Outline:

Pulkit Grover

"Distributed control: a new frontier of information theory"

ITA Workshop, San Diego, Feb. 3, 2010.

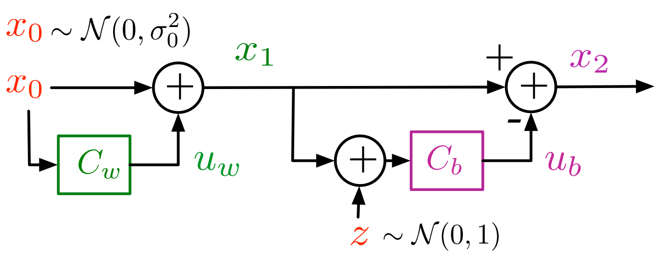
Joint work with Anant Sahai, Se Yong Park, and Aaron Wagner

Slides available: www.eecs.berkeley.edu/∼pulkit/files/ITA10Slides.pdf
This Handout: . . . /∼pulkit/files/ITA10Handout.pdf
Further discussion and references can be found in the papers.

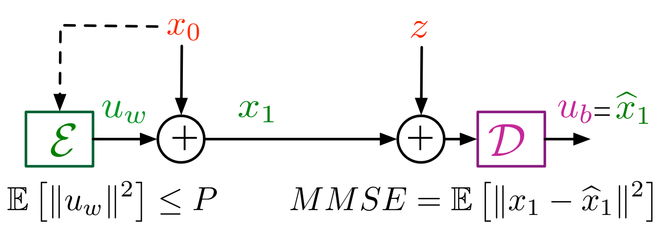
1 Distributed control and information theory

Consider multiple controllers maneuvering the state of a control system to a particular point. The purpose of this talk is to show that it helps if these controllers communicate *implicitly* with one another. Further, this unusual notion of communication can be understood *using tools from information theory*. This opens up a host of interesting new information-theoretic formulations that dovetail with exciting problems and approaches within information theory.

2 The Witsenhausen counterexample



$$\min\left\{k^2\mathbb{E}\left[u_w^2\right]+\mathbb{E}\left[x_2^2\right]\right\}$$



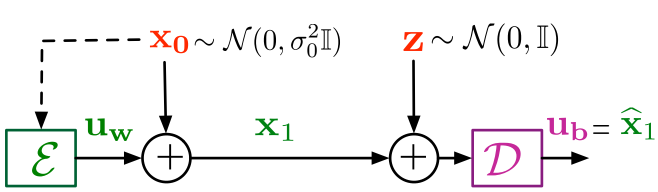
Implicit channel interpretation

The Witsenhausen counterexample [Witsenhausen'68] is a deceptively simple two-controller problem of maneuvering the state of a system to the origin.

Despite being a Linear-Quadratic-Gaussian (LQG) control problem, nonlinear control laws outperform the best linear strategy for the counterexample [Witsenhausen'68] (in fact, by an arbitrarily large factor [Mitter, Sahai '99]). The optimal control law for the problem is still elusive. The discrete relaxation of the problem is even NP-complete [Papadimitriou and Tsitsiklis '86].

The counterexample contains an implicit channel. The first controller modifies the initial state and attempts to communicate the *modified state* x_1 to the second controller through a noisy channel. Solving the problem of minimizing MMSE error in estimating $\widehat{x}_1$ with an average power constrained u_w is an equivalent formulation.

3 Vector version of Witsenhausen's counterexample



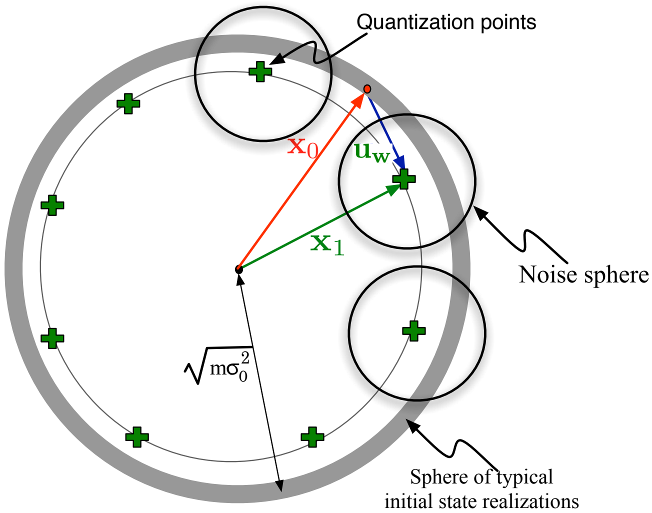
$$\frac{1}{m}\mathbb{E}\left[\|\mathbf{u_w}\|^2\right] \leq P$$

$$\mathbb{E}[\mathcal{C}] = k^2P + \frac{1}{m}\mathbb{E}\left[\|\mathbf{x}_1 - \widehat{\mathbf{x}}_1\|^2\right]$$

A natural vector extension to vector length m. The asymptotically infinite length extension offers simplification because it allows use of laws of large numbers. It also helps us avoid the complications associated with the geometry of finite-dimensional spaces.

