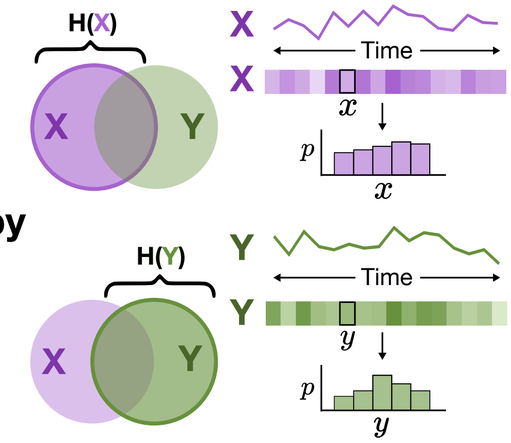
## 1. Single-process, order-independent measures

### Entropy

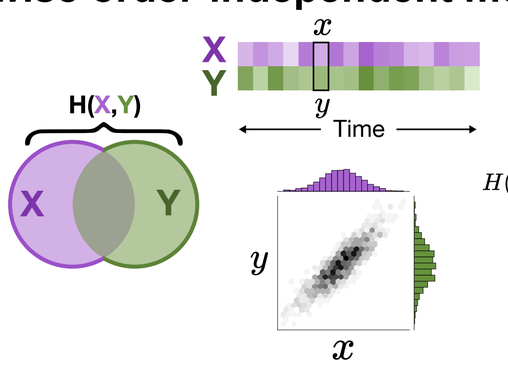



$$H(X) = -\sum_{x \in X} p(x) \log p(x)$$

$$H(Y) = -\sum_{y \in Y} p(y) \log p(y)$$

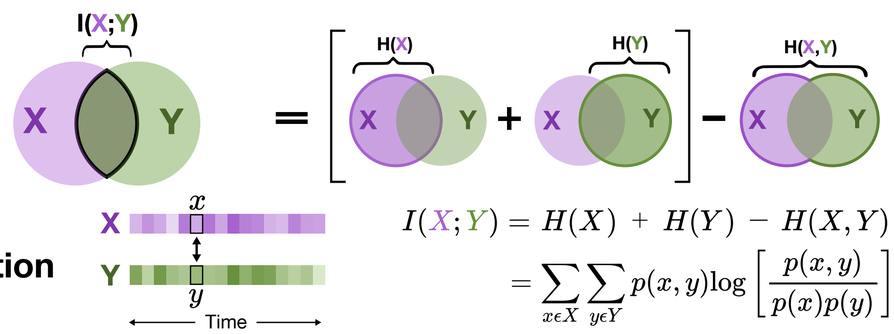
## 2. Pairwise order-independent measures, undirected

### Joint entropy (JE)




$$H(X,Y) = -\sum_{x \in X} \sum_{y \in Y} p(x,y) \log p(x,y)$$

### Mutual information (MI)

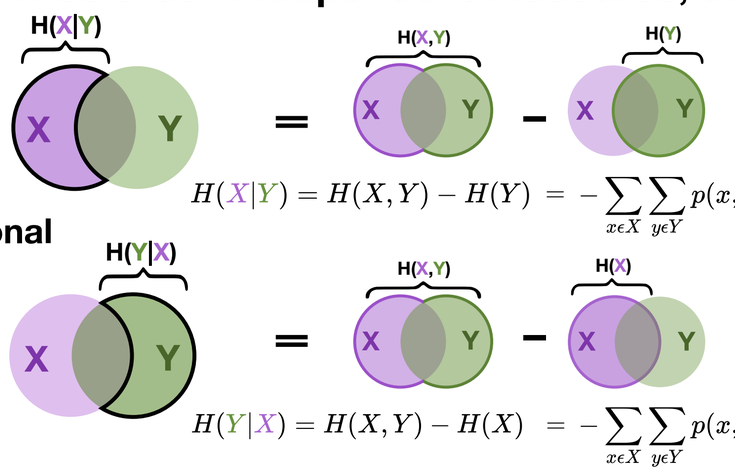



$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

$$= \sum_{x \epsilon X} \sum_{y \epsilon Y} p(x,y) \log \left[ \frac{p(x,y)}{p(x)p(y)} \right]$$

## 3. Pairwise order-independent measures, directed

### Conditional entropy (CE)

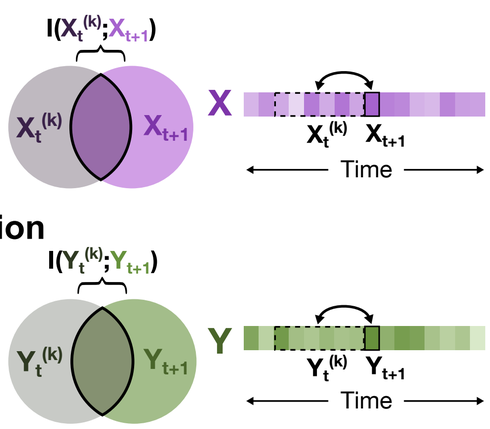



$$H(X|Y) = H(X,Y) - H(Y) = -\sum_{x \epsilon X} \sum_{y \epsilon Y} p(x,y) \log p(x \mid y)$$

$$H(Y|X) = H(X,Y) - H(X) = -\sum_{x \epsilon X} \sum_{y \epsilon Y} p(x,y) \log p(y \mid x)$$

