The role of common "context" in signaling

Pulkit Grover and Cedric Langbort
{pulkit@stanford, langbort@illinois}.edu

***Abstract*— In order to signal effectively, is it sufficient for decentralized agents to share a signaling protocol, a common language? In this paper, we argue that the agents also need a common signaling "context" that can depend on the environment. Paralleling traditional communication, where mapping of information-sequence to codewords can be thought of as a "language", the meaning attached to points in the state-space defines a signaling-language. In this framework, the coordinate-axes of an agent can be viewed as its signaling-context. By investigating the impact of context-misalignment in the minimalist signaling problem, the Witsenhausen counterexample, we show that significant context-misalignment can lead to lack of coordination despite agreement on the signaling-language.**

While Witsenhausen's counterexample is a single-shot problem, decentralized agents often act for longer time-horizons. By formulating a multi-shot extension of Witsenhausen's counterexample, we show that the agents can arrive at a common context by observing the costs arising due to their actions at previous time-steps, and are then able to use their common language to coordinate effectively.

I. Introduction

In decentralized control, it is well understood that coordination between control agents can often help improve the control performance. How do the agents establish coordination? In particular, if two agents are thrown on the ground to work as a team, what do they need in order to develop coordination? The purpose of this largely speculative paper is to explore this question.

Let us start with a simple example where it is patently obvious that the decentralized agents are acting in a coordinated manner: communication. For instance, a communication strategy often involves choosing discrete points in the signal-space and attaching meaning to those points (e.g. 0 and 1 for two phases in Binary Phase Shift Keying (BPSK) [1]). The meaning attached to the points is not derived merely from the problem structure[1], but arises because of *agreement among the agents on the assigned meaning*. A commonly agreed map of the signal-space can be called a "language". Indeed, a common language is necessary for communication. But is it sufficient? A well calibrated communication system when put on ground can fail because of channel uncertainties: channel parameters that not all participants in the system know *apriori* (e.g. the channel-fade coefficient may be known at the receiver, but is often unknown at the transmitter). It is these parameters, that might be known only locally and only at run-time, that we call the "context". Understanding the communication context *at run-time* is crucial to the performance of the communication system.

More precisely, a communication language can be thought of as a mapping from the set of observations $\mathcal{X}$ and the set of contexts $\mathcal{C}$ into the language set $\mathcal{L}$, *i.e.* $f : \mathcal{X} \times \mathcal{C} \to \mathcal{L}$. In order to communicate effectively, the agents may not merely need the map f, but also the contexts $C \in \mathcal{C}$ of other agents.

In traditional control literature, it has been observed that even when external communication channels are not present, control agents can often use coordinated strategies to 'signal' to each other [2], [3]. Intuitively, one agent is said to signal to another agent when the state modified by the former is observed by the latter at a later time (thus opening up the possibility of the first agent "writing on the state"). The possibility of signaling has been observed to introduce conceptual difficulties. From a control-theoretic perspective, control actions may have to balance between two roles, namely, control (e.g. reducing immediate costs) and signaling [4]. From an information-theoretic standpoint, signaling goes beyond traditional communication[2]: because no message is specified *apriori* in signaling, the message may be a matter of choice [6]! In [6]–[8], the authors use tools from information theory to investigate the minimalist problem of signaling: the Witsenhausen counterexample. The resulting bounds provide the first provably-approximately-optimal solutions to the problem, suggesting that these difficulties in understanding signaling could potentially be overcome in more complicated problems as well.

Paralleling the development of understanding of communication, a natural intellectual question arises in signaling: can decentralized agents signal in absence of any "signaling context"? Where should we start in order to understand this question? Witsenhausen's counterexample [9] has proven useful in understanding other aspects of signaling because it distills the concept of signaling to a simple two-user one-shot problem. Thus our investigation begins from Witsenhausen's counterexample as well. In general, if the optimizing signaling policy for the entire system is unique[3], each agent can choose its strategy as suggested by the overall optimal policy. What if the optimizing policy is not unique, or, as is the case

[1]Problem structure can help shape the communication strategy, as it almost always does. But this shaping is almost never unique, and the resulting strategies can be widely different. For instance, while communication of 'rain' or 'no rain' prompts a binary communication strategy, the assignment of '0' an '1' to 'rain' or 'no rain' is arbitrary.

[2]The closest relatives of signaling in information theory are dirty-paper coding [5] and its recent extensions.

