The Information Bottleneck for Optimal Prediction of the Voter Model

Robin Lamarche-Perrin
Sven Banisch
⇒ Eckehard Olbrich

Max-Planck-Institut für Mathematik in den Naturwissenschaften

Conference on Complex Systems 2015

Tempe, Arizona / September 28 – October 2

mathe MACS
Mathematics for Multilevel Anticipatory Complex Systems

MAX-PLANCK-GESELLSCHAFT

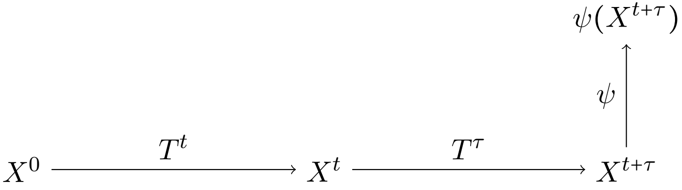
General Setting

$$X^0 \xrightarrow{\quad T^t \quad} X^t \xrightarrow{\quad T^\tau \quad} X^{t+\tau}$$

- Markovian Kernel $T(X^{t+1}|X^t)$

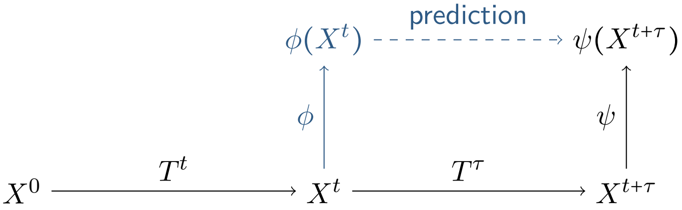
- Initial State $X^0 \in \Sigma$
- Current State $X^t \in \Sigma$ with Current Time $t \in \mathbb{N}$
- Future State $X^{t+\tau} \in \Sigma$ with Prediction Horizon $\tau \in \mathbb{N}$

General Setting



- Markovian Kernel $T(X^{t+1}|X^t)$
- Initial State $X^0 \in \Sigma$
- Current State $X^t \in \Sigma$ with Current Time $t \in \mathbb{N}$
- Future State $X^{t+\tau} \in \Sigma$ with Prediction Horizon $\tau \in \mathbb{N}$
- Post-measurement $\psi : \Sigma \to \mathcal{S}_\psi$ defined by $\Pr(\psi(X)|X)$

General Setting

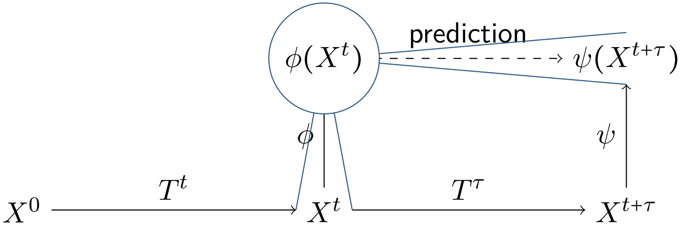


- Markovian Kernel $T(X^{t+1}|X^t)$

- Initial State $X^0 \in \Sigma$
- Current State $X^t \in \Sigma$ with Current Time $t \in \mathbb{N}$
- Future State $X^{t+\tau} \in \Sigma$ with Prediction Horizon $\tau \in \mathbb{N}$

- Post-measurement $\psi : \Sigma \to \mathcal{S}_\psi$ defined by $\Pr(\psi(X)|X)$
- Pre-measurement $\phi : \Sigma \to \mathcal{S}_\phi$ defined by $\Pr(\phi(X)|X)$

