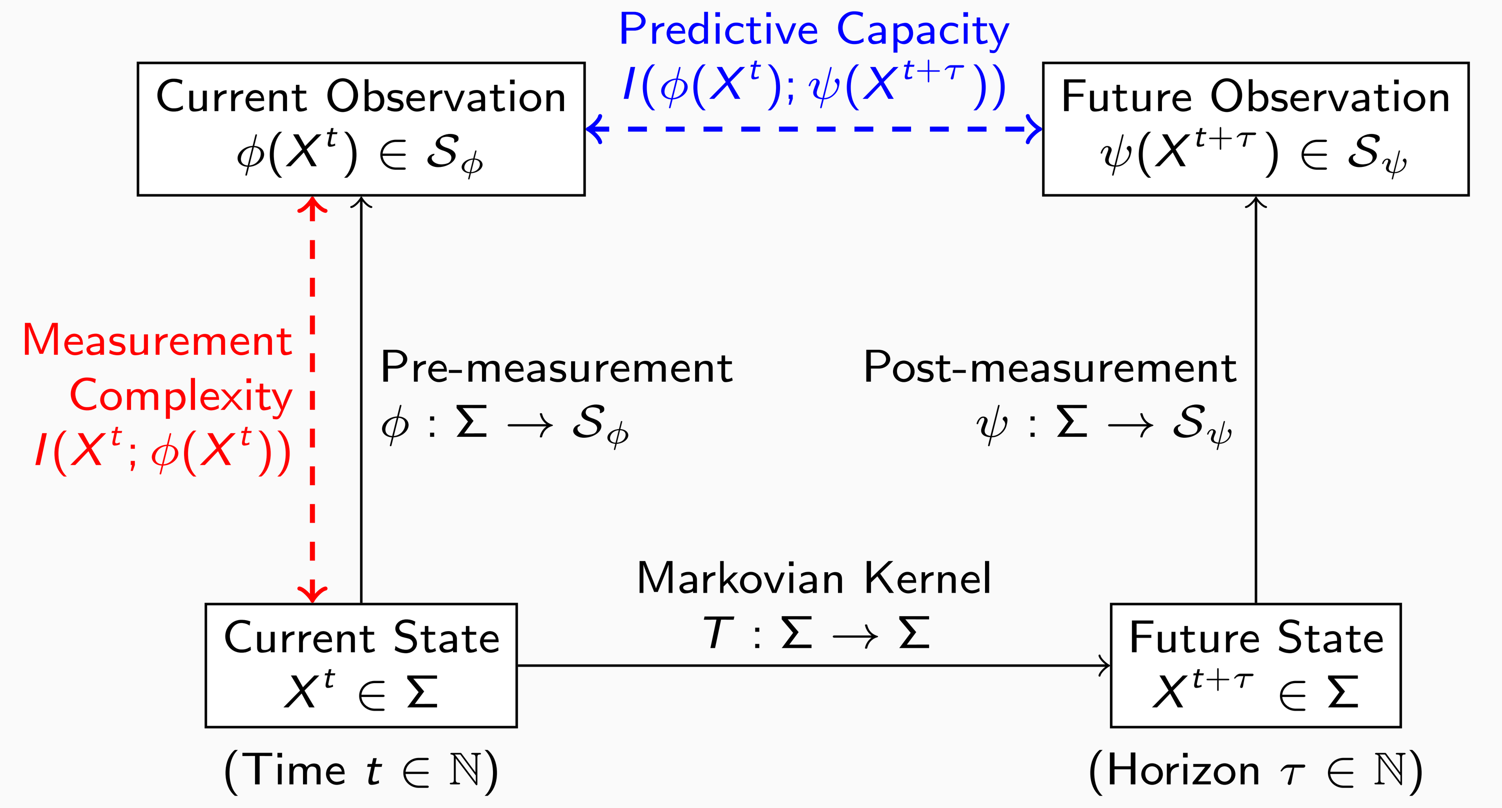
The Information Bottleneck Method for Optimal Prediction of Multilevel Agent-based Systems

Robin Lamarche-Perrin, Sven Banisch, and Eckehard Olbrich
{FirstName.LastName}@mis.mpg.de

Max Planck Institute for Mathematics in the Sciences
mathe MACS — Mathematics for Multilevel Anticipatory Complex Systems

