Understanding Neural Population Communication with Latent Channels

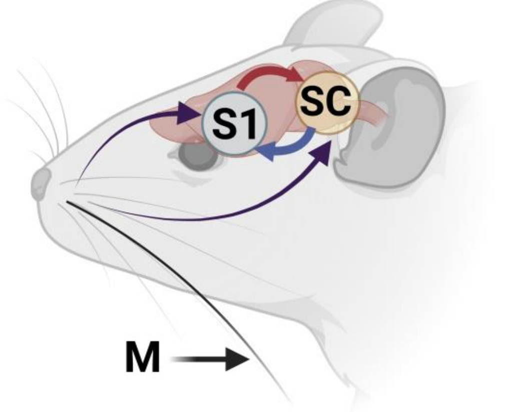
Amanda Merkley

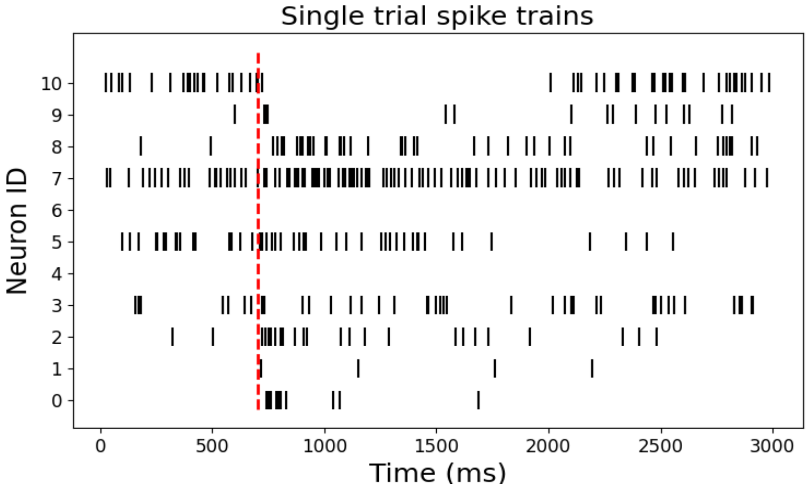
with Pulkit Grover

September 26, 2024

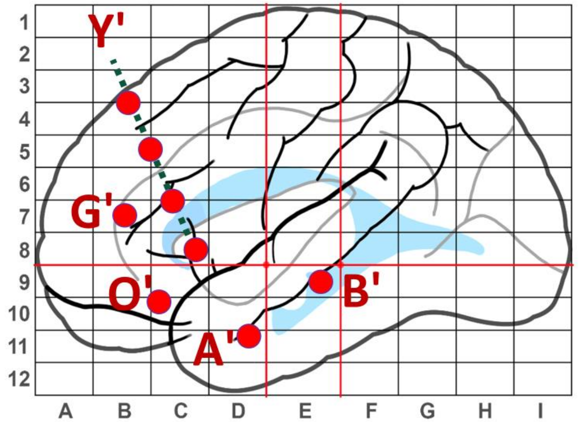
Large-scale brain recordings

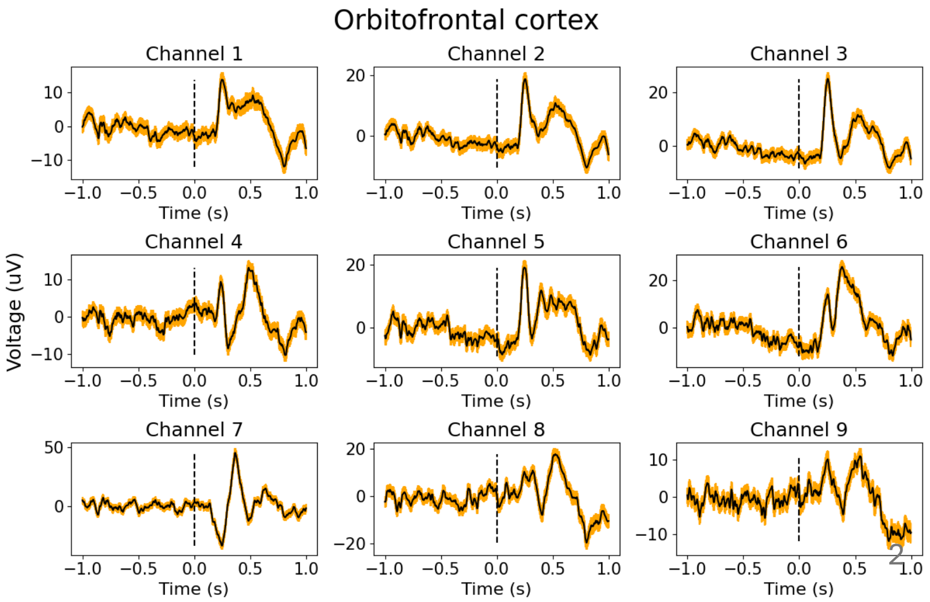
Neuroscience





Clinical translation

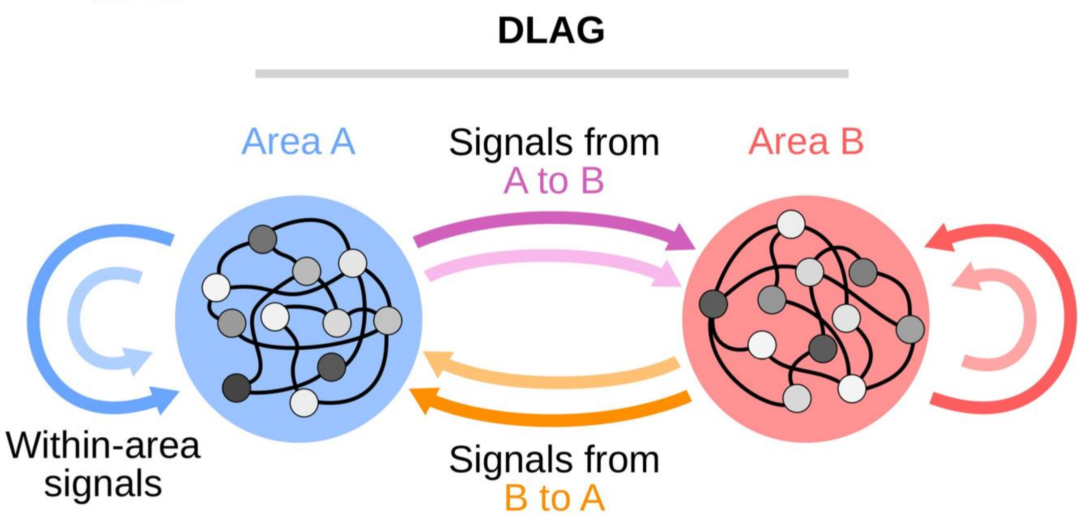




How do high-dimensional neural populations communicate messages?

