Fully Decentralized Policies for Multi-Agent Systems: An Information Theoretic Approach

Roel Dobbe*, David Fridovich-Keil*, and Claire J. Tomlin
UC Berkeley Hybrid Systems Lab - {dobbe,dfk}@eecs.berkeley.edu

Decentralized Optimal Control

- Many critical infrastructures, such as electric grids, or transportation networks, require *decentralized policies* to efficiently implement control schemes.
- Control is often hampered by partial observability and lack of coordination/communication.
- Rather than solving a central optimization in real-time, treat agent i's optimal action problem as a random variable u_i^* that can be fully decentralized
- u_i^* depends on the global state variables $x = (x_1, \ldots, x_n)$, i.e. we consider $p(u_i^*|x)$.
- View decentralization as a *compression* problem, by letting a policy only depend on x_i .
- Apply classical results from information theory to analyze performance limits based on fundamental limits of compression, which is well formulated as an instance of *rate distortion theory*.

Rate Distortion Theory

Consider X an information source and Y a reconstruction after transmission over a noisy channel.

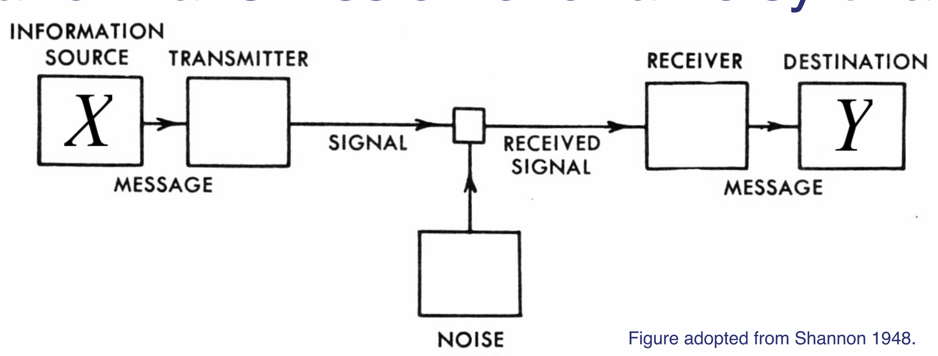
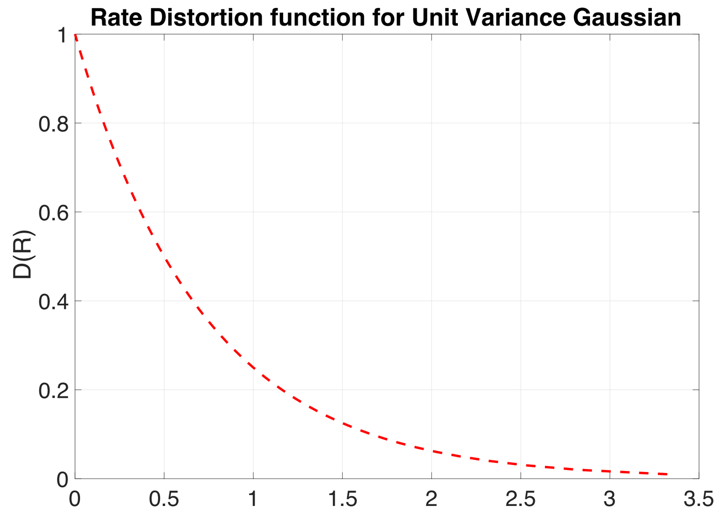


Figure adopted from Shannon 1948.

If mutual information $I(X;Y) \leq R$ then for distortion function $d(x,y)$ the minimum average distortion $D(R)$ solves a convex program.

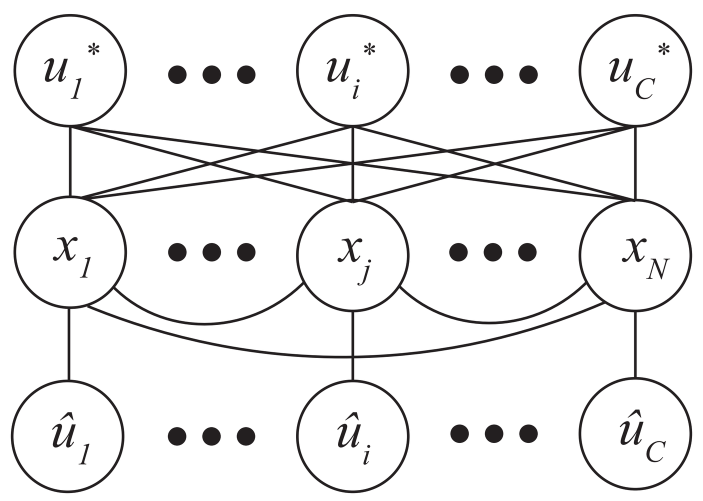
$$D(R) = \min_{p(X|Y)} \mathbb{E}[d(X,Y)], \quad \text{s.t. } I(X;Y) \leq R$$



R is called the "rate" of information transfer, and $D(R)$ gives a lower bound on average distortion.