The Optimal Prediction Problem



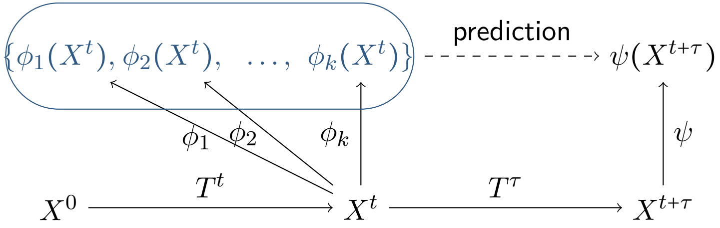
The Information Bottleneck Method [Tishby *et al.*, 1999]:

- **Maximize** Predictive Capacity $max_\phi \; I(\phi(X^t); \psi(X^{t+\tau}))$
- **Minimize** Measurement Complexity $min_\phi \; I(X^t; \phi(X^t))$

- **Minimize** the IB-variational

$$min_{\Pr(\hat{X}|X)} \quad I(X^t; \phi(X^t)) - \beta \; I(\phi(X^t); \psi(X^{t+\tau})) \quad \text{with } \beta \in \mathbb{R}^+$$

Constraining the Set of Feasible Measurements

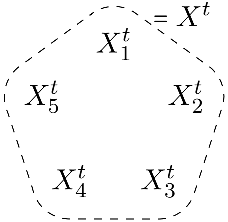


- Given a collection $\Phi = \{\phi_1, \ldots, \phi_k\}$ of *feasible* pre-measurements
- **Minimize** the IB-variational

$$min_{\phi \in \Phi} \quad I(X^t; \phi(X^t)) \ - \ \beta \ I(\phi(X^t); \psi(X^{t+\tau})) \quad \text{with } \beta \in \mathbb{R}^+$$

Application to Agent-based Systems

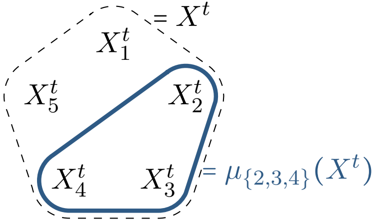
- Agent Set $\Omega = \{1, \ldots, N\}$
- Agent States $X_1^t \in S, \quad X_2^t \in S, \quad \ldots, \quad X_k^t \in S$
- System State $X^t = (X_1^t, X_2^t, \ldots, X_N^t) \in \Sigma = S^N$



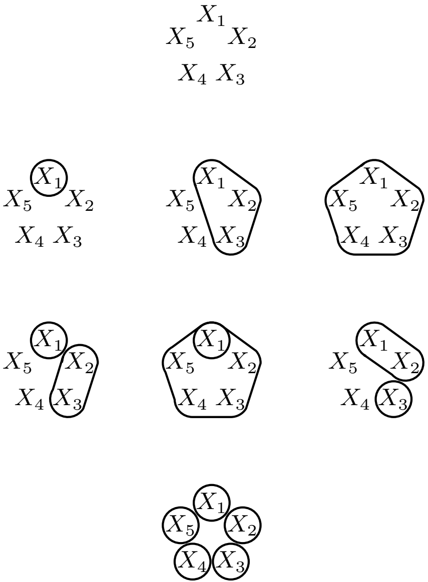
Application to Agent-based Systems

- Agent Set $\Omega = \{1, \ldots, N\}$
- Agent States $X_1^t \in S, \quad X_2^t \in S, \quad \ldots, \quad X_k^t \in S$
- System State $X^t = (X_1^t, X_2^t, \ldots, X_N^t) \in \Sigma = S^N$

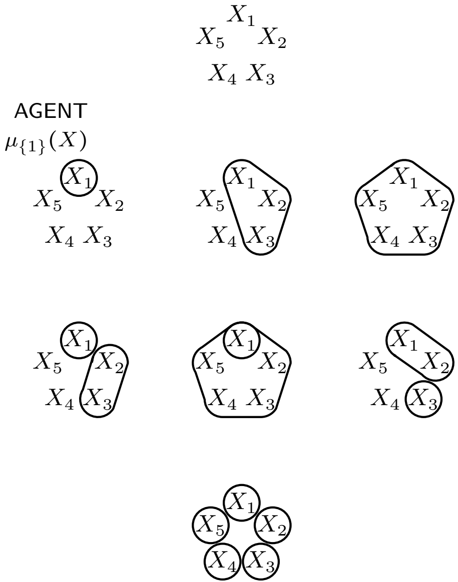
- Generic Measurement μ
  a familly of measurements $(\mu_A : \Sigma \to \mathcal{S}_\mu)$ for any $A \subset \Omega$
  such that $\Pr(\mu_A(X)|X) = \Pr(\mu_A(X)|(X_i)_{i\in A})$