1. General Setting: Information Bottleneck for Optimal Prediction



The Optimal Prediction Problem

Given:

- an initial state $X^0\in\Sigma$
- a Markovian kernel $T:\Sigma\to\Sigma$
- a post-measurement $\psi:\Sigma\to\mathcal{S}_\psi$
- a current time $t\in\mathbb{N}$
- an horizon prediction $\tau\in\mathbb{N}$
- a trade-off parameter $\beta\in\mathbb{R}^+$

Find:

- a pre-measurement $\phi:\Sigma\to\mathcal{S}_\phi$

that minimises the Information Bottleneck variational **[Tishby, 1999]**:

$$I(X^t;\phi(X^t)) \;-\; \beta\; I(\phi(X^t);\psi(X^{t+\tau}))$$

2. Observing Agent-based Systems with Additive Measurements

- Agent States $X_1^t\in\mathcal{S},\ X_2^t\in\mathcal{S},\ \dots,\ X_N^t\in\mathcal{S}$
- System State $X^t=(X_1^t,X_2^t,\dots,X_N^t)\ \in\ \Sigma=\mathcal{S}^N$

Definition: Additive Measurement

A family $(\mu_A:\Sigma\to\mathcal{S}_\mu)_{A\in\{1,\dots,N\}}$ of measurements parametrized by an agent set $A\subset\{1,\dots,N\}$ such that:

- $\mu_A(X)$ only depends on the state of agents in A: $\Pr(\mu_A(X)\mid X) \;=\; \Pr(\mu_A(X)\mid(X_i)_{i\in A})$
- from the observation of two disjoint agent sets, one can deduce the observation of their union:

$$A_1\cap A_2=\emptyset\ \Rightarrow\ \begin{cases}H(\mu_{A_1\cup A_2}(X)\mid\mu_{A_1}(X),\mu_{A_2}(X))=0\\ H(\mu_{A_1}(X)\mid\mu_{A_1\cup A_2}(X),\mu_{A_2}(X))=0\end{cases}$$

Properties of the Measurement Poset

- By combining several measurements, one refines the observation space: $(\mu_{A_1},\mu_{A_2})\prec\mu_{A_1}$
- The MICRO measurement refines any other measurement: $(\mu_{\{1\}},\dots,\mu_{\{N\}})\ \prec\ (\mu_{A_1},\dots,\mu_{A_N})$
- The EMPTY measurement is refined by any other measurement: $(\mu_{A_1},\dots,\mu_{A_N})\ \prec\ \mu_\emptyset$
- Observing two nested agent sets is equivalent to observing the smaller set and its complement:

$$A_1\subset A_2\ \Rightarrow\ \begin{cases}(\mu_{A_1},\mu_{A_2})\ \prec\ (\mu_{A_1},\mu_{A_2\setminus A_1})\\ (\mu_{A_1},\mu_{A_2})\ \succ\ (\mu_{A_1},\mu_{A_2\setminus A_1})\end{cases}$$

3. The Measurement Poset

Definition: Refinement Relation (partial order)