[3]In stochastic setups, this uniqueness is needed only on a set of measure 1, not the entire space of primitive random variables.

for Witsenhausen's counterexample, too hard to compute[4]? Allowing the agents to engage in "cheap talk" before being put on ground could help agree on the choice of strategies. The process of this "cheap-talk" could be, for instance, the hardcoding of a set of commands into the agents at the time that they are manufactured[5]. These commands can serve as the signaling protocol once the agents are put on the ground. The question of interest in this paper is: *would a common language suffice for signaling-coordination (see Fig. 1)? Or do the agents on ground need more than a common language in order to coordinate effectively?* In particular, what if some environmental parameters (or "context") is not known at the time of strategy design (and might be known only locally at each agent at run-time), even if the impact of these parameters on the signal space is known?

In order to answer this question for simplistic cases, in Section II we consider a variation on Witsenhausen's counterexample [9]. We take the coordinate axes of an agent as a manifestation of the environmental context: much as lack of synchronization in communication systems can lead to reduced performance, the misalignment between axes of different agents can lead to disagreement on the assigned meaning of the different parts of the signaling state-space. Intuitively, with large misalignment, all meaning of language may be lost. We show in Section II that this is indeed the case! We show that for sufficiently large misalignment between the axes, there is *no advantage to using any coordination*, *i.e.* it is *optimal* for each agent to operate oblivious to the strategy of the other agent. Further, we provide approximately-optimal strategies for all values of misalignment. We observe that signaling strategies are only necessary when the degree of misalignment is smaller than the observation noise of C_2.

While Witsenhausen's counterexample serves as a useful starting point, while examining lack of common context in signaling, it is intuitive that the agents should be able to evolve a common context in a repeated game. In order to investigate this, in Section III we consider a multi-shot version of our problem. We show that if the agents are told the attained cost after each time-step, then they can adaptively learn the misalignment through trial and error. The agents can thus learn to develop a common context. We also provide approximately optimal strategies for all degrees of misalignment and all time-horizons for this problem. We conclude in Section IV by pointing out some practical applications and alternative models for future work.

II. Single-shot case

1) Problem statement: The problem we consider in this section is a non-stochastic variation on Witsenhausen's

[4]Even when the optimizing policy is unique, the control agents may not be able to agree on a strategy because the computation of the optimal strategy might be hard! For instance, while the optimal strategy for the celebrated Witsenhausen counterexample [9] is not known, and can be NP-hard to compute by straightforward discretization [10], approximately-optimal solutions (that are not unique) are known explicitly [8]. Similarly, it was recently shown that finding a Nash equilibrium is PPAD complete [11].

[5]Recent work of Juba and Sudan [12] explores algorithms for arriving at a common language *without* cheap-talk, *i.e.* at run-time.

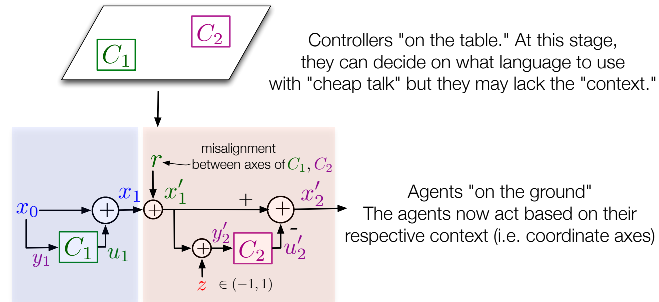

Fig. 1. In order to facilitate coordination, a two-step process is sometimes used. At the first step, the agents can talk *ad infinitum* and converge on a signaling language. However, the agents have no knowledge of the problem parameters (in this case, x_0, r, z, y) at this step. At the second step, the agents play their strategies to minimize the cost. The figure shows this implementation with Witsenhausen's counterexample [9] as an example.

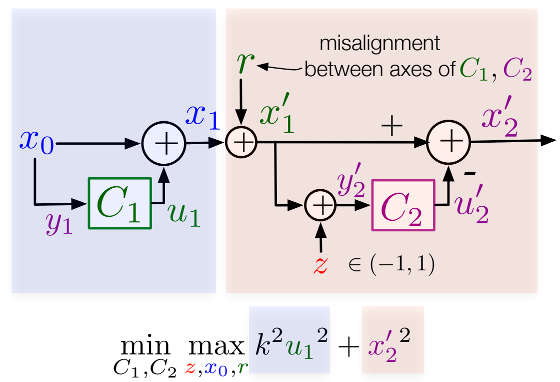

Fig. 2. A non-stochastic variation on Witsenhausen's counterexample with misalignment between the axes of the two control agents (denoted by r). The problem is not the same as that in [13] because r is unknown at the two agents.

counterexample (based on a similar variation that appeared in [13]), and is shown in Fig. 2. The initial state x_0 and the noise z are chosen arbitrarily, with the noise z restricted to fall within $(-1, 1)$. The misalignment between the axes of the two agents is denoted by $r \in (-a, a)$. The parameter r is not known at the two agents. Assuming the origin of the first agent as the reference origin, the origin of the second agent is displaced by r. In the reference frame of the second agent, the coordinates are denoted by a prime ($'$) above the relevant parameter. For instance, x_1'=$x_1 + r$ is x_1 in the reference frame of C_2.