## 4. Single-process, order-dependent measures

### Active information storage

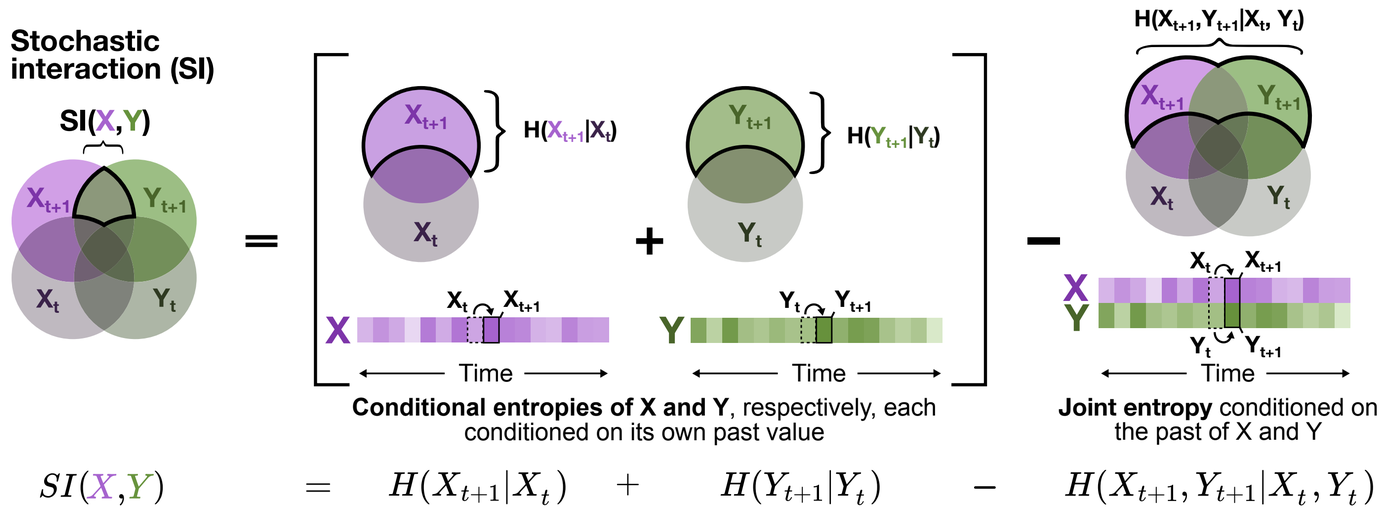



$$A(X)(k) = I(X_t^{(k)}; X_{t+1})$$

$$A(Y)(k) = I(Y_t^{(k)}; Y_{t+1})$$

## 5. Pairwise order-dependent measures, undirected

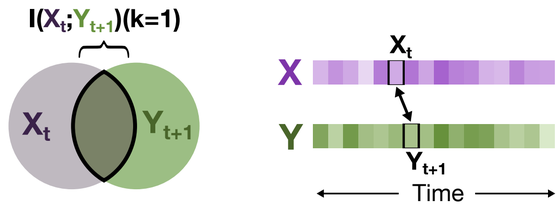
### Stochastic interaction (SI)




$$SI(X,Y) = H(X_{t+1}|X_t) + H(Y_{t+1}|Y_t) - H(X_{t+1}, Y_{t+1}|X_t, Y_t)$$

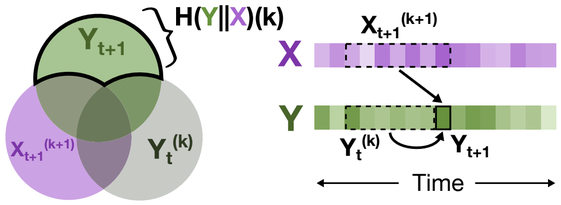
## 6. Pairwise order-dependent measures, directed