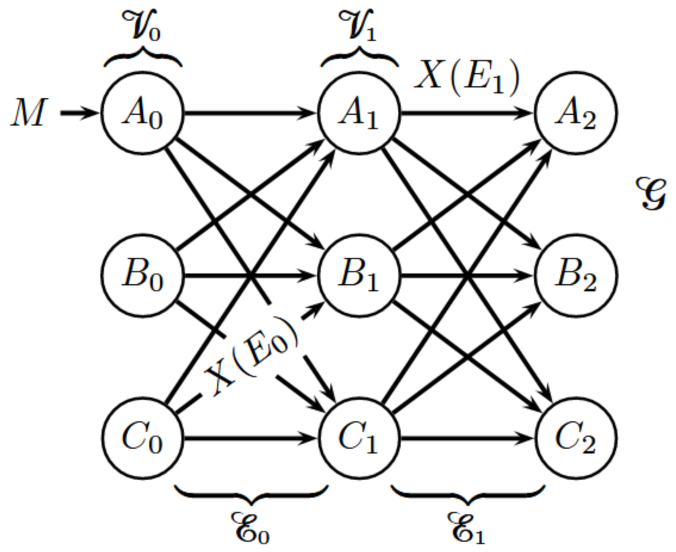
Previous work in population communication

Delayed Latents Across Groups



Identifies latents that predict activity in each population

M-Information Flow



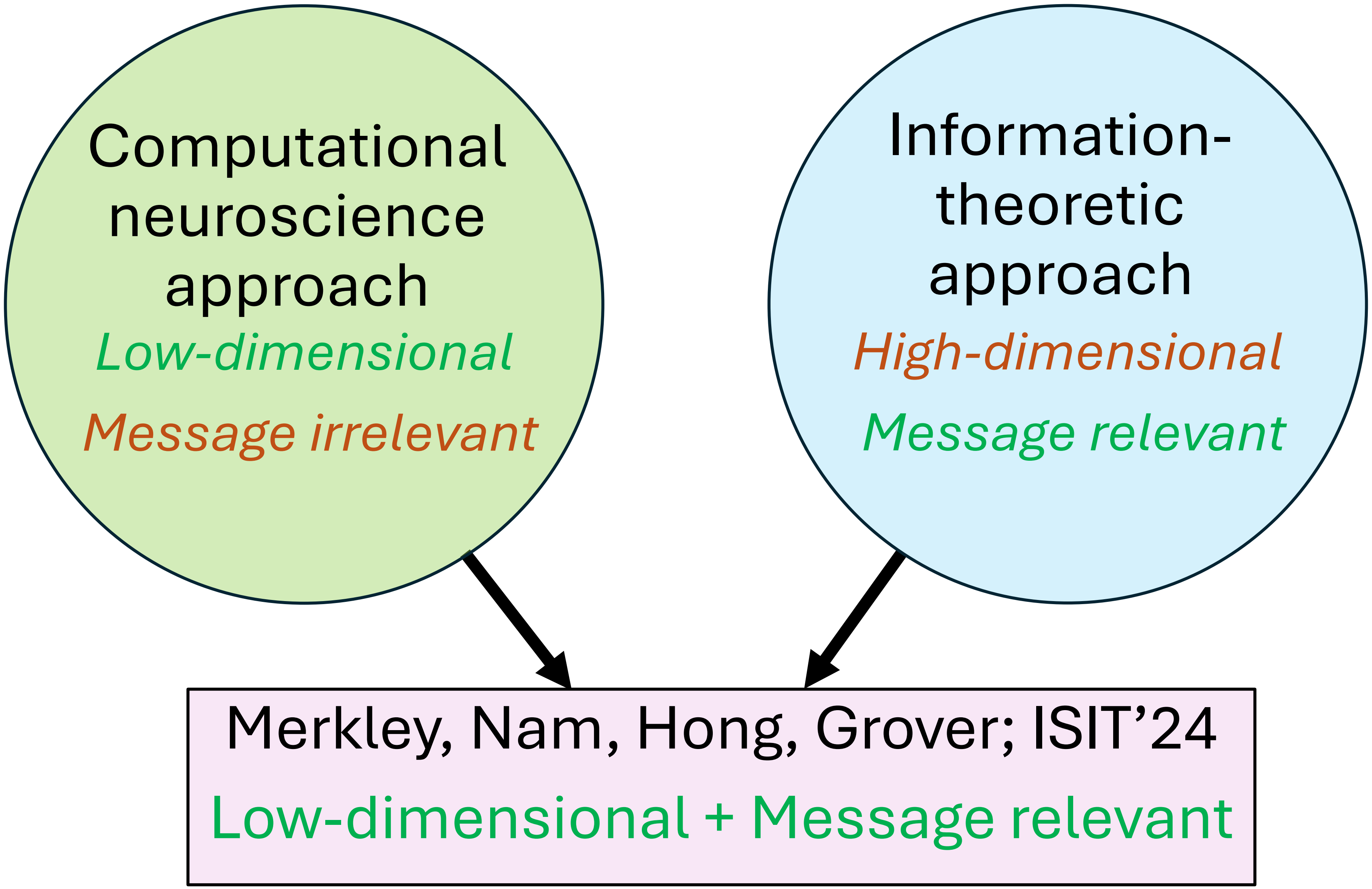
<u>Definition</u>: Information about M flows on an edge $E^{(t)} \subseteq \mathcal{E}^{(t)}$ if there is a set of edges $\mathcal{E}_0^{(t)} \subseteq \mathcal{E}^{(t)} \backslash \{E^{(t)}\}$ such that

$$I\left(M; E^{(t)} \middle| \mathcal{E}_0^{(t)}\right) > 0$$

Gokcen, Jasper, Semedo, Zandvakili, Kohn, Machens, Yu; 2022

Venkatesh, Dutta, Grover; ISIT'18, ISIT'19, TransIT'20

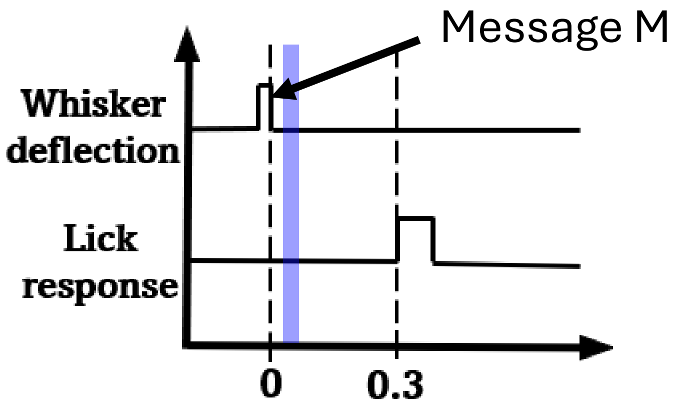
Previous work in population communication