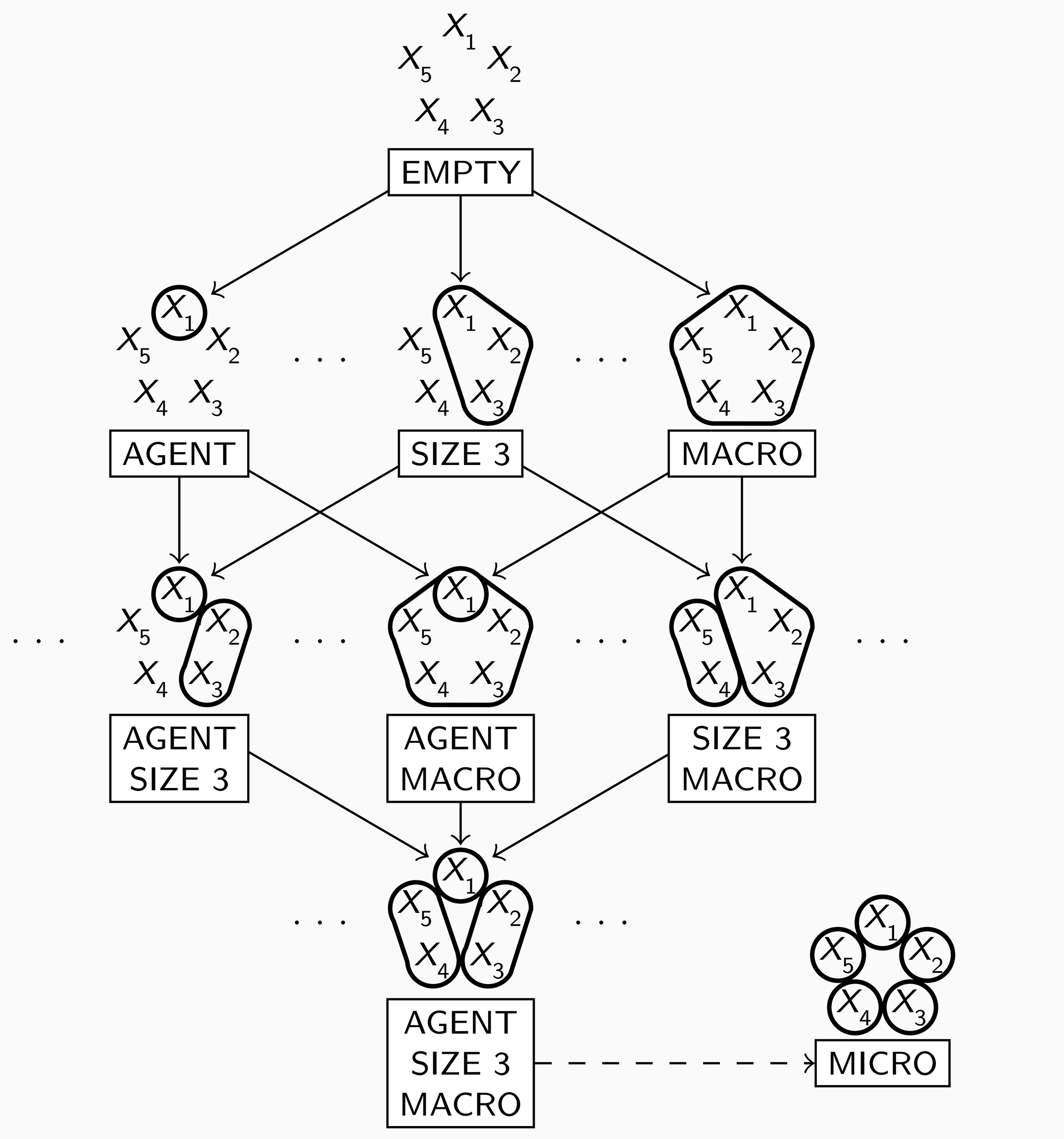
A measurement ϕ_1 refines a measurement ϕ_2 if and only if $\phi_1(X)$ contains all the micro-information about $\phi_2(X)$:

$$\phi_1\prec\phi_2\quad\Leftrightarrow\quad I(X\,;\,\phi_2(X)\mid\phi_1(X))\;=\;0$$
$$\Leftrightarrow\quad X\ \to\ \phi_1(X)\ \to\ \phi_2(X)\ \text{ is Markovian}$$

Theorem: Monotonicity of IB-measures

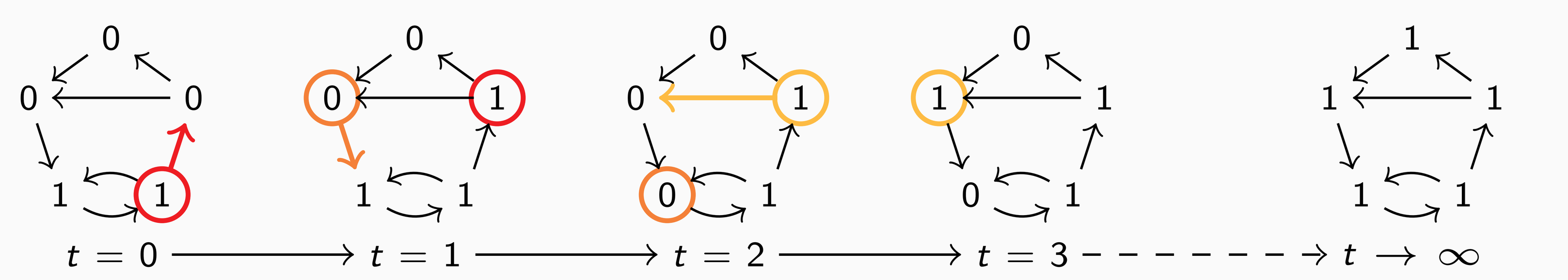
Refinements are always more complex and predictive:

$$\phi_1\prec\phi_2\ \Rightarrow\ I(X^t;\phi_1(X^t))\ge I(X^t;\phi_2(X^t))$$
$$\text{and}\quad I(\phi_1(X^t);\psi(X^{t+\tau}))\ge I(\phi_2(X^t);\psi(X^{t+\tau}))$$

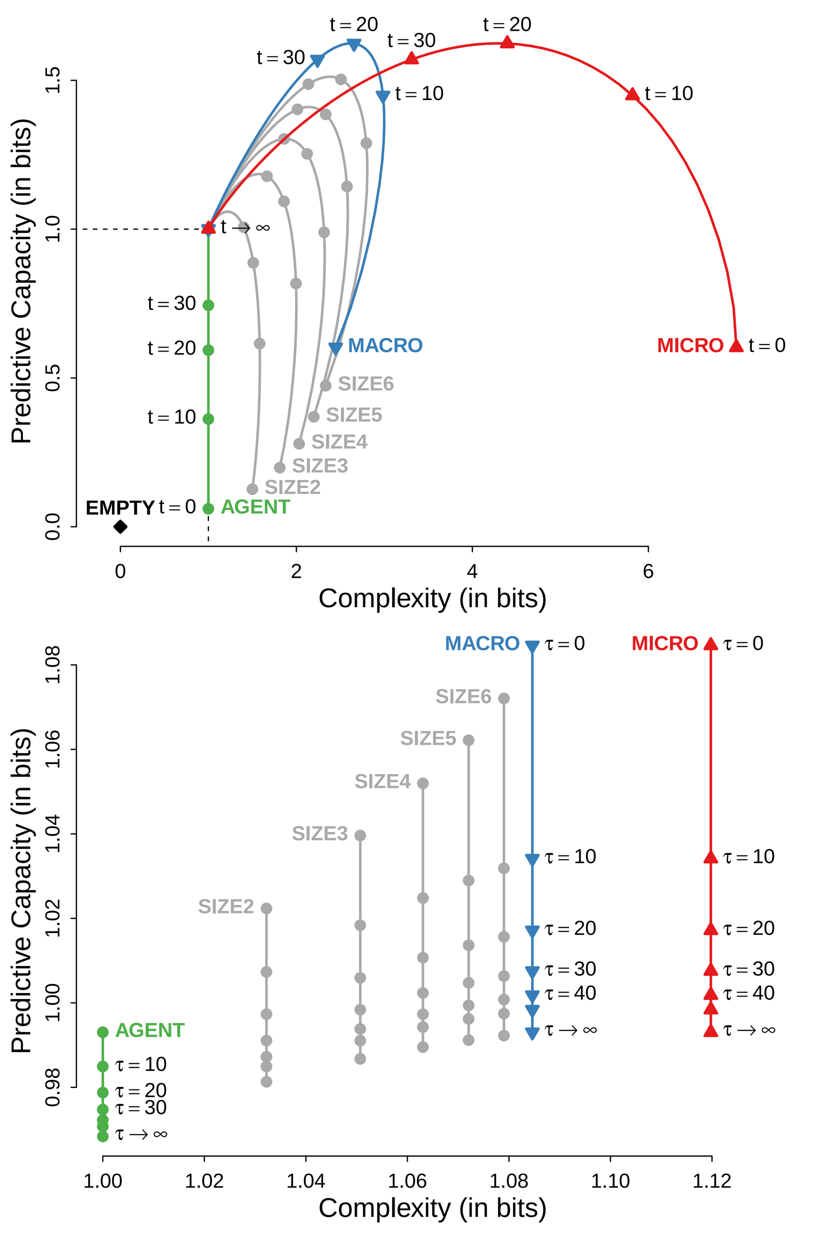


4. Application to the Voter Model: Predicting Synchronisation Processes in Interaction Graphs

- Agent States $X_1^t\in\{0,1\},\ X_2^t\in\{0,1\},\ \dots,X_N^t\in\{0,1\}$
- System State $X^t=(X_1^t,X_2^t,\dots,X_N^t)\ \in\ \Sigma=\{0,1\}^N$
- Transitions Kernel $T(X^{t+1}|X^t)$ determined by an interaction graph:
  arc (i,j) selected $\ \Rightarrow\ X_j^{t+1}=X_i^t$ and $\forall k\neq j,\ \ X_k^{t+1}=X_k^t$
- Additive Measurements $\forall A\subset\{1,\dots,N\},\quad \mu_A(X^t)=\sum_{i\in A}X_i^t$



