Outline:

Pulkit Grover

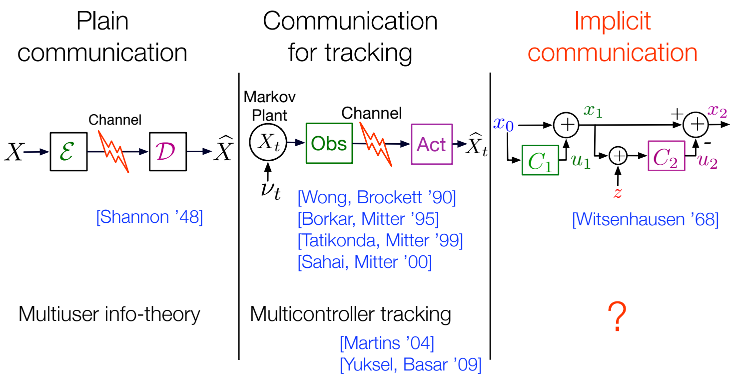
"Synergistic Implicit and Explicit Communication in Decentralized Control"

48th Annual Allerton Conference, Sept. 29, 2010.

Joint work with Anant Sahai.

Slides: www.eecs.berkeley.edu/~pulkit/files/Allerton10Talk.pdf
This Handout: …/~pulkit/files/Allerton10.H.pdf
Further discussions/references : www.eecs.berkeley.edu/~pulkit

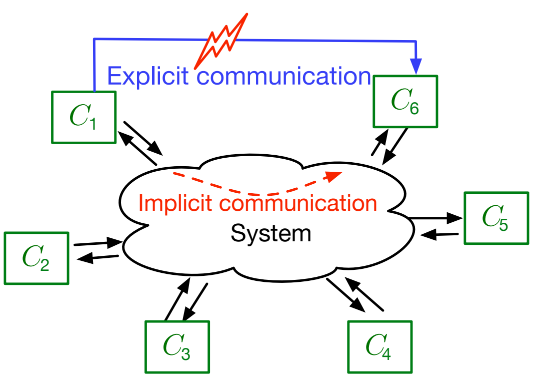
1 Coordination via communication



Coordination is of central importance in decentralized control. Communication is a natural way to create coordination.

Toy problems here that have been studied can be classified into three types: plain communication, communication for tracking an unstable source, and implicit communication. Of the three, implicit communication is the least understood.

1.1 Implicit and communication

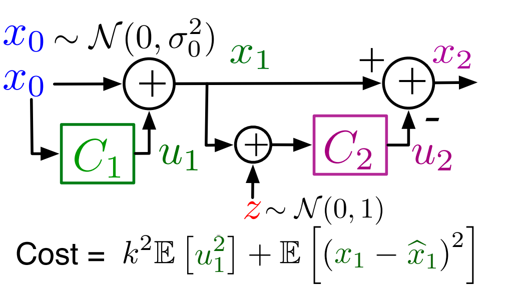


We define implicit communication as one of two possibilities:

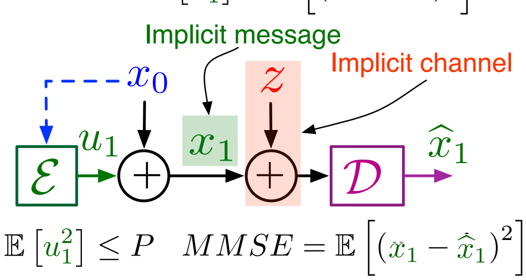
- Implicit Message: the message is generated endogenously by the system.
- Implicit Channel: the system state itself is used to send messages.

One often has the engineering freedom to attach an explicit communication channel between controllers. In presence of such a channel, is implicit communication at all useful? What are the good strategies for communicating implicitly *and* explicitly? These are the questions we are interested in.

2 A toy problem of implicit communication: Witsenhausen's counterexample

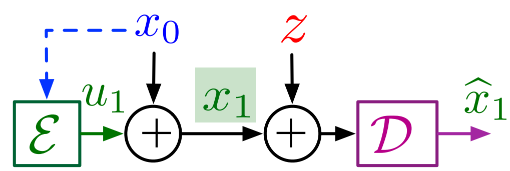


The Witsenhausen counterexample [Witsenhausen'68] is a deceptively simple two-controller two-time-step problem of maneuvering the state of a system to the origin. Costs are imposed on the input u_1 of the first controller, and the state x_2 after action of the two controllers. Since there is no cost on the input u_2 of the second controller, it is best to choose $u_2 = \widehat{x}_1$, the MMSE estimate of x_1.