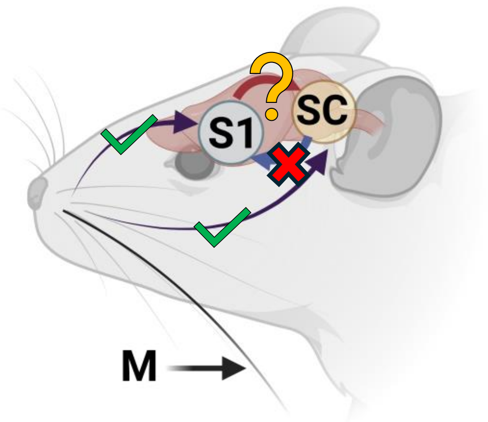


Previous work: Message forwarding

Does S1 communicate M to SC?

Does SC communicate M to S1?





Step 1:

Message-relevant dimension reduction of each population via correlation maximization

Step 2:

Infer the existence of communication structures via hypothesis test

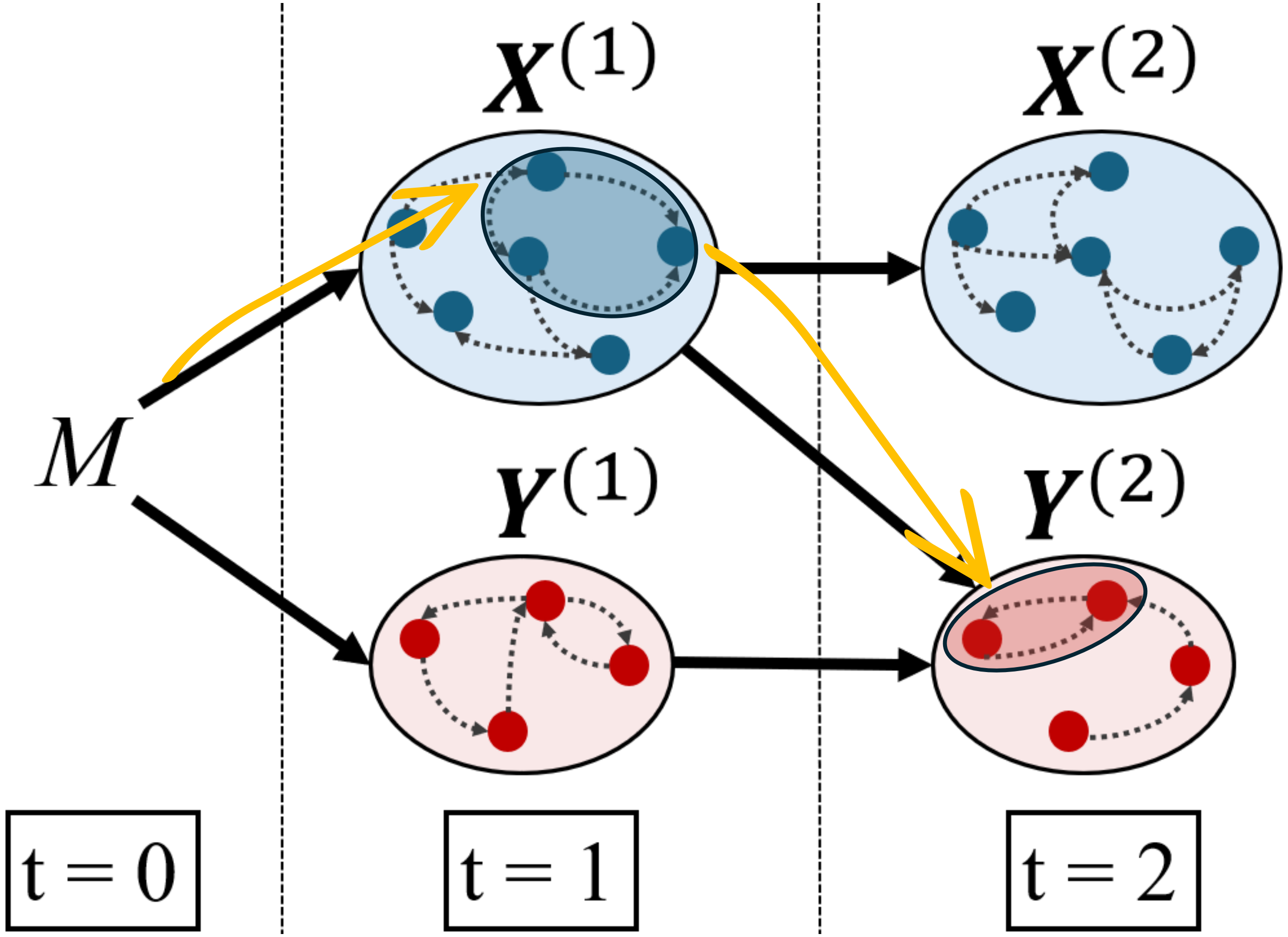
***M*-forwarding**

$$M \to A \to B$$

Definition: ***A*** *M*-forwards to ***B*** if $I(M; \boldsymbol{B} | A) = 0$.