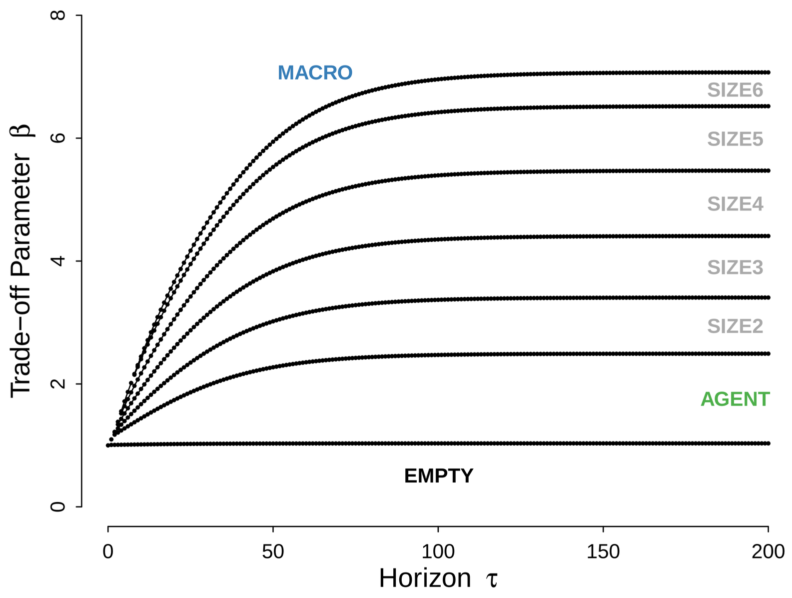
Predicting MACRO in the Complete Graph ($N=7$)



Results:

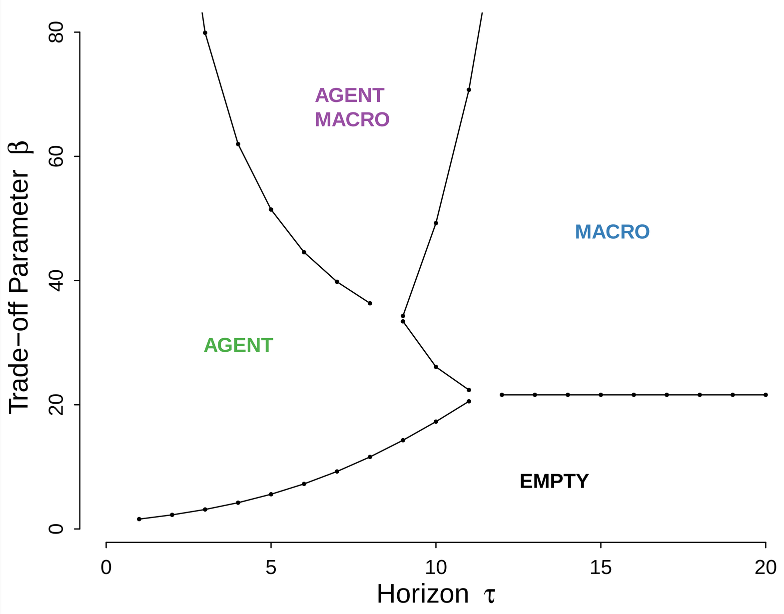
- The predictive capacity of all measurements decreases as the horizon increases
- MICRO is more complex than MACRO
- MACRO is as predictive as MICRO
- Estimating the current state by sampling might provide more efficient prediction

Optimal Pre-measurements



Predicting AGENT in the Complete Graph ($N=7$)