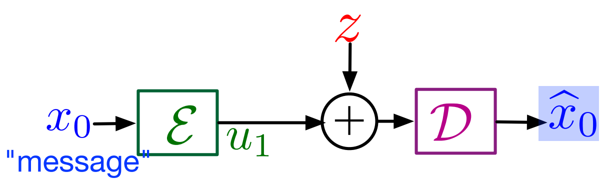
The counterexample is the simplest example of a problem that displays both implicit message and implicit channel. The communication interpretation of the problem, where the controllers are interpreted as an *encoder* and a *decoder*, brings this out. The goal is to communicate the implicit message x_1 to the decoder through the implicit channel, while using minimum power of the encoder input u_1.

3 Implicit versus explicit communication



$$k^2\mathbb{E}\left[u_1^2\right] + \mathbb{E}\left[\left(x_1 - \widehat{x}_1\right)^2\right]$$

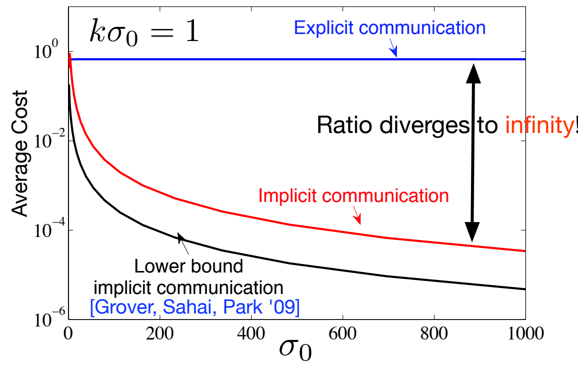
Implicit communciation



$$k^2\mathbb{E}\left[u_1^2\right] + \mathbb{E}\left[\left(x_0 - \widehat{x}_0\right)^2\right]$$

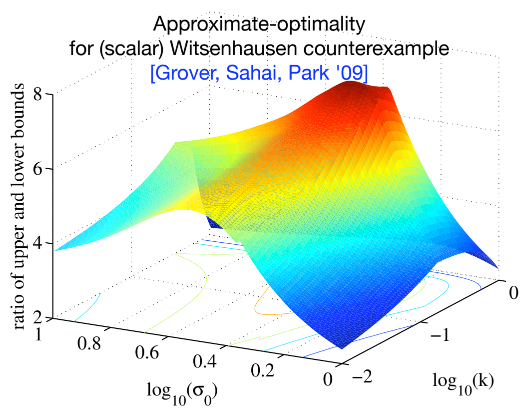
Explicit communciation

The first problem is Witsenhausen's counterexample — purely implicit communication. The second is Shannon's point-to-point explicit communication problem. Which one has lower costs?

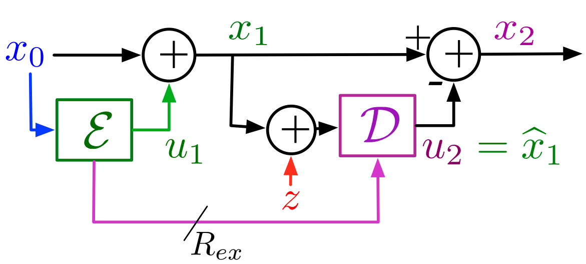


It turns out that quantization-based implicit communication can outperform explicit-communication by an arbitrarily large factor. However, the optimal cost using implicit communication in the first problem is still unknown. Our results [Grover, Sahai, Park '09] provide the optimal costs within a constant factor of 8 for all problem parameters.

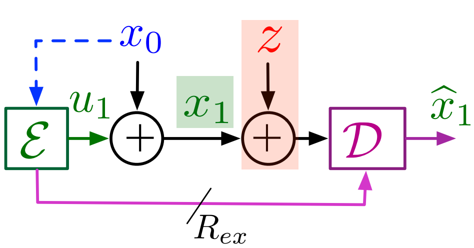
What if both implicit and explicit communication are possible in a problem — can we build on this understanding?