Merkley, Nam, Hong, Grover; ISIT'24

Latent Channels: abstracting communication



Definition: A Latent Channel (LC) from population $\boldsymbol{X}^{(1)}$ to population $\boldsymbol{Y}^{(2)}$ is

$$M - f\left(\boldsymbol{X}^{(1)}\right) - g\left(\boldsymbol{Y}^{(2)}\right)$$

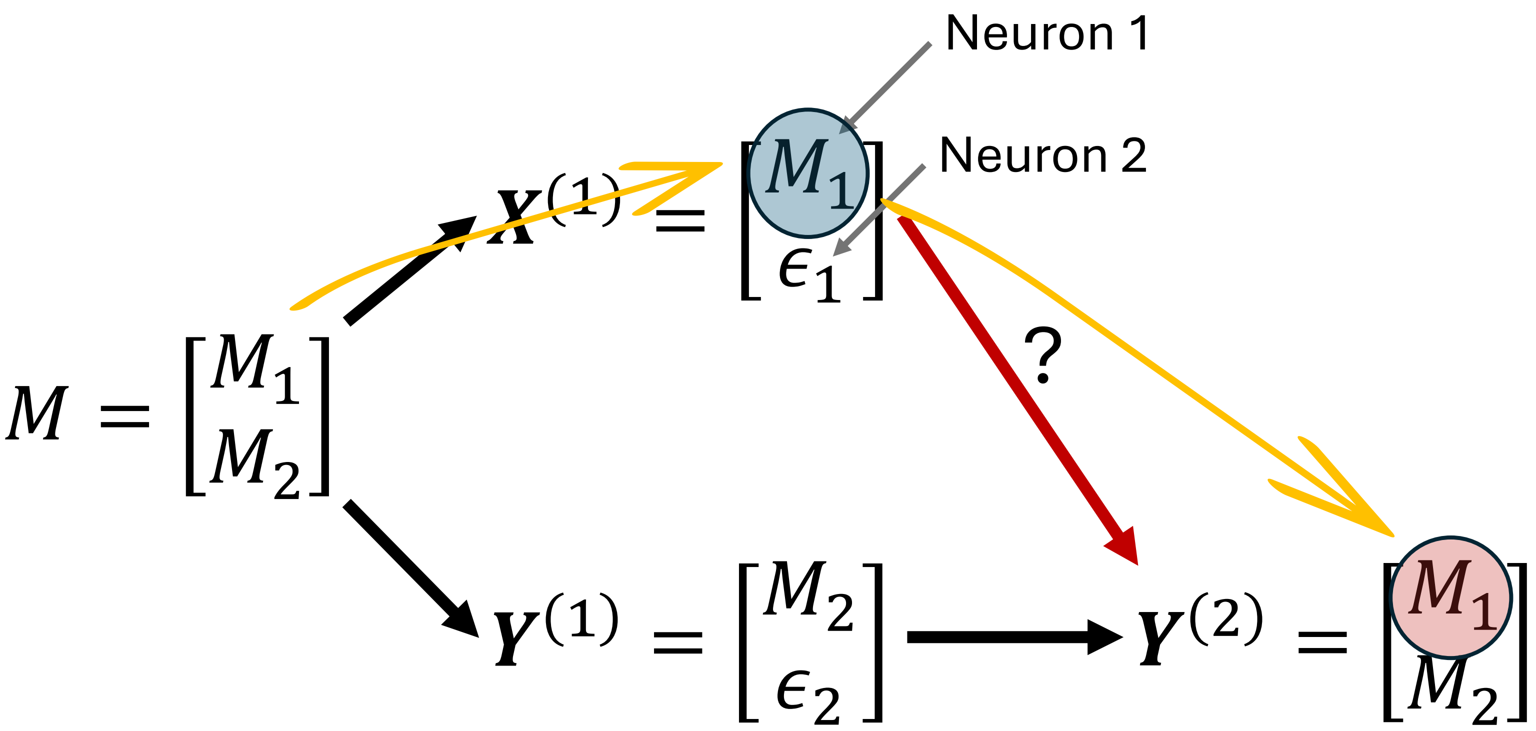
for projections, f and g, of $\boldsymbol{X}^{(1)}$ and $\boldsymbol{Y}^{(2)}$.

Which LCs are meaningful?

Goal: characterize types of LCs to define a *valid* LC

Motivating example

Consider a 2-dimensional message $M = [M_1, M_2]$ and 2 populations of 2 neurons. M_i, ϵ_i for $i = 1{,}2$ are mutually independent.



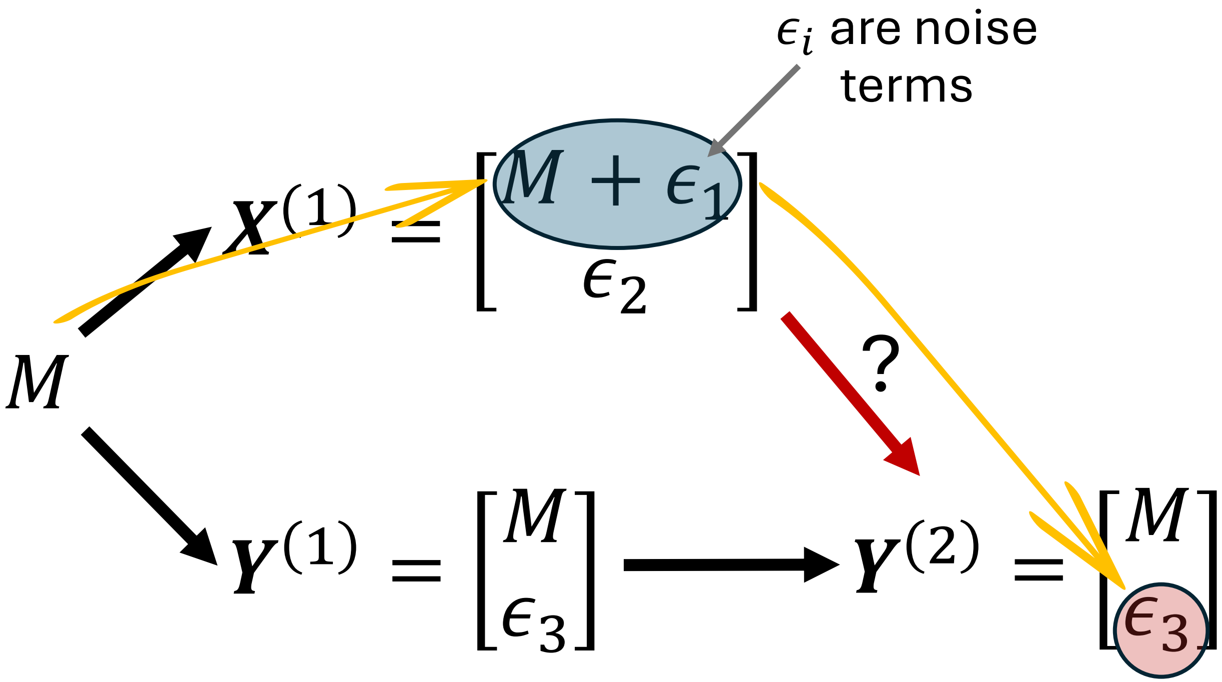
Linearly project of $\boldsymbol{X}^{(1)}$ and $\boldsymbol{Y}^{(2)}$: $M - f(\boldsymbol{X}^{(1)}) - g(\boldsymbol{Y}^{(2)})$

Desirable property:

Projections of $\boldsymbol{X}^{(1)}$ and $\boldsymbol{Y}^{(2)}$ should maximize information with $\boldsymbol{M}$

Phantom Latent Channels

Consider jointly Gaussian $P\left(M, X^{(1)}, Y^{(2)}\right)$ where each population has 2 neurons. Find an LC from $X^{(1)}$ to $Y^{(2)}$.