Closely related problems have been addressed in the information theory literature, starting from Dirty-Paper Coding (DPC) [Costa '83]. Some recent related extensions of DPC that have been explored are those of state amplification (*communicating* $\mathbf{x}_0$) [Kim, Sutivong, Cover '08], state masking (*hiding* $\mathbf{x}_0$) [Merhav, Shamai '06], and some cases of distributed dirty-paper coding [Somekh-Baruch, Shamai, Verdu '08][Kotagiri, Laneman '08].

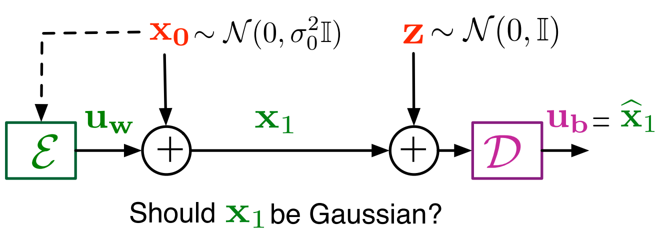
3.1 An upper bound using vector quantization



Inspired from quantization-based strategies for the scalar case [Witsenhausen '68][Mitter and Sahai '99], we propose a vector quantization strategy for the vector Witsenhausen problem [Grover, Sahai '08].

The first controller forces the initial state to the nearest quantization point. Provided the number of quantization points is sufficiently small, the second controller can decode them correctly and force the state to zero. Asymptotic cost is k^2P and 0 for the first and the second stage respectively. The required power P for perfect recovery of $\mathbf{x}_1$ turns out to equal to the noise variance, which is 1. Thus the total cost is just k^2.

3.2 A lower bound



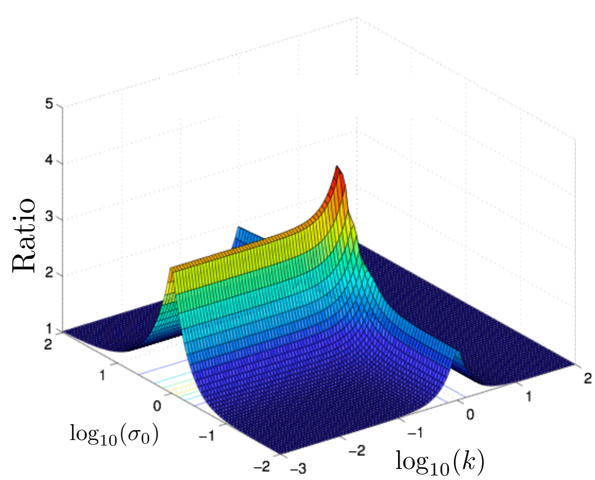
The Gaussian input distribution is capacity achieving for power-constrained Gaussian channels. However, a Gaussian source is also the hardest to estimate (in a rate-distortion sense) across the same channel. Since $\mathbf{x}_1$ is the channel input as well as the quantity that has to be estimated, there is tension in the choice of distribution for $\mathbf{x}_1$.

Ignoring this tension, the following lower bound holds [Grover, Sahai '08] on the optimal costs

$$\bar{\mathcal{C}}_{\min} \geq \inf_{P \geq 0} k^2 P + \left((\sqrt{\kappa(P)} - \sqrt{P})^+ \right)^2,$$

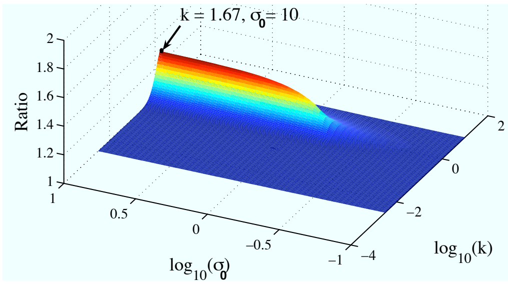
where $\kappa(P) = \frac{\sigma_0^2}{\sigma_0^2 + 2\sigma_0\sqrt{P} + P + 1}$.