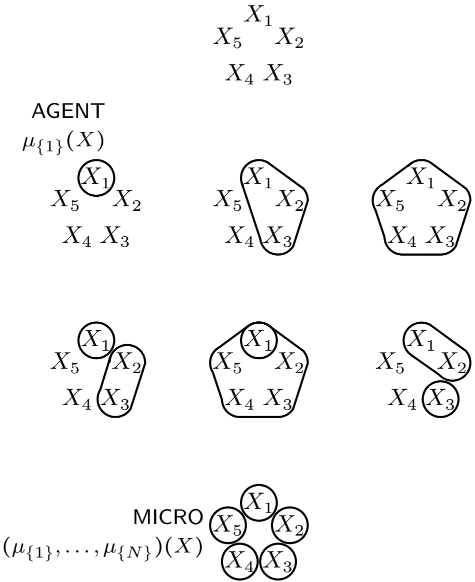
The Poset of Feasible Measurements



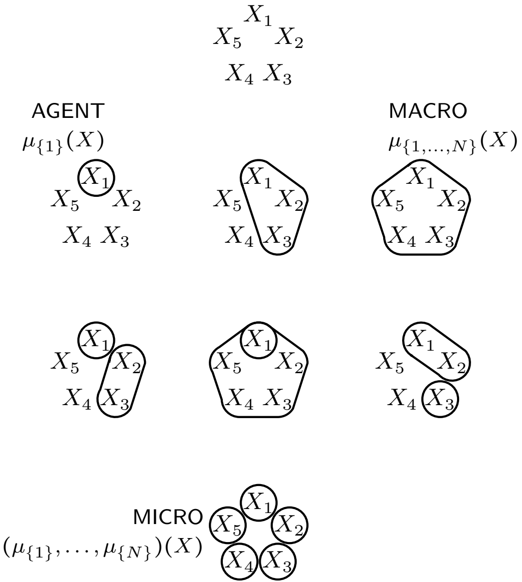
The Poset of Feasible Measurements



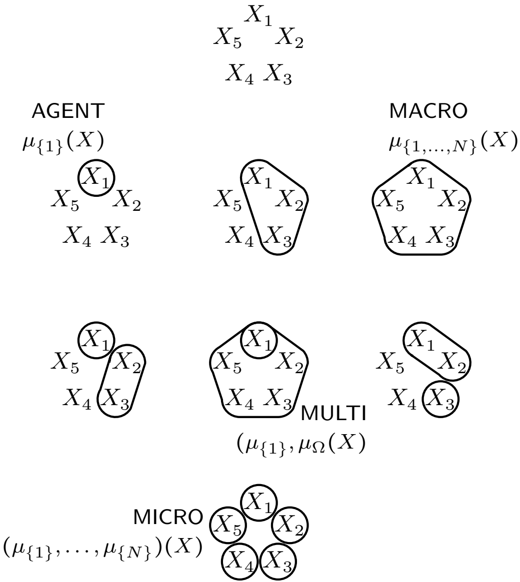
The Poset of Feasible Measurements



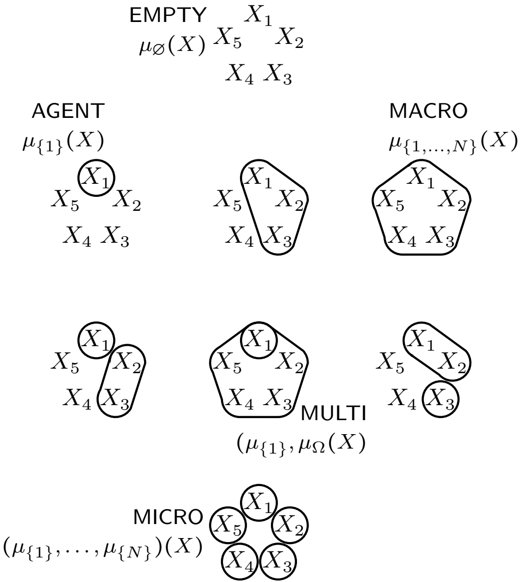
The Poset of Feasible Measurements



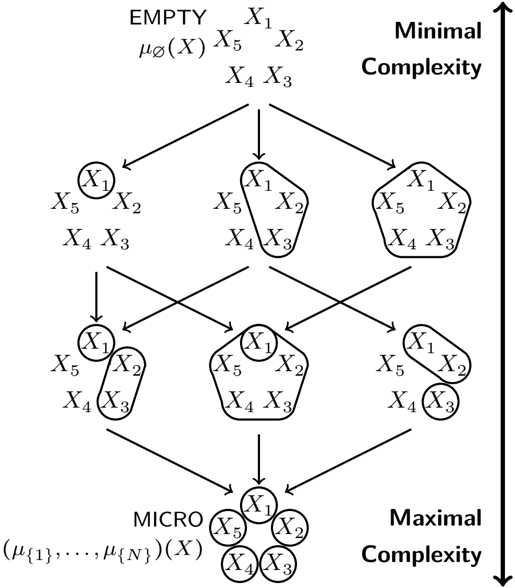
The Poset of Feasible Measurements



The Poset of Feasible Measurements



The Poset of Feasible Measurements



Definition 1 (Additivity)

μ additive iff $\forall A \cap B = \varnothing$,

$$H(\mu_{A\cup B}(X) \mid \mu_A(X), \mu_B(X)) = 0$$
$$H(\mu_A(X) \mid \mu_{A\cup B}(X), \mu_B(X)) = 0$$

Definition 2 (Refinement)

$\phi_1 \prec \phi_2$

iff $X \to \phi_1(X) \to \phi_2(X)$ is Markovian

iff $I(X; \phi_2(X) | \phi_1(X)) = 0$

Theorem 1 (Monotonicity)

$\phi_1 \prec \phi_2 \;\Rightarrow$

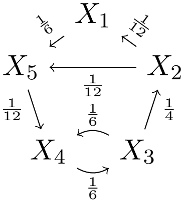
$I(\phi_1(X^t); \psi(X^{t+\tau})) \geq I(\phi_2(X^t); \psi(X^{t+\tau}))$

and $I(X^t; \phi_1(X^t)) \geq I(X^t; \phi_2(X^t))$

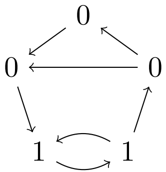
The Voter Model

- Set of Agents $\Omega = \{1, \ldots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \ldots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X^{t+1} | X^t)$ given by an interaction graph:

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j,$ $X_k^{t+1} = X_k^t$



$t = 0$

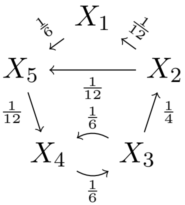


[Kimura & Weiss, 1964] [Banisch & Lima, 2012]

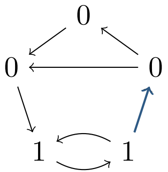
The Voter Model

- Set of Agents $\Omega = \{1, \ldots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \ldots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X^{t+1} | X^t)$ given by an interaction graph:

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j,$ $X_k^{t+1} = X_k^t$



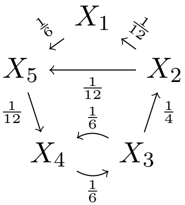
$t = 0$ $\rightarrow$ arc $(3, 2)$



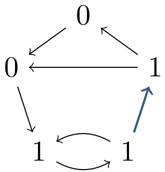
[Kimura & Weiss, 1964] [Banisch & Lima, 2012]

- Set of Agents $\Omega = \{1, \ldots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \ldots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X^{t+1}|X^t)$ given by an interaction graph:

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j,$ $X_k^{t+1} = X_k^t$



$t = 1$

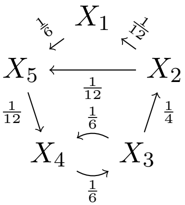


[Kimura & Weiss, 1964] [Banisch & Lima, 2012]

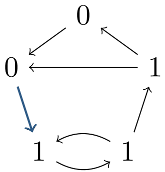
The Voter Model

- Set of Agents $\Omega = \{1, \ldots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \ldots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X^{t+1} | X^t)$ given by an interaction graph:

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j,$ $X_k^{t+1} = X_k^t$



$t = 1$ $\rightarrow$ arc $(5, 4)$

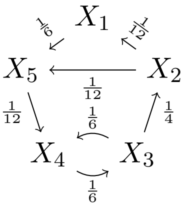


[Kimura & Weiss, 1964] [Banisch & Lima, 2012]

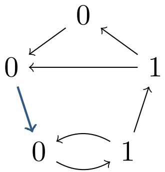
The Voter Model

- Set of Agents $\Omega = \{1, \dots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \dots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X^{t+1}|X^t)$ given by an interaction graph:

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j,$ $X_k^{t+1} = X_k^t$



$t = 2$

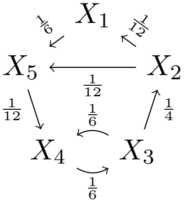


[Kimura & Weiss, 1964] [Banisch & Lima, 2012]

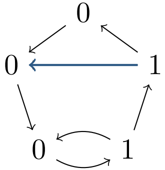
The Voter Model

- Set of Agents $\Omega = \{1, \ldots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \ldots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X^{t+1}|X^t)$ given by an interaction graph:

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j,$ $X_k^{t+1} = X_k^t$



$t = 2$ $\rightarrow$ arc $(2, 5)$

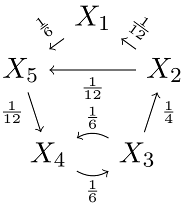


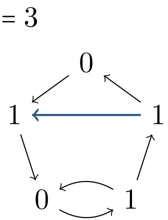
[Kimura & Weiss, 1964] [Banisch & Lima, 2012]

The Voter Model

- Set of Agents $\Omega = \{1, \ldots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \ldots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X^{t+1}|X^t)$ given by an interaction graph:

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j,$ $X_k^{t+1} = X_k^t$



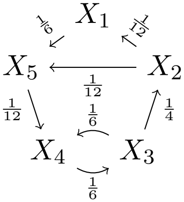


[Kimura & Weiss, 1964] [Banisch & Lima, 2012]

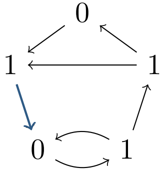
The Voter Model

- Set of Agents $\Omega = \{1, \ldots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \ldots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X^{t+1}|X^t)$ given by an interaction graph:

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j,$ $X_k^{t+1} = X_k^t$



$t = 3$ $\rightarrow$ arc $(5, 4)$

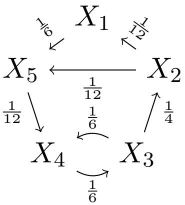


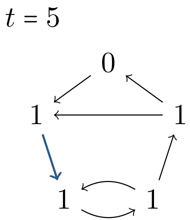
[Kimura & Weiss, 1964] [Banisch & Lima, 2012]