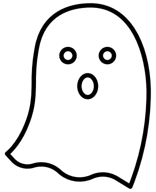


The only LC is

$$M \; - (M + \epsilon_1) - \epsilon_3$$

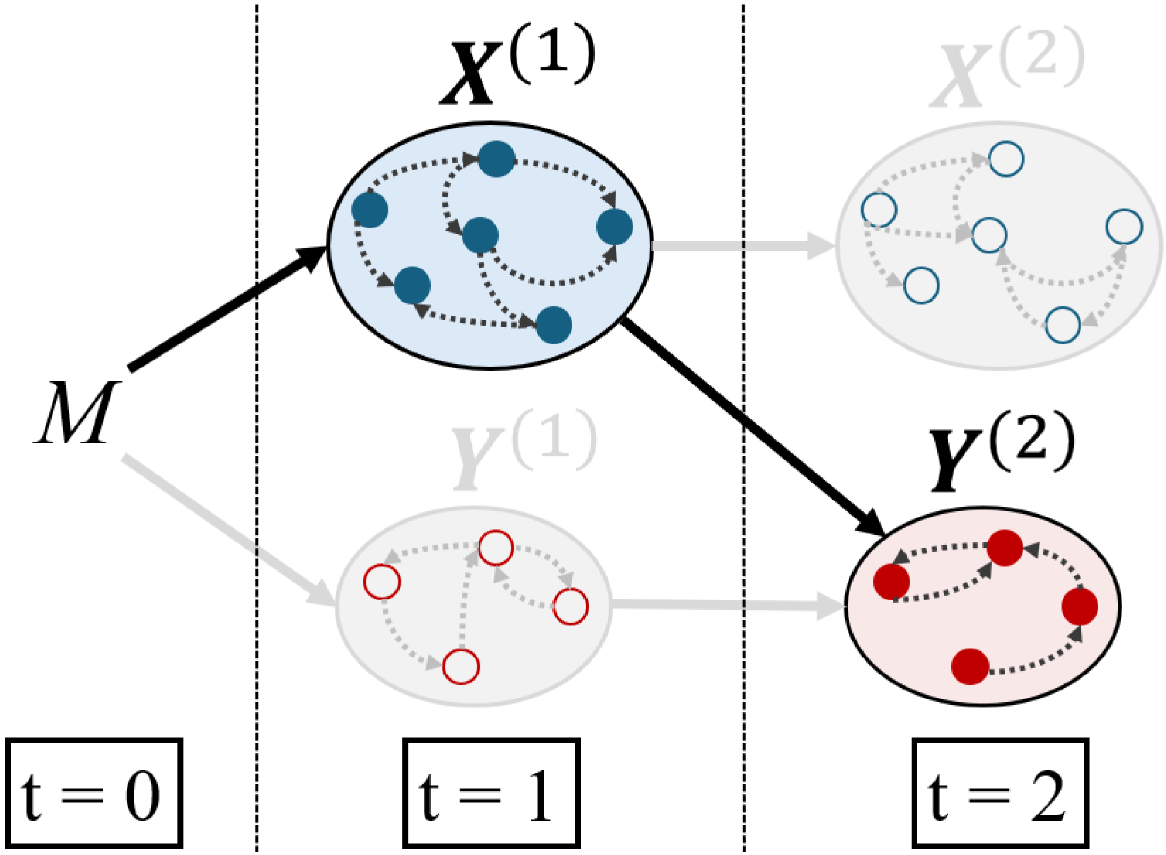
Note that $I\left(M; g\left(\boldsymbol{Y}^{(2)}\right)\right) = 0.$
Captures degenerate forwarding.

<ins>Definition</ins>: The LC given by $M - f\left(\boldsymbol{X}^{(1)}\right) - g\left(\boldsymbol{Y}^{(2)}\right)$ is a phantom LC if $I\left(M; g\left(\boldsymbol{Y}^{(2)}\right)\right) = 0.$



Marginal Latent Channels

So far, focus is on M-forwarding in anatomically distinct populations $\boldsymbol{X}$ and $\boldsymbol{Y}$.