3.3 An approximately optimal solution to the infinite length problem



The ratio of the vector quantization upper bound (complemented by the optimal linear strategy) and our lower bound is bounded by 4.13, providing an approximately optimal solution to the asymptotic vector Witsenhausen counterexample [Grover, Sahai '08]. Observe that the ratio is quite close to 1 for "most" values of k and σ_0^2.

3.4 Improved upper and lower bounds



An improved upper bound is obtained using a combination of dirty-paper coding and linear strategies.

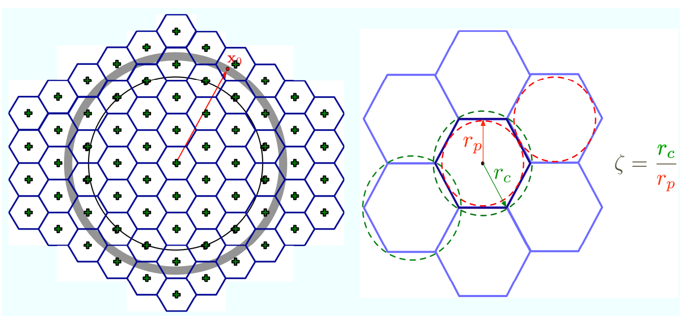
Focusing attention on the problem in the limit of zero MMSE yields an improved lower bound [Grover, Wagner, Sahai '10] that does not ignore the tension in the Gaussianity of $\mathbf{x}_1$. This new bound is still loose, but it is sufficient to show that dirty-paper coding attains within a factor of 1.3 of optimal for the asymptotic Witsenahausen counterexample.

The new lower bound also shows that DPC is optimal in the limit of zero $MMSE$ error in the estimation of the second controller (similar to thae observation in [Sumszyk, Steinberg '09] in the discrete case).

4 The finite-dimensional problem

While an asymptotic solution is sometimes considered adequate for information theory problems, it is far from adequate for control problems — vector lengths in control problems are much smaller (< 10).

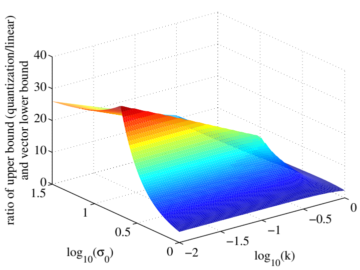
4.1 Lattice-quantization based upper bound



Since the initial state can fall outside the shell of typical realizations with non-zero probability, we need to tile the space with quantization points. Lattices provide a natural set of such quantization points. Lattices that are good for quantization and simultaneously have error resilience have a small $\zeta = \frac{r_c}{r_p}$, where r_c is the covering radius and r_p is the packing radius of the lattice.

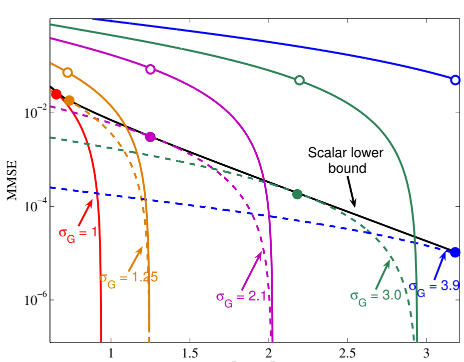
The strategy thus is : the first controller C_w drives the state to the nearest quantization point. The second controller C_b estimates the state to be the nearest quantization point, and forces it to zero [Grover, Sahai, Park '09].

4.2 The vector lower bound is insufficient



The figure shows that in the scalar case, the ratio of the cost attained by the quantization-based strategies and our vector lower bound runs off to infinity in the low k regime. As we shall see, this is because of the looseness in the lower bound, rather than a mediocre performance of quantization-based strategies.