The state-evolves as $x_1 = x_0 + u_1$ and $x_2' = x_1' - u_2'$. The observation of C_1 is $y_1 = x_0$ and of C_2 is $y_2 = x_1' + z$. The goal is to minimize the worst-case cost, *i.e.*

$$\mathcal{J}_{opt} = \min_{C_1,C_2} \max_{z,x_0,r} k^2 u_1^2 + x_2'^2. \tag{1}$$

The cost for any strategy $\gamma = (\gamma_1, \gamma_2)$, where γ_i is the mapping from observations to control actions of agent i, is denoted by $\mathcal{J}^{(\gamma)}$. The optimal cost is denoted by $\mathcal{J}_{opt}$ (as above).

In [13], the problem with no misalignment between the coordinate axes was considered (*i.e.* $a = 0$). It was shown that quantization-based signaling strategies (complemented by linear strategies) attain within a constant factor of the optimal for all problem parameters. Are similar signaling strategies still useful for our problem? The next section explores this question.

2) Approximately-optimal strategies: The following theorem provides upper bounds on the costs that are obtained using quantization strategy complemented by the zero-input strategy.

Theorem 1 (Upper bound):

$$\mathcal{J}_{opt} \leq \min\{k^2(1+a)^2 + a^2, 1\} \tag{2}$$

Proof: The first term in the upper bound is obtained

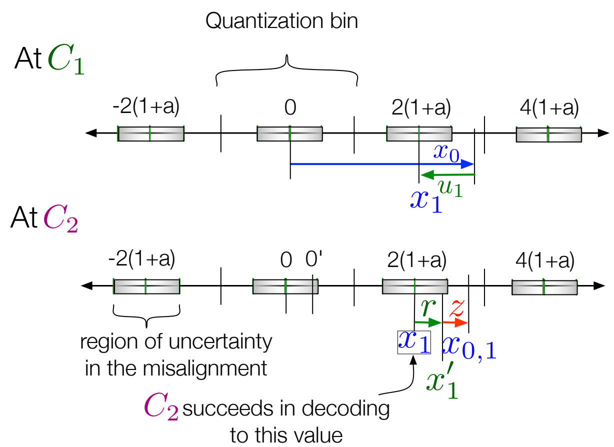

Fig. 3. An illustration of the quantization-based signaling strategy for the upper bound on costs. Using a quantization bin-size of $2(1+a)$, one can always ensure that C_2 gets the bin right, even though the bin-center is only known to within an error of a. Thus C_2 decodes to x_1, even though the state from its reference point is x_1'. As a quick example, consider the case when $a = 0.5$. Say $x_1 = 2(1+a) = 3$. If $r = 0.25$ (unknown to the two agents), then the observation at C_2 is $y_2' = x_1 + r + z = 3.25 + z$. Because $z < 1$, this will never exceed 4.25, which is smaller than 4.5, the largest point in the bin. Thus C_2 will succeed in decoding to 3. However, the state $x_1' = x_1 + r = 3.25$, thereby incurring an error of 0.25.

by using a quantization-based strategy that uses uniform quantization bins of size $2(1+a)$, thereby costing at most $k^2(1+a)^2$ at the first stage. Because the noise is always bounded above by 1, and the misalignment is bounded above by a, using a bin-size of $2(1+a)$ we can ensure that C_2 can always decide in what interval x_1 lies (see Fig. 3). The resulting error in estimating the state at C_2 is bounded by a, resulting in a second stage cost of a^2.

The second term in the upper bound is obtained by using the zero-input strategy: the first agent uses zero-input, and the second agent uses the observation as the state estimate, incurring a maximum second stage cost of 1. ■

The following provides a fundamental lower bound on the cost for any strategy.