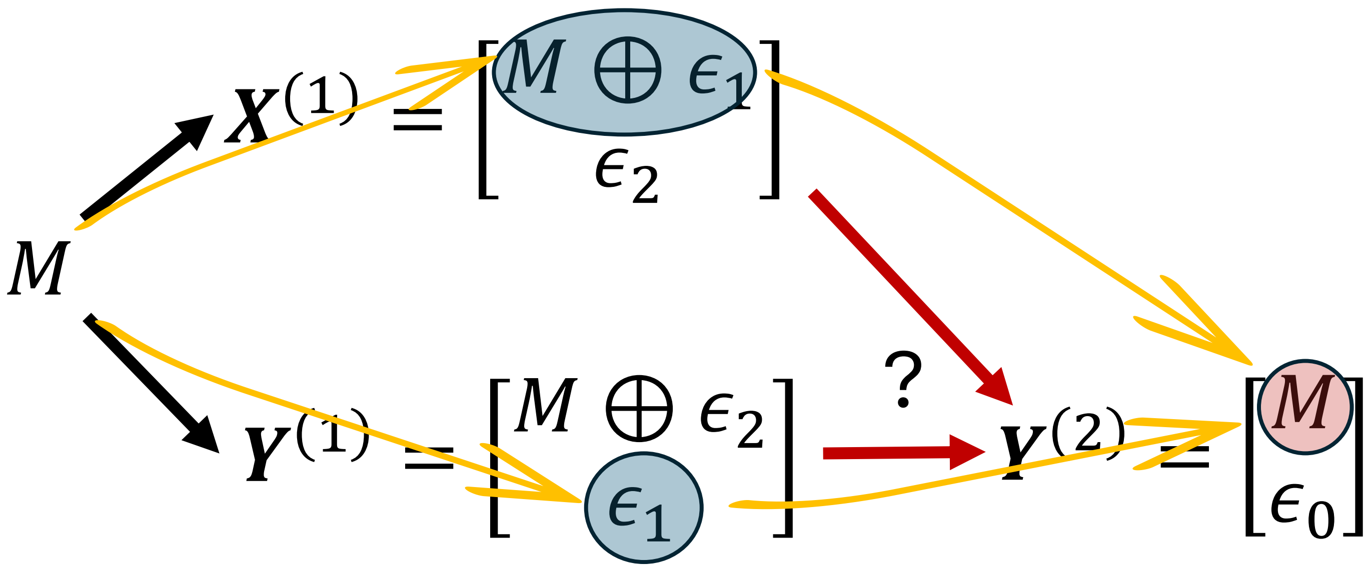
Definition: A marginal LC is an LC, $M - f\left(\boldsymbol{X}^{(1)}, \boldsymbol{Y}^{(1)}\right) - g(\boldsymbol{X}^{(2)}, \boldsymbol{Y}^{(2)})$, where f and g depend on only one of $\boldsymbol{X}^{(i)}$ or $\boldsymbol{Y}^{(i)}$.

ISIT'24 work only identifies marginal LCs

Intuitive to extend to LCs based on joint population activity

Joint Latent Channels

Consider $M, \epsilon_i \sim Bern(0.5)$ that are mutually independent.



Note: $I\left(M; \boldsymbol{X}^{(1)}\right) = 0$ and $I\left(M; \boldsymbol{Y}^{(1)}\right) = 0$, but $I\left(M; \boldsymbol{Y}^{(2)}\right) > 0$

One possible LC is:

$$M - \left(X_1^{(1)} \oplus Y_2^{(1)}\right) - Y_1^{(2)}$$

<u>Definition:</u> A joint LC is an LC, $M - f\left(\boldsymbol{X}^{(1)}, \boldsymbol{Y}^{(1)}\right) - g(\boldsymbol{X}^{(2)}, \boldsymbol{Y}^{(2)})$, where f or g depend on both $\boldsymbol{X}^{i}$ and $\boldsymbol{Y}^{i}$.

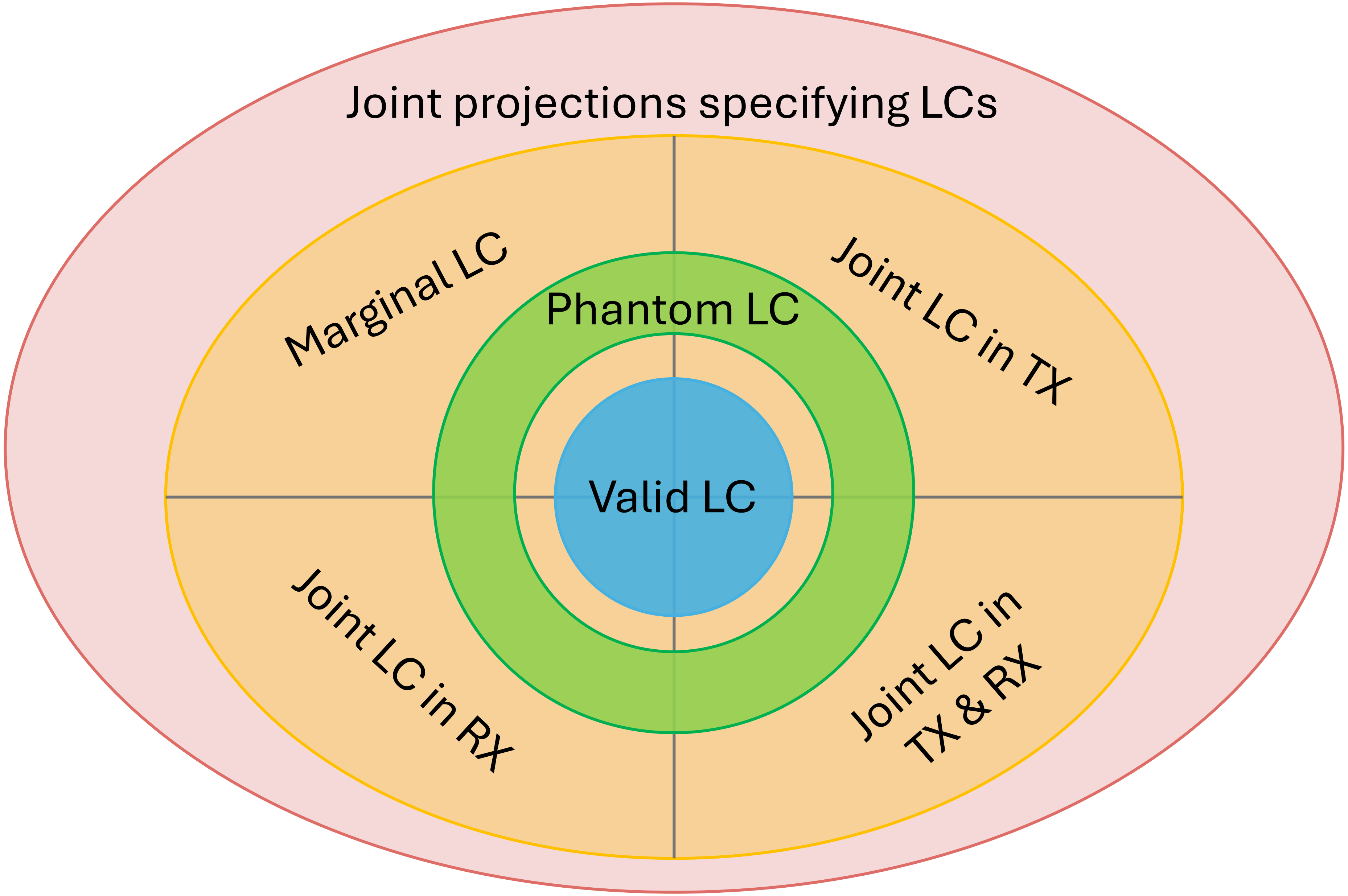
Valid Latent Channels

General LC Algorithm

1. Find information-maximizing projection of TX population
2. Identify the set of projections of RX population such that M-forwarding holds
3. Find information-maximizing projection of RX population from above

Definition: A valid LC is an LC found through the General LC Algorithm that is not a phantom LC.