The Voter Model

- Set of Agents $\Omega = \{1, \ldots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \ldots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X^{t+1}|X^t)$ given by an interaction graph:

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j,$ $X_k^{t+1} = X_k^t$



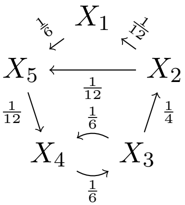


[Kimura & Weiss, 1964] [Banisch & Lima, 2012]

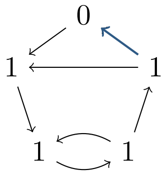
The Voter Model

- Set of Agents $\Omega = \{1, \ldots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \ldots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X^{t+1} | X^t)$ given by an interaction graph:

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j,$ $X_k^{t+1} = X_k^t$



$t = 5$ $\rightarrow$ arc $(2, 1)$

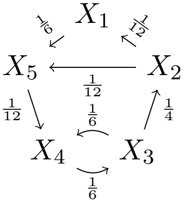


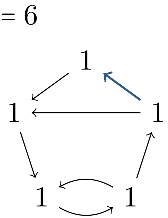
[Kimura & Weiss, 1964] [Banisch & Lima, 2012]

The Voter Model

- Set of Agents $\Omega = \{1, \ldots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \ldots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X^{t+1} | X^t)$ given by an interaction graph:

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j, \quad X_k^{t+1} = X_k^t$



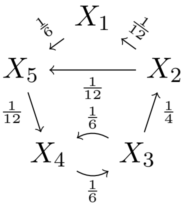


[Kimura & Weiss, 1964] [Banisch & Lima, 2012]

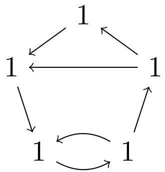
The Voter Model

- Set of Agents $\Omega = \{1, \ldots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \ldots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X^{t+1} | X^t)$ given by an interaction graph:

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j$, $X_k^{t+1} = X_k^t$



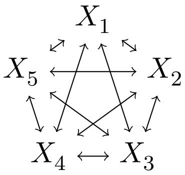
$t = 6$



[Kimura & Weiss, 1964] [Banisch & Lima, 2012]

Aggregated-states in the Complete Graph

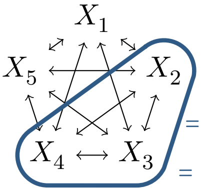
- All arcs are equally likely $\forall (i,j) \in \Omega^2, \quad \Pr(\text{arc } (i,j)) = \frac{1}{N\,(N-1)}$
- Uniform Initial State $\forall x \in \{0,1\}^N, \quad p(X^0 = x) = 2^{-N}$



Aggregated-states in the Complete Graph

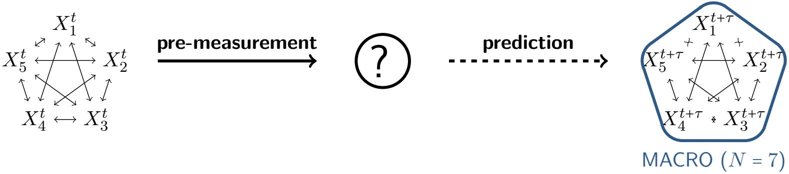
- All arcs are equally likely $\forall (i,j) \in \Omega^2, \quad \Pr(\text{arc } (i,j)) = \frac{1}{N\,(N-1)}$
- Uniform Initial State $\forall x \in \{0,1\}^N, \quad p(X^0 = x) = 2^{-N}$

- Generic Measurement $\forall A \subset \Omega, \quad \eta_A(x) = \sum_{i \in A} x_i$
  (Aggregated-state Measurement)