4.3 A tighter "sphere-packing" lower bound

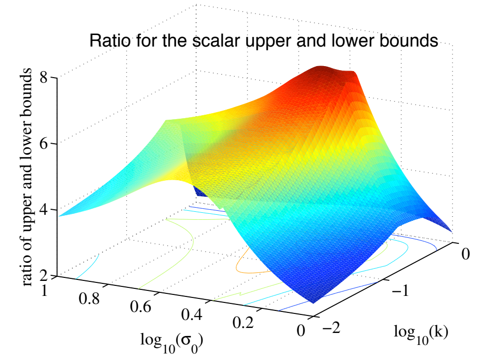


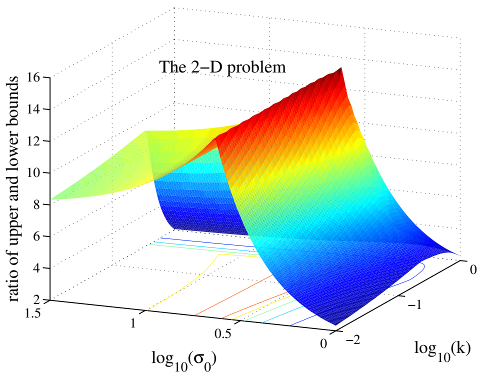
A new lower bound that is specific to a given vector length. The derivation [Grover, Sahai, Park '09] is a large deviations perspective on the "sphere-packing" bounding technique. The observation noise can behave as if its variance is much larger, thereby increasing the lower bound. For finite dimensions, there is a non-zero probability of such atypical behavior.

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 behaves as a Gaussian with variance σ_G^2 provides a lower bound to our problem.

The figure illustrates this intuition in the scalar case.

4.4 Approximate optimality for the finite-length Witsenhausen counterexample



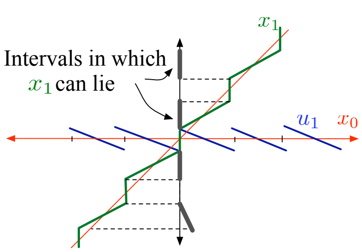


The bound inspired from the "sphere-packing" technique is sufficient to show that lattice-quantization strategies (complemented by linear strategies) attain within a factor of 8 of the optimal cost uniformly over all k and σ_0 for the original (scalar) Witsenhausen counterexample [Grover, Sahai, Park '09].

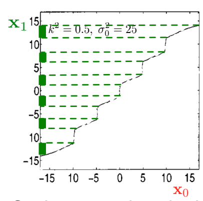
More generally, we show that lattices can attain within a constant factor of the optimal costs for any finite dimension uniformly over all k, σ_0 and m, the number of dimensions. The approximation factor ($100\zeta^2$, where $\zeta \leq 4$ for each m) is loose because the analysis is crude.

Numerical calculations show that this factor is in fact much smaller. For example, quantization to a hexagonal lattice ($\zeta = \frac{2}{\sqrt{3}}$) attains within a factor of 15 of the optimal cost for the 2-dimensional Witsenhausen counterexample.

4.5 Conjectured optimal strategy : dirty-paper coding



Scalar version of our DPC-based strategy

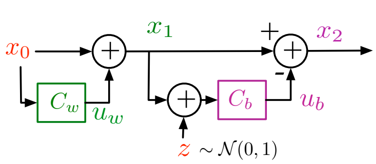


Scalar strategies obtained by semi-exhaustive search in [Baglietto, Parisini, Zoppoli] [Lee, Lau and Ho]

Empirical evidence found by semi-exhaustive search in function spaces [Baglietto, Parisini and Zoppoli][Lee, Lau and Ho] yields a "slopey" quantization strategy.

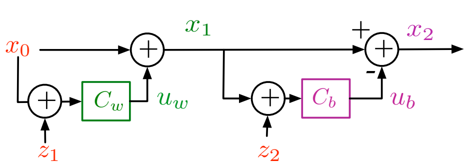
This slopey-quantization strategy is in fact equivalent to scalar dirty-paper coding [Grover and Sahai '08]. This fact, along with the proximity to optimality of DPC for infinite-length problem, leads us to conjecture that DPC-based strategy is optimal for the Witsenhausen counterexample.

5 More general distributed control problems



$\min \left\{ \mathbb{E}\left[k_0^2 x_0^2 + k_1^2 x_1^2 + k_w^2 u_w^2 + k_b^2 u_b^2 + x_2^2\right] \right\}$

Witsenhausen problem with costs on all states and inputs



$\min \left\{ k^2 \mathbb{E}\left[u_w^2\right] + \mathbb{E}\left[x_2^2\right] \right\}$

Witsenhausen problem with noise in the observations of both controllers

It may appear that Witsenhausen's formulation is an isolated case where implicit communication is useful. After all, it can be thought of as an encoder-decoder system.

In order to test this, we consider two extensions of Witsenhausen's counterexample. In [Grover, Park, Sahai '09] we impose costs on all states and inputs. Since the second controller has a cost on its input, it is not clear whether it can be thought of as a decoder.

In [Grover, Sahai '10], we consider an extension with noise in the observation of the first controller as well. No longer is it clear if the first controller be thought of as an encoder.

For both of these problems, we provide implicit communication based strategies that attain within a constant factor of optimal.