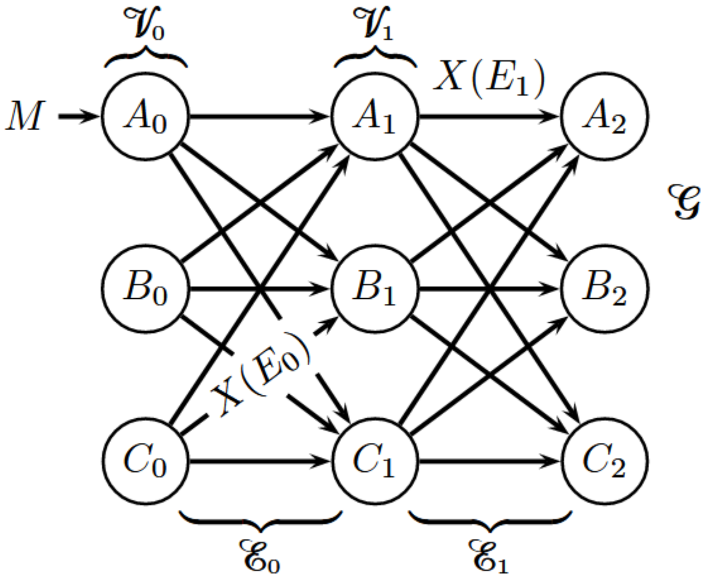
All joint projections $P(M, f(\boldsymbol{X}), g(\boldsymbol{Y}))$



Valid LC > M-information flow

*Detail slide

Definition: Information about M flows on an edge $E^{(t)} \subseteq \mathcal{E}^{(t)}$ if there is a set of edges $\mathcal{E}_0^{(t)} \subseteq \mathcal{E}^{(t)} \setminus \{E^{(t)}\}$ such that $I\left(M; E^{(t)} \middle| \mathcal{E}_0^{(t)}\right) > 0$.



Assumption: Activity $X^{(t)}$ observed at a neuron (vertex) is the same as activity $E^{(t)}$ observed on its outgoing edge, i.e. $X^{(t)} = E^{(t)}$.

Theorem:
If $M - f(\boldsymbol{X}^{(1)}) - g(\boldsymbol{Y}^{(2)})$ is a valid LC, then there is M-information flow between $\boldsymbol{X}^{(1)}$ and $\boldsymbol{Y}^{(2)}$.

Proof sketch:

$$0 < I\big(M; g(\boldsymbol{Y}^{(2)})\big) \leq I\big(M; f(\boldsymbol{X}^{(1)})\big) \leq I\big(M; \boldsymbol{X}^{(1)}\big) = I(M; \boldsymbol{X}^{(1)} | \emptyset)$$

Converse does not hold since $I\left(M; E^{(t)} \middle| \mathcal{E}_0^{(t)}\right) > 0$ does not imply $I(M; E^{(t)}) > 0$.

Structure between populations

Partial Information Decomposition (PID)

Unique information X has about M wrt Y

Unique information in Y

Redundant information in X and Y about M

Synergistic information X and Y have jointly about M

$$I(M; X, Y) = UI(X) + UI(Y) + RI + SI$$

$$I(M; X) = UI(X) + RI$$

$$I(M; Y) = UI(Y) + RI$$

Blackwellian PID measure (Venkatesh, Schamberg, 2022): If $Y|M$ is stochastically degraded wrt $X|M$, $UI(Y) = 0$.

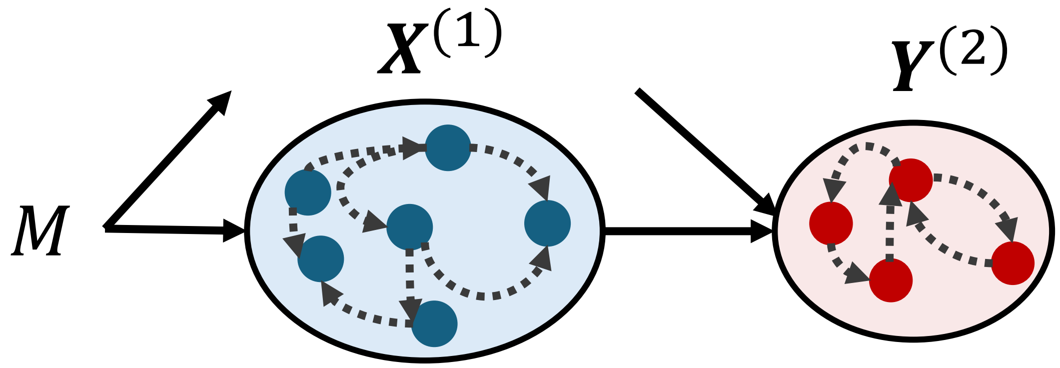
If $M - X - Y$ holds (physically degraded), then $SI = 0$ too.

Williams, Beer; 2010. Bertschinger, Rauh, Olbrich, Jost, Ay; 2014. Blackwell; 1953. Goswami, Merkley; 2024.

*Detail slide

Structure through PID

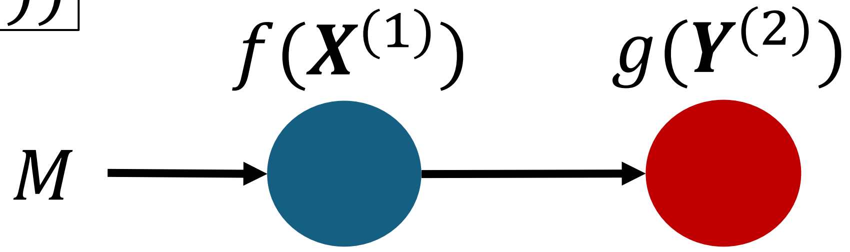
$\boxed{P(M, \boldsymbol{X}^{(1)}, \boldsymbol{Y}^{(2)})}$



$$I_P\left(M; \boldsymbol{X}^{(1)}, \boldsymbol{Y}^{(2)}\right) = UI_P\left(\boldsymbol{X}^{(1)}\right) + \underbrace{UI_P\left(\boldsymbol{Y}^{(2)}\right)}_{=0} + RI_P + SI_P$$

