian sequences) as tapers (smoothing = 1 Hz) implemented in the MATLAB toolbox FieldTrip [26].

We calculated correlations between information-theoretic measures and conventional measures for individual isoflurane levels and recording sites or direction of interaction, respectively. For each correlation, we pooled data over recording sessions for the respective isoflurane level and calculated Spearman's rank correlation. We tested the correlation for significance using a restricted permutation test, where permutations were allowed only *within* one recording, accounting for the nested experimental design (see [69] and section *Statistical testing using linear mixed models*, below). Additionally, we calculated the correlation as well as the variance explained, R^2 , for individual recordings, because calculating R^2 for the coefficient calculated from *pooled* data does not yield interpretable results.

Supporting information

S1 Dataset. Group statistical data and estimated information-theoretic values entering statistical tests.

(ZIP)

S1 Table. Correlation of AIS with autocorrelation decay time (ACT). \tilde{R}^2 and $\max(R^2)$ indicate the median and maximum of R^2 over recordings per condition, respectively.

(PDF)

S2 Table. Correlation of H with signal variance. \tilde{R}^2 and $\max(R^2)$ indicate the median and maximum of R^2 over recordings per condition, respectively.

(PDF)

S3 Table. Correlation of TE_{SPO} with source autocorrelation decay time (ACT) and source signal variance. \tilde{R}^2 and $\max(R^2)$ indicate the median and maximum of R^2 over recordings per condition, respectively.

(PDF)

S4 Table. Correlation of AIS with band power. \tilde{R}^2 and $\max(R^2)$ indicate the median and maximum of R^2 over recordings per condition, respectively.

(PDF)

S5 Table. Correlation of H with band power. \tilde{R}^2 and $\max(R^2)$ indicate the median and maximum of R^2 over recordings per condition, respectively.
(PDF)

S6 Table. Correlation of TE_{SPO} with band power. \tilde{R}^2 and $\max(R^2)$ indicate the median and maximum of R^2 over recordings per condition, respectively.
(PDF)

S7 Table. Average running times. Practical average running times for estimation of information-theoretic measures from one recording session and two recording sites or directions of interaction in animal 1 (see also supporting information [S1 Text](#)).
(PDF)

S1 Fig. The effect of sampling on the estimation of active information storage (AIS). AIS estimates from data sampled at different rates. Data were chosen from three random recordings in animal 1 under three levels of Isoflurane. AIS was estimated from data with the sampling rate used for analysis in the present work (1000 Hz) and from data re-sampled at 2000 and 600 Hz respectively; note that qualitative results did not change due to re-sampling, but absolute estimates increased for higher sampling rates; the number of data points was held approximately constant by selecting a subset of trials for estimation such that the number of points entering the analysis was equal to the smallest number of points over all isoflurane levels.
(TIF)

S1 Text. Theoretical and practical running time of nearest-neighbor based estimators.
(PDF)

Acknowledgments

This study was partly inspired by the work of Vasily A. Vakorin on the interplay of complexity and information transfer in [74]. The authors thank Saskia Helbling for fruitful discussions on data analysis and for making the permutation ANOVA code available.

Author Contributions

Conceptualization: MW VP AH.

Data curation: KKS PW AH MW.

Formal analysis: PW LR.

Funding acquisition: FF MW VP.

Investigation: KKS FF.

Methodology: PW MW LR VP.

Project administration: MW FF AH.

Resources: MW VP FF.

Software: PW LR MW VP.

Supervision: MW.

Validation: AH.

Visualization: PW LR.

Writing – original draft: PW MW LR.

Writing – review & editing: MW PW VP AH KKS FF LR.