3.1 A toy problem of implicit and explicit communication



Cost = $k^2\mathbb{E}\left[u_1^2\right] + \mathbb{E}\left[\left(x_1 - \widehat{x}_1\right)^2\right]$

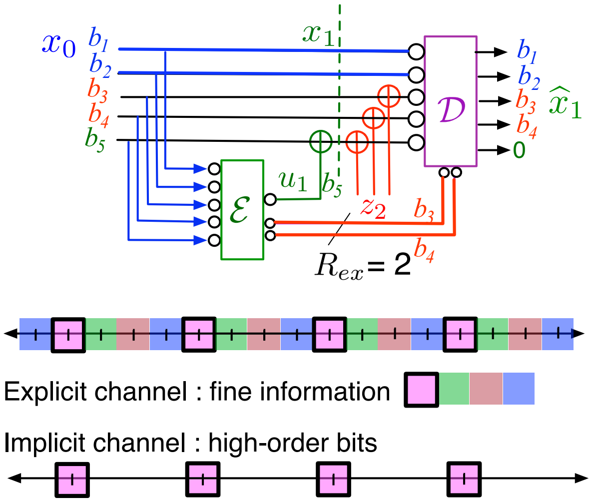


The problem is an extension of Witsenhausen's counterexample where a rate-limited external channel connects the two controllers. Similar extensions (with Gaussian external channel) were considered in [Shoarinejad '02] and [Martins '06].

A natural strategy:
When there is no possibility of implicit communication, the optimal strategy is to communicate the state [Goblick '65]. If the external channel is not present, quantization is approximately optimal [Grover, Sahai, Park '09].

Therefore, a seemingly natural strategy (used in [Martins '06]) is to use quantization for implicit communication, and communicate the state on the external channel. Can we do better?

3.2 A deterministic abstraction, and resulting strategies

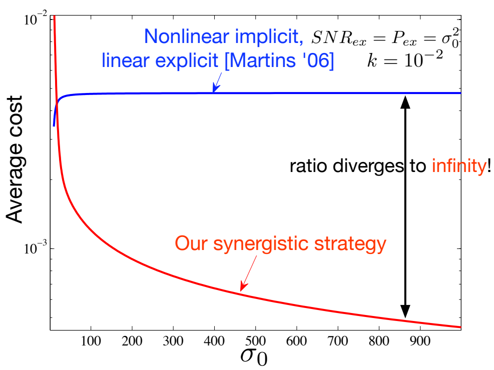


A deterministic abstraction of the problem, inspired by deterministic models in information theory [Avestimehr, Diggavi and Tse '08].

Bits b_1, b_2 are received noiselessly. Since noise corrupts the lower order bits, the external channel can be used to send b_3, b_4. Since the capacity of the external channel is now exhausted, the encoder can use its input to force b_5 to zero. The decoder now has a perfect estimate of x_1.

The scheme is known as a binning strategy in information theory. Fine information, one of four bin-colors, is sent over the external channel. Implicit communication is used to send the higher order bits and force the state to a fine quantization point.

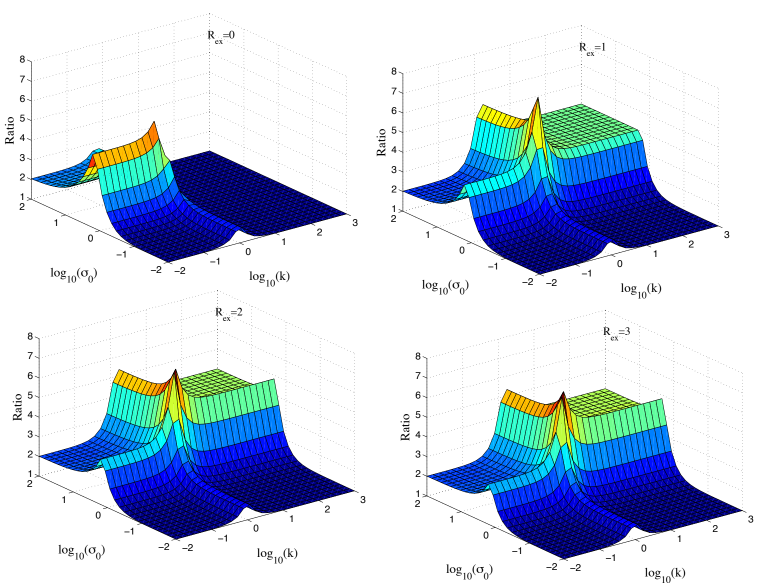
3.3 Binning based strategies are much better!



The external channel in [Martins '06] is an additive Gaussian noise channel with an average power constraint P_{ch} and noise variance 1. For comparison with results in [Martins '06], we plot the performance of our strategy when used over this Gaussian external channel. Bin-size is assumed to be a free parameter in the optimization for both strategies.