Theorem 2 (Lower bound): The cost for any strategy γ is lower bounded by

$$\mathcal{J}_{opt} \geq$$
$$\max\left\{\min\{1, a^2\}, \inf_{P \geq 0} k^2 P + \left(\left(\sqrt{\tfrac{2}{\pi e}} - \sqrt{P}\right)^+\right)^2\right\}.$$

Proof: Constraining r, z such that $r + z = 0$, the cost for any strategy γ is lower bounded as follows

$$\begin{aligned}
&\mathcal{J}^{(\gamma)} \\
\geq\ & \sup_{r\in(-a,a), z\in(-1,1), x_0} k^2\left(C_1^{(\gamma)}(x_0)\right)^2 \\
&\quad + \left(x_1' - C_2^{(\gamma)}(x_1 + r + z)\right)^2 \\
\geq\ & \sup_{r\in(-a,a), z\in(-1,1), x_0} \left(x_1' - C_2^{(\gamma)}(x_1 + r + z)\right)^2 \\
\overset{(x_1'=x_1+r)}{=}\ & \sup_{r\in(-a,a), z\in(-1,1), x_0} \left(x_1 + r - C_2^{(\gamma)}(x_1 + r + z)\right)^2 \\
\geq\ & \sup_{\substack{r\,\in\,(-a,a),\, z\,\in\,(-1,1),\, x_0,\\ r+z=0}} \left(x_1 + r - C_2^{(\gamma)}(x_1 + r + z)\right)^2 \\
=\ & \sup_{\substack{r\,\in\,(-a,a),\, z\,\in\,(-1,1),\, x_0,\\ r+z=0}} \left(x_1 + r - C_2^{(\gamma)}(x_1)\right)^2 \\
=\ & \sup_{\substack{r\,\in\,(-a,a),\, z\,\in\,(-1,1),\, x_0,\\ r+z=0}} \left(x_1 - C_2^{(\gamma)}(x_1) + r\right)^2 \\
\overset{(a)}{\geq}\ & \left(|x_1 - C_2^{(\gamma)}(x_1)| + \min\{1, a\}\right)^2 \\
\geq\ & \min\{1, a^2\}.
\end{aligned}$$

where (a) holds because r has to satisfy $|r| < a$ and further, because $r + z = 0$, and $|z| < 1$, it also satisfies $|r| < 1$. Thus $r < \min\{1, a\}$.

The trick in the above proof is the observation that for any choice of r such that $|r| < \{1, a\}$, we can keep u_2' to a constant value $C_2^{(\gamma)}(x_1)$ by adjusting z so that the second controller sees exactly the same observation y_2'.

This proves the first term in the lower bound. The second term in Theorem 2 is obtained as follows: if r is known exactly at C_2, then it knows the axes of the first agent as well and the problem becomes the version of Witsenhausen counterexample with arbitrary state and bounded noise that was considered in [13]. Thus, the lower bound of [13, Theorem 4] applies[6], which is the second term in Theorem 2. ■

Theorem 3 (Single-shot approximate optimality): The following bounds characterize the optimal cost to within a constant factor

$$\frac{1}{\mu}\min\{k^2(1+a)^2+a^2, 1\} \leq \mathcal{J}_{opt} \leq \min\{k^2(1+a)^2+a^2, 1\}, \tag{3}$$

where $\mu \leq 76$. That is, quantization-based strategies, complemented by linear strategies, attain within a factor of 76 of the optimal.

Proof: Case 1: $a \geq 1$. In this case, the upper and lower bounds are both 1, and the ratio is also 1.

Case 2: $k^2 \geq 1, a < 1$.

[6]The bound of [13, Theorem 4] applies with a slight modification. In [13], the noise $z \in (-\sqrt{3}, \sqrt{3})$. Here the noise $z \in (-1, 1)$. As would be expected, this modification does not require any significant change in the derivation which is why we do not re-derive the result here.

In this case, we use the upper bound of 1, obtained by using the zero-input strategy. For the lower bound, let P^* denote the optimizing value of P.

Case 2a: If $P^* > \frac{1}{2\pi e}$, then the lower bound is larger than $k^2 P^* = \frac{k^2}{2\pi e} > \frac{1}{2\pi e}$. Using the zero-input upper bound of 1, the ratio of upper and lower bounds is smaller than $2\pi e < 18$.

Case 2b: If $P^* \leq \frac{1}{2\pi e}$, then the lower bound is larger than $\left(\sqrt{\frac{2}{\pi e}} - \sqrt{P^*}\right)^2 \geq \frac{1}{2\pi e}$. Using the zero-input upper bound of 1, the ratio is again smaller than $2\pi e < 18$.

Case 3: $k^2 < 1, a < 1$. If $P^* < \frac{1}{2\pi e}$, an upper bound is 1. The lower bound is larger than $\left(\sqrt{\frac{2}{\pi e}} - \sqrt{P^*}\right)^2 > \frac{1}{2\pi e}$. Thus the ratio is smaller than $2\pi e < 18$.