References

1. Dehaene S, Changeux JP. Experimental and theoretical approaches to conscious processing. *Neuron*. 2011; 70(2):200–227. <https://doi.org/10.1016/j.neuron.2011.03.018> PMID: 21521609
2. Tononi G. An information integration theory of consciousness. *BMC Neurosci*. 2004; 5:42. <https://doi.org/10.1186/1471-2202-5-42> PMID: 15522121
3. Imas OA, Ropella KM, Ward BD, Wood JD, Hudetz AG. Volatile anesthetics disrupt frontal-posterior recurrent information transfer at gamma frequencies in rat. *Neurosci Lett*. 2005; 387(3):145–150. <https://doi.org/10.1016/j.neulet.2005.06.018> PMID: 16019145
4. Hudetz AG. Suppressing consciousness: mechanisms of general anesthesia. *Semin Anesth*. 2006; 25(4):196–204. <https://doi.org/10.1053/j.sane.2006.09.003>
5. Alkire MT, Hudetz AG, Tononi G. Consciousness and anesthesia. *Science*. 2008; 322(5903):876–880. <https://doi.org/10.1126/science.1149213> PMID: 18988836
6. Ku SW, Lee U, Noh GJ, Jun IG, Mashour GA. Preferential inhibition of frontal-to-parietal feedback connectivity is a neurophysiologic correlate of general anesthesia in surgical patients. *PLoS One*. 2011; 6(10):e25155. <https://doi.org/10.1371/journal.pone.0025155> PMID: 21998638
7. 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(6):1264–1275. <https://doi.org/10.1097/ALN.0b013e31829103f5> PMID: 23695090
8. Jordan D, Ilg R, Riedl V, Schorer A, Grimberg S, Neufang S, et al. Simultaneous electroencephalographic and functional magnetic resonance imaging indicate impaired cortical top-down processing in association with anesthetic-induced unconsciousness. *Anesthesiology*. 2013; 119(5):1031–1042. <https://doi.org/10.1097/ALN.0b013e3182a7ca92> PMID: 23969561
9. Untergerher G, Jordan D, Kochs EF, Ilg R, Schneider G. Fronto-parietal connectivity is a non-static phenomenon with characteristic changes during unconsciousness. *PLoS One*. 2014; 9(1):e87498. <https://doi.org/10.1371/journal.pone.0087498> PMID: 24475298
10. Schreiber T. Measuring information transfer. *Phys Rev Lett*. 2000; 85(2):461–464. <https://doi.org/10.1103/PhysRevLett.85.461> PMID: 10991308
11. Ay N, Polani D. Information flows in causal networks. *Adv Complex Syst*. 2008; 11:17–41. <https://doi.org/10.1142/S0219525908001465>
12. Lizier JT, Prokopenko M. Differentiating information transfer and causal effect. *Eur Phys J B*. 2010; 73:605–615. <https://doi.org/10.1140/epjb/e2010-00034-5>
13. Chicharro D, Ledberg A. When two become one: the limits of causality analysis of brain dynamics. *PLoS One*. 2012; 7:e32466. <https://doi.org/10.1371/journal.pone.0032466> PMID: 22438878
14. Krasowski M, Harrison N. General anaesthetic actions on ligand-gated ion channels. *Cell Mol Life Sci*. 1999; 55(10):1278–1303. <https://doi.org/10.1007/s000180050371> PMID: 10487207
15. Wibral M, Pampu N, Priesemann V, Siebenhühner F, Seiwert H, Lindner M, et al. Measuring information-transfer delays. *PLoS One*. 2013; 8(2):e55809. <https://doi.org/10.1371/journal.pone.0055809> PMID: 23468850
16. Wollstadt P, Sellers KK, Hutt A, Fröhlich F, Wibral M. Anesthesia-related changes in information transfer may be caused by reduction in local information generation. In: Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE. IEEE; 2015. p. 4045–4048.
17. Nemenman I, Shafee F, Bialek W. Entropy and Inference, Revisited. In: Advances in Neural Information Processing Systems 14. Cambridge, MA: MIT Press; 2002. p. 471–478.
18. Nemenman I, Shafee F, van Steveninck RR. Entropy and information in neural spike trains: progress on the sampling problem. *Phys Rev E*. 2004; 69:056111. <https://doi.org/10.1103/PhysRevE.69.056111> PMID: 15244887
19. Lizier JT, Prokopenko M, Zomaya AY. Local measures of information storage in complex distributed computation. *Inform Sciences*. 2012; 208:39–54.
20. Archer E, Park IM, Pillow JW. Bayesian entropy estimation for countable discrete distributions. *J Mach Learn Res*. 2014; 15:2833–2868.