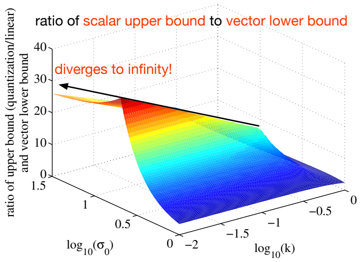
As $\sigma_0^2 \to \infty$, the ratio of costs diverges to infinity as well.

3.4 An approximately optimal solution to the asymptotically infinite-length problem



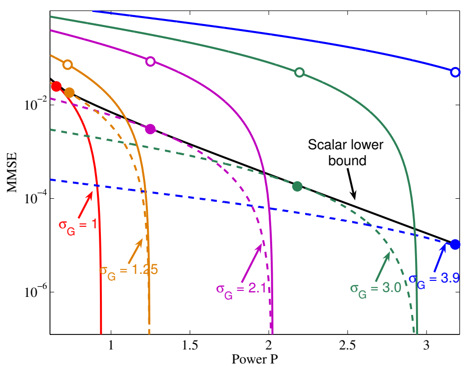
Binning-based strategies attain within a factor of 8 of the optimal cost in the limit of infinite lengths.

3.5 Asymptotic lower-bound is insufficient for the scalar case



The asymptotic lower bound is insufficient to show approximate-optimality for the scalar case. We need tighter lower bounds that take vector length into account.

3.6 A tighter KL-divergence-based "sphere-packing" lower bound for the scalar problem

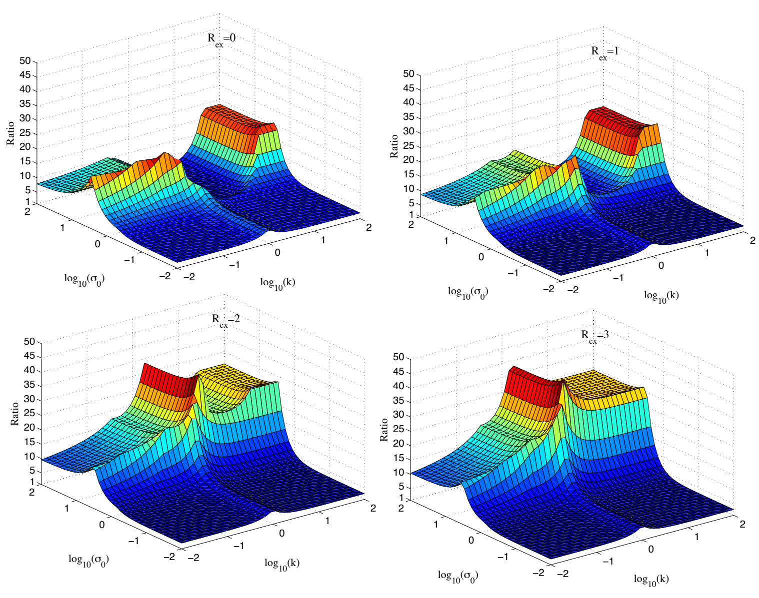


A tighter lower bound that is specific to a given vector length. The derivation is based on that for Witsenhausen's counterexample [Grover, Sahai, Park '09]. It is a large deviations perspective on the "sphere-packing" bounding technique. The observation noise can behave as if its variance is much larger, thereby increasing the lower bound. For finite dimensions, the probability of such atypical behavior is approximately characterized by the KL-divergence between the typical and the atypical distributions.

Intuitively, for a fixed power P, we can find a lower bound on the $MMSE$ assuming a test noise variance of $\sigma_G^2 > 1$. Multiplying this lower bound on MMSE with the probability that the channel noise behaves as Gaussian with variance σ_G^2 provides a lower bound to our problem.

The figure illustrates this intuition in the scalar case for $R_{ex} = 0$.

3.7 Approximate optimality for the scalar problem?



With the sphere-packing bound, binning-based strategies can be shown to attain within a constant factor of the optimal cost for any *fixed* rate R_{ex} on the external channel. However, as $R_{ex} \to \infty$, the ratio diverges to infinity as well.

We believe that tightening of the upper bound (*i.e.* a better achievable strategy) in the regime of large-k, large-σ_0 is required to attain within a constant factor of the optimal cost, and to not have the constant depend on R_{ex}.

4 Summary

This talk intends to bring out the following ideas:

- Implicit communication has the potential to be extremely useful way of building coordination— if substituted with explicit communication, the system performance is much worse.
- Good control strategies use implicit and explicit communication synergistically. Deterministic models can yield valuable insights on how to design such strategies.
- KL-divergence is an indispensable tool for proving approximate-optimality at finite-lengths.