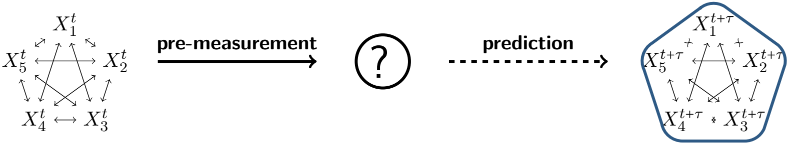


$= \mu_{\{2,3,4\}}(X)$
$= X_2 + X_3 + X_4$

Predicting the Macroscopic Measurement

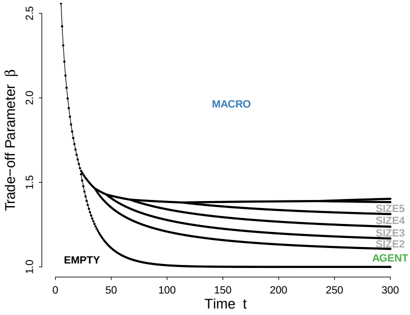


Predicting the Macroscopic Measurement

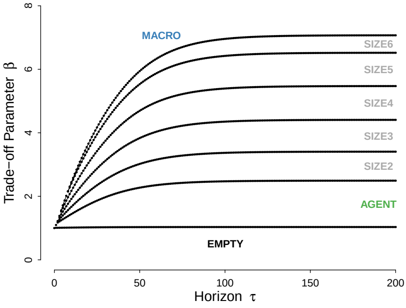


MACRO ($N = 7$)

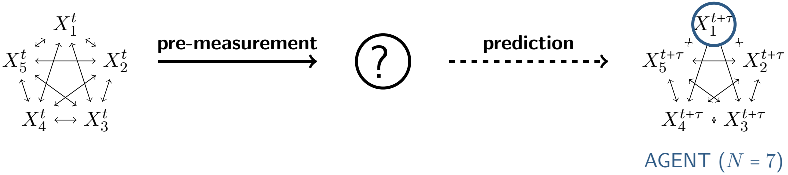
Fixed horizon $\tau = 3$



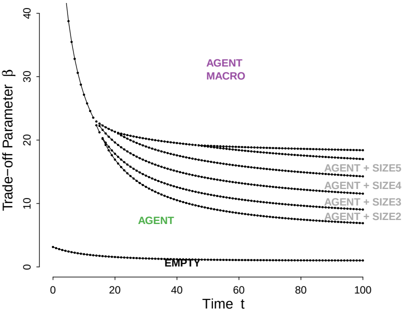
Fixed time $t = 100$



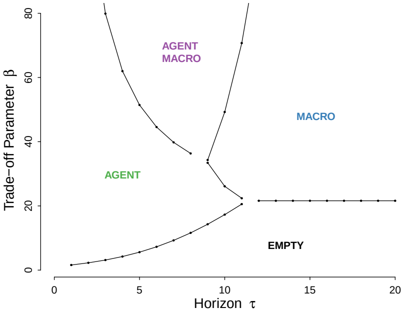
Predicting the Agent Measurement



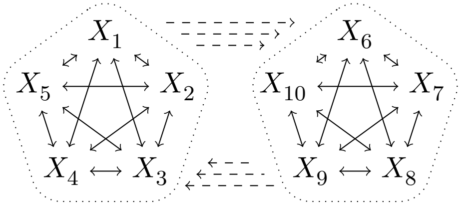
Fixed horizon $\tau = 3$



Fixed time $t = 0$



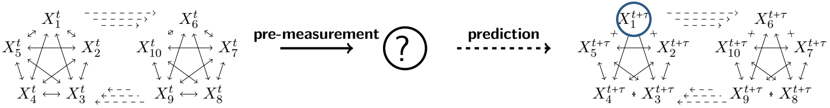
The Two-community Graph



First Community Ω_1 Second Community Ω_2

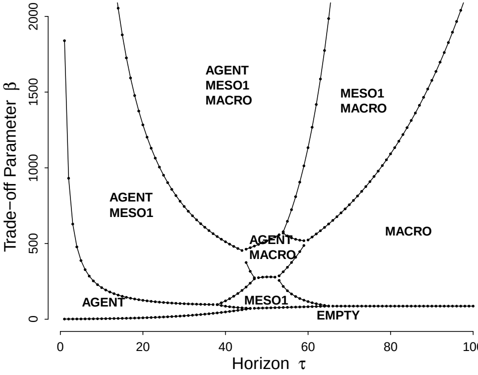
Coupling Parameter $\rho = \frac{\mathrm{Pr}(\text{inter edge})}{\mathrm{Pr}(\text{intra edge})} < 1$