21. Ragwitz M, Kantz H. Markov models from data by simple nonlinear time series predictors in delay embedding spaces. *Phys Rev E*. 2002; 65(5):056201. <https://doi.org/10.1103/PhysRevE.65.056201> PMID: 12059674
22. Vicente R, Wibral M, Lindner M, Pipa G. Transfer entropy—a model-free measure of effective connectivity for the neurosciences. *J Comput Neurosci*. 2011; 30(1):45–67. <https://doi.org/10.1007/s10827-010-0262-3> PMID: 20706781
23. Lindner M, Vicente R, Priesemann V, Wibral M. TRENTOOL: a Matlab open source toolbox to analyse information flow in time series data with transfer entropy. *BMC Neurosci*. 2011; 12(1):119. <https://doi.org/10.1186/1471-2202-12-119> PMID: 22098775
24. Barnett L, Seth AK. Behaviour of Granger causality under filtering: Theoretical invariance and practical application. *J Neurosci Methods*. 2011; 201(2):404–419. <https://doi.org/10.1016/j.jneumeth.2011.08.010> PMID: 21864571
25. Florin E, Gross J, Pfeifer J, Fink GR, Timmermann L. The effect of filtering on Granger causality based multivariate causality measures. *Neuroimage*. 2010; 50(2):577–588. <https://doi.org/10.1016/j.neuroimage.2009.12.050> PMID: 20026279
26. Oostenveld R, Fries P, Maris E, Schoffelen JM. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci*. 2011; 2011:1–9. <https://doi.org/10.1155/2011/156869> PMID: 21253357
27. Gómez C, Lizier JT, Schaum M, Wollstadt P, Grützner C, Uhlhaas PJ, et al. Reduced predictable information in brain signals in autism spectrum disorder. *Frontiers in Neuroinformatics*. 2014; 8:9. <https://doi.org/10.3389/fninf.2014.00009> PMID: 24592235
28. Brodski-Guerniero A, Paasch GF, Wollstadt P, Oezdemir I, Lizier JT, Wibral M. Activating task relevant prior knowledge increases active information storage in content specific brain areas; 2016. Preprint. Available from: bioRxiv: <https://doi.org/10.1101/089300>. Cited 24 March 2017.
29. Wibral M, Lizier JT, Priesemann V. Bits from brains for biologically inspired computing. *Frontiers in Robotics and AI*. 2015; 2:5. <https://doi.org/10.3389/frobt.2015.00005>
30. Purdon PL, Pierce ET, Mukamel EA, Prerau MJ, Walsh JL, Wong KFK, et al. Electroencephalogram signatures of loss and recovery of consciousness from propofol. *PNAS*. 2013; 110(12):E1142–E1151. <https://doi.org/10.1073/pnas.1221180110> PMID: 23487781
31. Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cort*. 1991; 1(1):1–47. <https://doi.org/10.1093/cercor/1.1.1> PMID: 1822724
32. Salin PA, Bullier J. Corticocortical connections in the visual system: structure and function. *Physiol Rev*. 1995; 75(1):107–154. PMID: 7831395
33. Tomioka R, Okamoto K, Furuta T, Fujiyama F, Iwasato T, Yanagawa Y, et al. Demonstration of long-range GABAergic connections distributed throughout the mouse neocortex. *Eur J Neurosci*. 2005; 21(6):1587–1600. <https://doi.org/10.1111/j.1460-9568.2005.03989.x> PMID: 15845086
34. Clark A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci*. 2013; 36(03):181–204. <https://doi.org/10.1017/S0140525X12000477> PMID: 23663408
35. Hohwy J. *The Predictive Mind*. New York: Oxford University Press; 2013. <https://doi.org/10.1093/acprof:oso/9780199682737.001.0001>
36. Hawkins J, Blakeslee S. *On Intelligence*. New York: Macmillan; 2007.
37. Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. Canonical microcircuits for predictive coding. *Neuron*. 2012; 76(4):695–711. <https://doi.org/10.1016/j.neuron.2012.10.038> PMID: 23177956
38. Mashour GA. Top-down mechanisms of anesthetic-induced unconsciousness. *Front Syst Neurosci*. 2014; 8:115. <https://doi.org/10.3389/fnsys.2014.00115> PMID: 25002838
39. Brodski A, Paasch GF, Helbling S, Wibral M. The faces of predictive coding. *J Neurosci*. 2015; 35(24): 8997–9006. <https://doi.org/10.1523/JNEUROSCI.1529-14.2015> PMID: 26085625
40. Rivolta D, Heidegger T, Scheller B, Sauer A, Schaum M, Birkner K, et al. Ketamine dysregulates the amplitude and connectivity of high-frequency oscillations in cortical–subcortical networks in humans: evidence from resting-state magnetoencephalography-recordings. *Schizophr Bull*. 2015; 41(5): 1105–1114. <https://doi.org/10.1093/schbul/sbv051> PMID: 25987642
41. Spinney RE, Prokopenko M, Lizier JT. Transfer entropy in continuous time, with applications to jump and neural spiking processes; 2016. Preprint. Available from: arXiv:161008192v1. Cited 27 December 2016.
42. Fröhlich F, McCormick DA. Endogenous electric fields may guide neocortical network activity. *Neuron*. 2010; 67(1):129–143. <https://doi.org/10.1016/j.neuron.2010.06.005> PMID: 20624597