Graphical Model Representation

Each agent's optimal control u_i^* depends on all state variables x but the control applied in practice can only depend on local information x_i.



Data Processing Inequality

Suppose $X - Y - Z$ form a Markov chain. Then the following constraint holds:

$$I(X;Z) \leq I(X;Y)$$

In decentralized optimal control, the DPI applies between controls and state variables:

$$I(\hat{u}_i; u_j^*) \leq I(x_i; u_j^*) \text{ and } I(\hat{u}_i; \hat{u}_j) \leq I(x_i; x_j)$$

Application of Rate Distortion

In distributed control, rate R is effectively specified by the information contained in local state variables. We formulate the problem as:

$$\begin{aligned} D^* = \quad & \min_{p(\hat{u}|u^*)} && \mathbb{E}[d(u^*, \hat{u})] \\ & \text{s.t.} && I(\hat{u}_i; u_j^*) \leq I(x_i; u_j^*) \\ & && I(\hat{u}_i; \hat{u}_j) \leq I(x_i; x_j) \end{aligned}$$

- Constraints depend upon the problem structure.
- Mutual information must be estimated from data.
- Problem is convex in function space, but not necessarily convex for particular parameterizations.
- Solution is non-prescriptive, i.e. does not specify precisely how to perform minimum-distortion decoding (control).

For the Gaussian case with squared error distortion, there is a closed form solution for D^* and a corresponding optimal decoder:

$$D^* = \sum_i \sigma_{u_i^*}^2 \left(1 - \rho_{u_i^* x_i}^2\right) \text{ and } \hat{u}_i = \mathbb{E}[u_i^*] + \frac{\rho_{u_i^* x_i} \sigma_{u_i^*}}{\sigma_{x_i}} (x_i - \mathbb{E}[x_i])$$

Communication Strategy

Suppose controller i is allowed to observe k other state variables $\mathcal{S}_i$.

Theorem

Setting $\mathcal{S}_i$ as follows minimizes D^* for any distortion d:

$$\mathcal{S}_i = \arg\max_{\mathcal{S}} I(u_i^*; x_i, \{x_j : j \in \mathcal{S}\}) : |\mathcal{S}| = k$$

The proof follows from the monotonicity of $D(R)$.

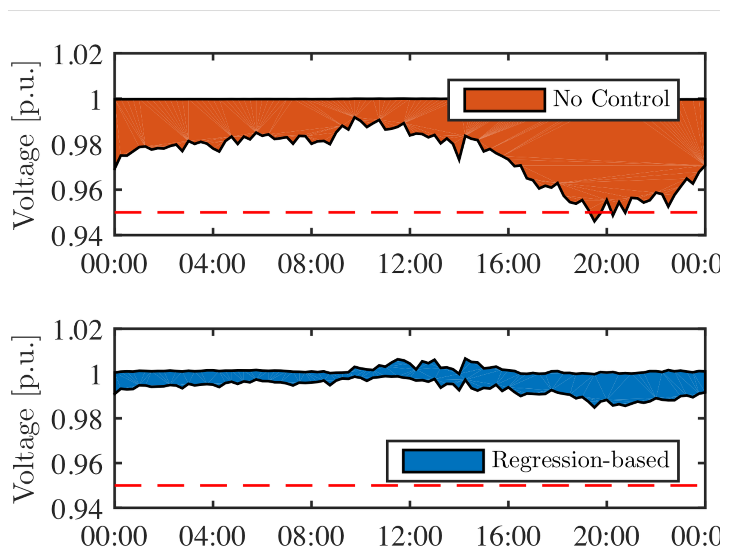
Case: Optimal Power Flow (OPF)

- We address voltage variability in an electric grid $\mathcal{G} := (\mathcal{N}, \mathcal{E})$ due to intermittent renewable energy.
- Optimize reactive power outputs q_i^g , $i \in \mathcal{C}$ of a set of distributed energy resources (DERs), $\mathcal{C}$.
- Adhere to physics of power flow and constraints due to energy capacity and safety.

$$\begin{aligned} u^* = \arg \min_{q_i^g,\ \forall i \in \mathcal{C}} \quad & \sum_{i \in \mathcal{N}} |v_i - v_{\text{ref}}| \,, \\ \text{s.t.} \quad & P_{ij} = \sum_{(j,k) \in \mathcal{E}, k \neq i} P_{jk} + p_j^c - p_j^g \,, \\ & Q_{ij} = \sum_{(j,k) \in \mathcal{E}, k \neq i} Q_{jk} + q_j^c - q_j^g \,, \\ & v_j = v_i - 2\left(r_{ij} P_{ij} + \xi_{ij} Q_{ij}\right) , \\ & |q_i^g| \leq \bar{q}_i \ , \ \underline{v} \leq v_i \leq \bar{v} \,. \end{aligned}$$

- Collect data set of offline simulations and optima u_i^*.
- Generalize u_i^* for all DER using an interpretable stepwise regression model $\hat{\pi}_i(x_i) = \theta_i^\top \phi_i(x_i)$.
- Use communication strategy to improve performance with additional variables $\mathcal{S}_i$.

Voltage Variability Result



Communication Strategies