Optimal Pre-measurements

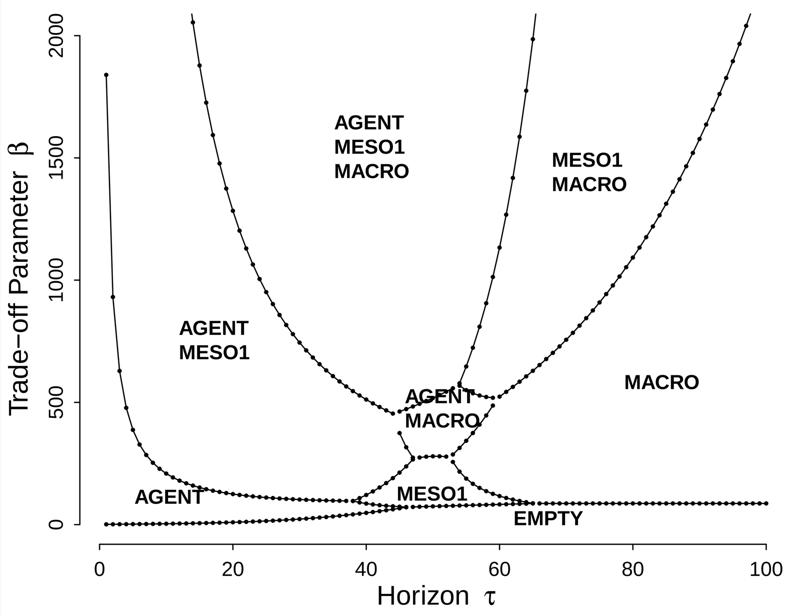


Results:

- Local dynamics: AGENT is more efficient for short-term prediction
- Global dynamics: MACRO is more efficient for long-term prediction
- Multilevel dynamics: AGENT + MACRO is more efficient for middle-term prediction when a higher complexity level is allowed

Predicting AGENT in Two Communities ($N=10+10$)

Optimal Pre-measurements

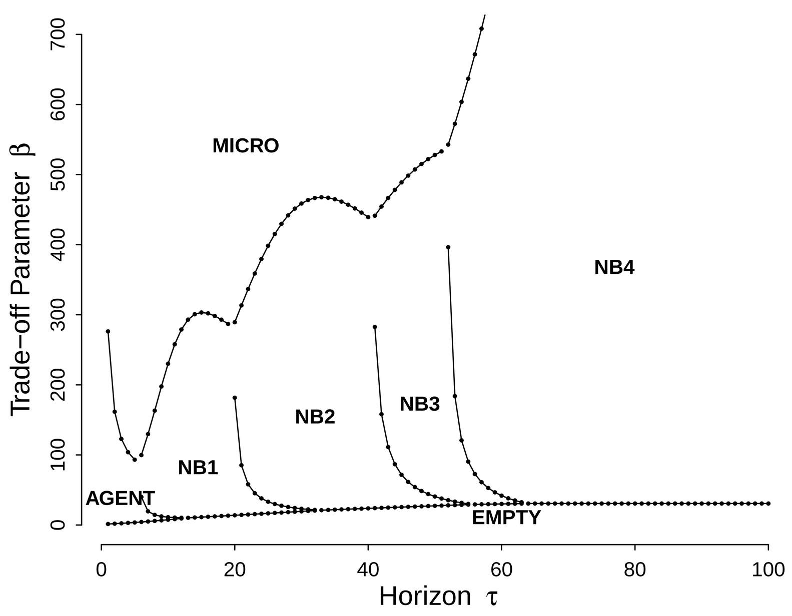


Results:

- From local to global dynamics: MICRO, MESO, and MACRO are respectively more efficient for short, middle, and long-term prediction
- Multilevel dynamics: These three measurements can be combined for higher predictive power depending on the horizon (including the three-level measurement MICRO + MESO + MACRO)

Predicting AGENT in the Ring ($N=9$)

Optimal Pre-measurements



Results:

- The size of the neighbourhood to be observed depends on the prediction horizon (the higher the horizon is, the larger the observed neighbourhood should be)
- For a given prediction horizon, their is an optimal neighbourhood size that does not depend on the allowed complexity level

Bibliography

Robin Lamarche-Perrin, Sven Banisch, and Eckehard Olbrich. The Information Bottleneck Method for Optimal Prediction of Multilevel Agent-based Systems. In *Advances in Complex Systems*, 2016.
Preprint: http://www.mis.mpg.de/publications/preprints/2015/prepr2015-55.html