43. Tao T. Szemerédi's regularity lemma revisited; 2005. Preprint. Available from: arXiv:math/0504472. Cited 24 November 2016.
44. Roux F, Wibral M, Mohr HM, Singer W, Uhlhaas PJ. Gamma-band activity in human prefrontal cortex codes for the number of relevant items maintained in working memory. *Journal of Neuroscience*. 2012; 32(36):12411–12420. <https://doi.org/10.1523/JNEUROSCI.0421-12.2012> PMID: 22956832
45. Sellers KK, Fröhlich F. Data from: Breakdown of local information processing may underlie isoflurane anesthesia effects; 2017. Dryad Digital Repository. Available from: <http://dx.doi.org/10.5061/dryad.kk40s>.
46. Sellers KK, Bennett DV, Hutt A, Fröhlich F. Anesthesia differentially modulates spontaneous network dynamics by cortical area and layer. *J Neurophysiol*. 2013; 110(12):2739–2751. <https://doi.org/10.1152/jn.00404.2013> PMID: 24047911
47. Sellers KK, Bennett DV, Hutt A, Williams J, Fröhlich F. Awake versus anesthetized: layer-specific sensory processing in visual cortex and functional connectivity between cortical areas. *J Clin Neurophysiol*. 2015; 113:3798–3815.
48. Cover TM, Thomas JA. Elements of Information Theory. New York: Wiley; 2006. <https://doi.org/10.1002/047174882X>
49. Wibral M, Vicente R, Lindner M. Transfer entropy in neuroscience. In: Wibral M, Vicente R, Lizier JT editors. Directed Information Measures in Neuroscience. Heidelberg: Springer; 2014. pp. 3–58. https://doi.org/10.1007/978-3-642-54474-3_1
50. Williams PL, Beer RD. Generalized measures of information transfer; 2011. Preprint. Available from: arXiv:11021507. Cited 24 November 2016.
51. Takens F. Detecting strange attractors in turbulence. In: Dynamical Systems and Turbulence, Warwick 1980. vol. 898 of Lecture Notes in Mathematics. Berlin: Springer; 1981. p. 366–381. <https://doi.org/10.1007/BFb0091924>
52. Wibral M, Lizier JT, Vögler S, Priesemann V, Galuske R. Local active information storage as a tool to understand distributed neural information processing. *Front Neuroinform*. 2014; 8. <https://doi.org/10.3389/fninf.2014.00001> PMID: 24501593
53. Lizier JT. JIDT: an information-theoretic toolkit for studying the dynamics of complex systems. *Front Robot AI*. 2014; 1(11). <https://doi.org/10.3389/frobt.2014.00011>
54. Faes L, Porta A. Conditional entropy-based evaluation of information dynamics in physiological systems. In: Wibral M, Vicente R, Lizier JT editors. Directed Information Measures in Neuroscience. Heidelberg: Springer; 2014. pp. 61–86. https://doi.org/10.1007/978-3-642-54474-3_3
55. Staniak M, Lehnertz K. Symbolic transfer entropy. *Phys Rev Lett*. 2008; 100:158101. <https://doi.org/10.1103/PhysRevLett.100.158101> PMID: 18518155
56. Pompe B, Runge J. Momentary information transfer as a coupling measure of time series. *Phys Rev E*. 2011; 83(5):051122. <https://doi.org/10.1103/PhysRevE.83.051122> PMID: 21728505
57. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. *Phys Rev E*. 2004; 69(6): 066138. <https://doi.org/10.1103/PhysRevE.69.066138>
58. Frenzel S, Pompe B. Partial mutual information for coupling analysis of multivariate time series. *Phys Rev Lett*. 2007; 99(20):204101. <https://doi.org/10.1103/PhysRevLett.99.204101> PMID: 18233144
59. Kraskov A. Synchronization and interdependence measures and their application to the electroencephalogram of epilepsy patients and clustering of data. PhD Thesis. University of Wuppertal; 2004.
60. Khan S, Bandyopadhyay S, Ganguly AR, Saigal S, Erickson DJ III, Protopopescu V, et al. Relative performance of mutual information estimation methods for quantifying the dependence among short and noisy data. *Phys Rev E*. 2007; 76(2):026209. <https://doi.org/10.1103/PhysRevE.76.026209> PMID: 17930123
61. Wollstadt P, Martínez-Zarzuela M, Vicente R, Díaz-Pernas FJ, Wibral M. Efficient transfer entropy analysis of non-stationary neural time series. *PLoS One*. 2014; 9(7):e102833. <https://doi.org/10.1371/journal.pone.0102833> PMID: 25068489
62. Kozachenko LF, Leonenko NN. On statistical estimation of entropy of random vector. *Problems Inform Transm*. 1987; 23:95–101.
63. Wolpert DH, Wolf DR. Estimating functions of probability distributions from a finite set of samples. *Phys Rev E*. 1995; 52(6):6841. <https://doi.org/10.1103/PhysRevE.54.6973.2>
64. Faes L, Nollo G, Porta A. Information-based detection of nonlinear Granger causality in multivariate processes via a nonuniform embedding technique. *Phys Rev E*. 2011; 83(5):051112. <https://doi.org/10.1103/PhysRevE.83.051112> PMID: 21728495
65. Anderson M, Braak CT. Permutation tests for multi-factorial analysis of variance. *J Stat Comput Simul*. 2003; 73(2):85–113. <https://doi.org/10.1080/00949650215733>