If P is stochastically degraded

$\boxed{S(M, f(\boldsymbol{X}^{(1)}), g(\boldsymbol{Y}^{(2)}))}$



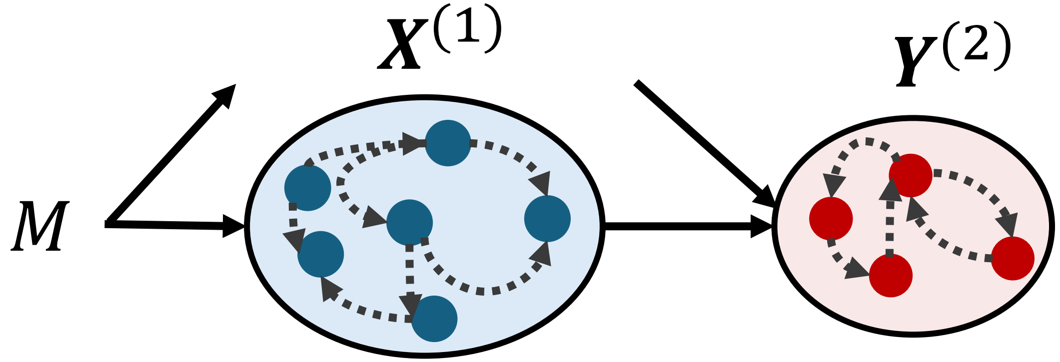
Amount of information transmitted

$$I_S\left(M; f(\boldsymbol{X}^{(1)}), g(\boldsymbol{Y}^{(2)})\right) = UI_S\left(f(\boldsymbol{X}^{(1)})\right) + \underbrace{UI_S\left(g(\boldsymbol{Y}^{(2)})\right)}_{=0} + RI_S + \underbrace{SI_S}_{=0}$$

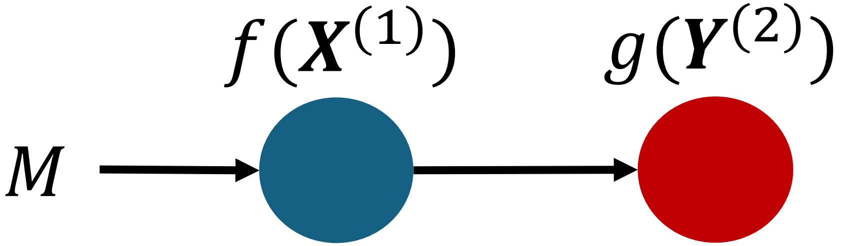
*Detail slide

Structure through PID

$P(M, \boldsymbol{X}^{(1)}, \boldsymbol{Y}^{(2)})$



$S(M, f(\boldsymbol{X}^{(1)}), g(\boldsymbol{Y}^{(2)}))$



Structure in P

If $M - \boldsymbol{X} - \boldsymbol{Y}$ and f is a sufficient projection of $\boldsymbol{X}$, then $M - f(\boldsymbol{X}) - g(\boldsymbol{Y})$ holds for all projections g of $\boldsymbol{Y}$. If g is sufficient, the PID is invariant from P to S.

Structure in S

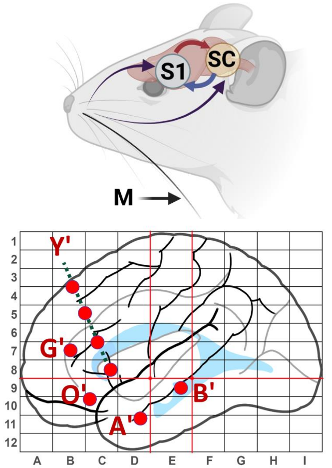
If $M - f(\boldsymbol{X}) - g(\boldsymbol{Y})$ holds for sufficient projections f and g, then $\frac{UI_S(f(\boldsymbol{X}))}{RI_S} \leq \frac{UI_P(\boldsymbol{X})}{RI_P}$.

Structure in P & S

If $M - f(\boldsymbol{X}) - g(\boldsymbol{Y})$ holds for sufficient f and arbitrary g and P is a stochastically degraded system, $\frac{UI_P(\boldsymbol{X})}{RI_P} \leq \frac{UI_S(f(\boldsymbol{X}))}{RI_S}$.

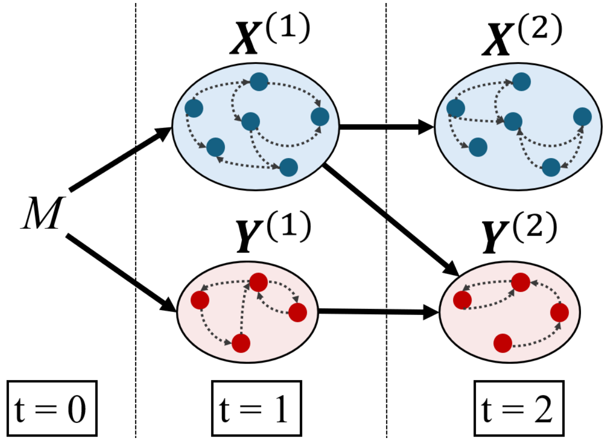
Summary

Defining neural population communication



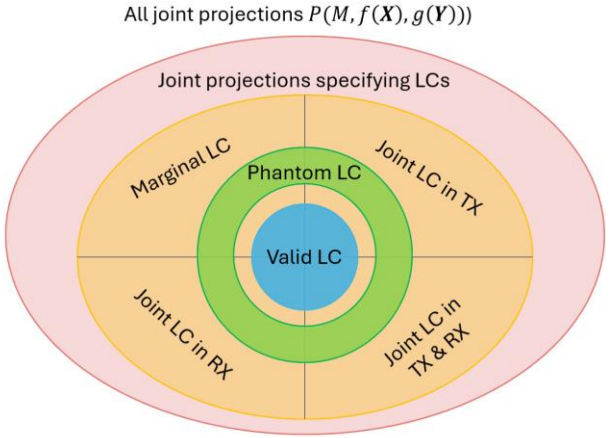
M-forwarding

Latent Channels



Low-dimensional + Message relevant

Taxonomy of Latent Channels



Structure through PID

$$I(M;X,Y) = UI(X) + UI(Y) + RI + SI$$

$$I(M;X) = UI(X) + RI$$

$$I(M;Y) = UI(Y) + RI$$

Acknowledgements

Dr. Pulkit Grover

NIH NSF

Contact me at
merkley@cmu.edu