### Time-lagged mutual information (TLMI)



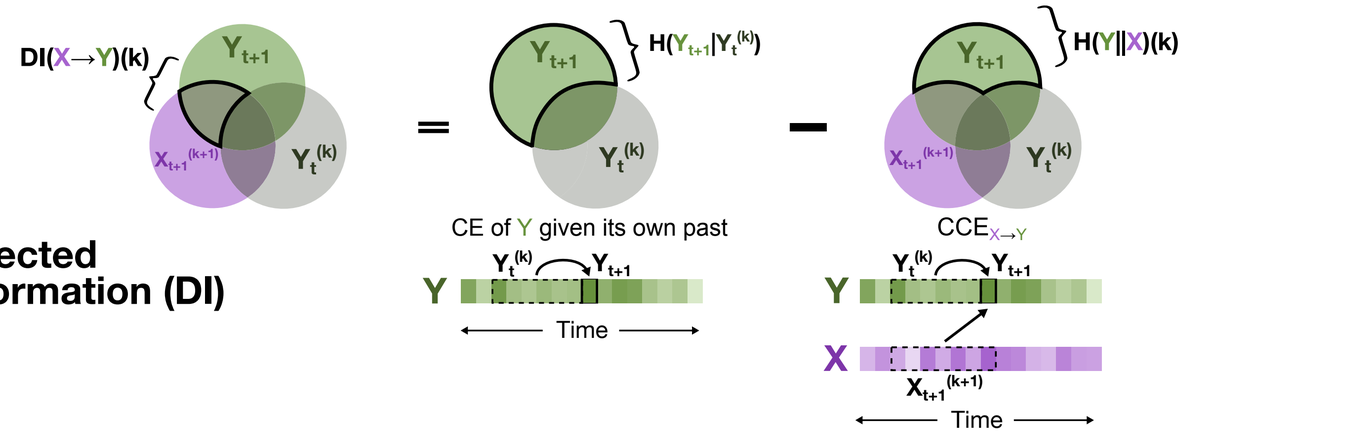

$$I(X_t; Y_{t+1}) = H(X_t) + H(X_t, Y_{t+1}) - H(Y_{t+1})$$

### Causal conditioned entropy (CCE)



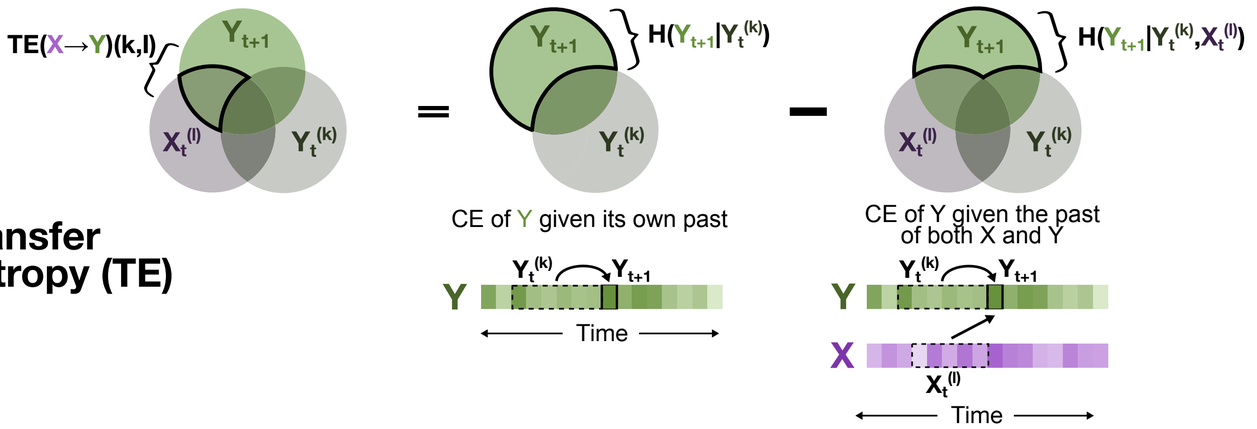

$$H(Y||X)(k) = H(Y_{t+1}|Y_t^{(k)}, X_{t+1}^{(k+1)})$$

### Directed information (DI)



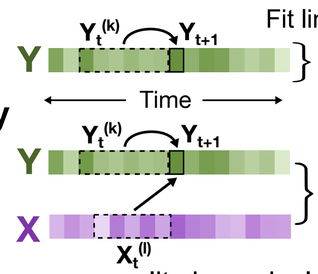

$$DI(X \to Y)(k) = I(Y_{t+1}; X_{t+1}^{(k+1)} | Y_t^{(k)}) = H(Y_{t+1}|Y_t^{(k)}) - H(Y||X)(k)$$

### Transfer entropy (TE)




$$TE(X \to Y)(k,l) = I(Y_{t+1}; X_t^{(l)} | Y_t^{(k)}) = H(Y_{t+1}|Y_t^{(k)}) - H(Y_{t+1}|Y_t^{(k)}, X_t^{(l)})$$

### Granger causality (GC)




Fit linear autoregressive models:

$Y_{t+1} \sim Y_t^{(k)}$ [reduced]

$Y_{t+1} \sim Y_t^{(k)}$ & $X_t^{(l)}$ [full]

Compute log-ratio of residual variance between the **reduced** vs. **full** models:

$$GC(X \to Y)(k,l) = \log \left[ \frac{\sum_{t=1}^{T} (\hat{y}[\text{reduced}]_{t+1} - y_{t+1})^2}{\sum_{t=1}^{T} (\hat{y}[\text{full}]_{t+1} - y_{t+1})^2} \right]$$

*Note: Granger causality is equivalent to transfer entropy with a Gaussian density estimator.