Predicting the Agent Measurement



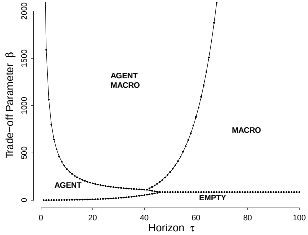
AGENT ($N = 10 + 10$ and $\rho = \frac{1}{5}$)

Fixed time $t = 0$

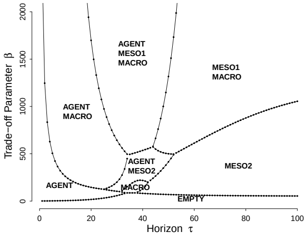


Some Other Heterogeneous Interaction Graphs

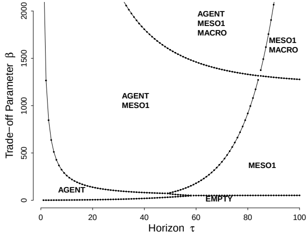
Homogeneous Case: $\rho_{1\to2} = \rho_{2\to1} = 1$



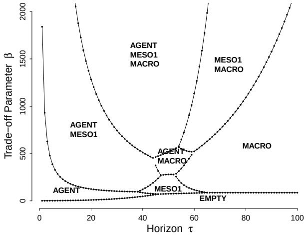
Follower Case: $\rho_{1\to2} = 1/5$ and $\rho_{2\to1} = 1$



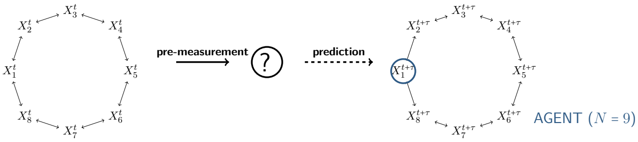
Leader Case: $\rho_{1\to2} = 1$ and $\rho_{2\to1} = 1/5$



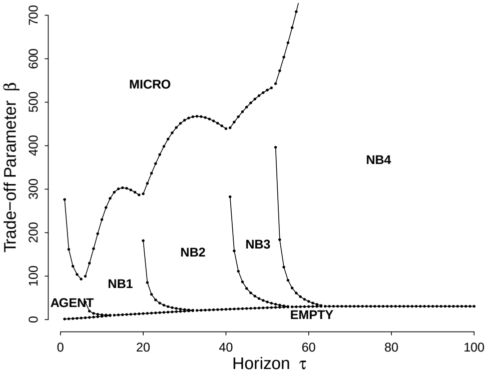
Symmetrical Case: $\rho_{1\to2} = \rho_{2\to1} = 1/5$



Predicting the Agent Measurement in the Ring



Fixed time $t = 0$



Application Perspectives

- Application Perspectives
  - Optimal prediction of **economic indicators** based on multilevel structures of the international trade network
  - Optimal prediction of **population dynamics** in ecology based on multilevel representation of the inter-species mutual dependencies
- Theoretical Perspectives
  - Taking into account the **problem of inferability** of low-level predictors when solving the optimisation problem (*e.g.*, Bayesian model selection)
  - Taking into account **domain-depend costs and rewards** as practical objective functions to be optimised in real-world scenarios

Thank you for your attention

Mail: Eckehard.Olbrich@mis.mpg.de

Web: http://www.mpipks-dresden.mpg.de/ olbrich/