66. Suckling J, Bullmore E. Permutation tests for factorially designed neuroimaging experiments. *Hum Brain Mapp*. 2004; 22(3):193–205. <https://doi.org/10.1002/hbm.20027> PMID: 15195286
67. Helbling S. Advances in MEG methods and their applications to investigate auditory perception. PhD Thesis. Goethe-University, Frankfurt; 2015.
68. Helbling S. permANOVA toolbox; 2017. Available from: <https://github.com/sashel/permANOVA>. Cited 26 January 2017.
69. Aarts E, Verhage M, Veenvliet JV, Dolan CV, van der Sluis S. A solution to dependency: using multilevel analysis to accommodate nested data. *Nat Neurosci*. 2014; 17(4):491–496. <https://doi.org/10.1038/nn.3648> PMID: 24671065
70. Fahrmeir L, Kneib T, Lang S. Regression. Berlin: Springer; 2007.
71. R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing; 2008. Vienna, Austria. <http://www.R-project.org>.
72. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 2015; 67(1):1–48.
73. Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hypothesis testing: keep it maximal. *J Mem Lang*. 2013; 68(3):255–278. <https://doi.org/10.1016/j.jml.2012.11.001> PMID: 24403724
74. Vakorin VA, Krakovska O, McIntosh AR. On complexity and phase effects in reconstructing the directionality of coupling in non-linear systems. In: Wibral M, Vicente R, Lizier JT editors. Directed Information Measures in Neuroscience. Heidelberg: Springer; 2014. pp. 137–158. https://doi.org/10.1007/978-3-642-54474-3_6