If $P^* \geq \frac{1}{2\pi e}$, the lower bound is larger than $\max\{a^2, \frac{k^2}{2\pi e}\}$. This means that the lower bound is also larger than any convex combination of these two quantities. Thus,

$$\mathcal{J}_{opt} \geq \frac{a^2}{10} + \frac{9}{10}\frac{k^2}{2\pi e}. \quad (4)$$

The upper bound is smaller than $k^2(1+a)^2 + a^2 \leq 4k^2 + a^2$. Thus the ratio is smaller than $\max\{10, 2\pi e \times \frac{10}{9} \times 4\} < 76$.

Thus the ratio of upper and lower bounds is bounded by 76 for all problem parameters. Our bounding technique is quite crude, and the actual ratio is likely much smaller. ∎

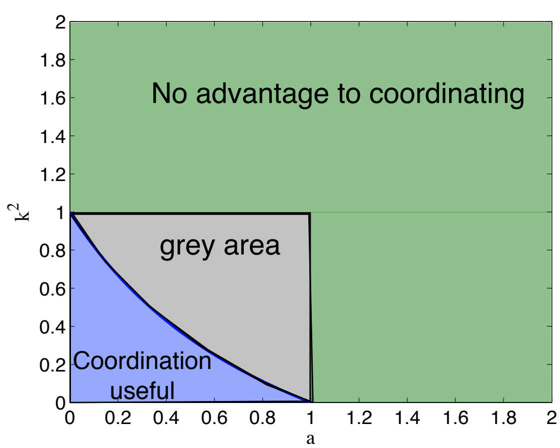

Fig. 4. The regimes in which coordination is useful in the one-shot case. Our results do not provide conclusive evidence of whether coordination is useful in the "grey area" because both quantization-based signaling and non-coordinating zero-input strategies are order-optimal in this region.

3) When is coordinating useful?: In order to understand when coordinating is useful in this one-shot problem, we first need to see what strategies do not assume any coordination. A case where both agents use strategies that are completely oblivious to the strategy of the other can be thought of as uncoordinated. If C_1 is oblivious to the strategy of C_2, then it should use no input at all, because any input has an associated cost. For C_2, since it is ignoring the strategy of C_1, it should always decode the observation as the state estimate. In that case, the error x_2 is always bounded above by 1. The following corollary shows that, remarkably, this uncoordinated strategy is sometimes optimal.

Corollary 1 (Of Theorem 1 and Theorem 2): For $a > 1$, the optimal strategy is the non-coordinating zero-input strategy.

Proof: The corollary follows from the simple observation that the upper bound of zero-input and the lower bound of Theorem 2 equal 1 for $a > 1$. ∎

To obtain a more refined understanding of when coordination is useful, we now compare the performance of our strategies with each other and with the lower bound.

The quantization-based signaling strategy attains an upper bound of $k^2(1 + a)^2 + a^2$, while the zero-input no-coordination strategy attains an upper bound of 1. Clearly, for $a > 1$ or $k > 1$, the cost of quantization-strategy is larger than 1, and thus the zero-input is the better strategy, and there is no advantage of signaling in this one-shot problem.

Comparing the two costs when $a < 1$ and $k^2 < 1$,

$$\begin{aligned}
& k^2(1+a)^2 + a^2 \leq 1 \\
& \textit{i.e. } k^2(1+a)^2 \leq 1 - a^2 = (1+a)(1-a) \\
& \overset{(\text{since } a<1)}{\Rightarrow} \quad k^2(1+a) \leq 1 - a \\
& \Rightarrow k^2 \leq \frac{1-a}{1+a}, \ \textit{i.e. } a \leq \frac{1-k^2}{1+k^2}.
\end{aligned}$$

Thus, signaling is certainly useful in the one-shot case when $a \leq \frac{1-k^2}{1+k^2}$ with $a, k < 1$. Also, no coordination is useful when $a > 1$ or $k > 1$.

4) What provides context in this one-shot problem?: Even when the control agents can talk to each other indefinitely before they are put on the ground, if there is arbitrary misalignment between the coordinate axes of the two agents, then in the one-shot problem of this section, there is no advantage to coordinating.

Clearly, the context in this one-shot problem is provided by the alignment of the reference axes. Our bounds show that *regardless of the signaling strategy*, the two agents share a useful common context only if the misalignment between the axes of the two agents is no larger than the sensor noise of the second agent. Does this strong conclusion hold true when we have a multi-shot problem? The next section examines this question.

III. Multi-shot Case

The conceptual difference that the multi-shot version of the problem creates is that it introduces the possibility of *learning* the misalignment.

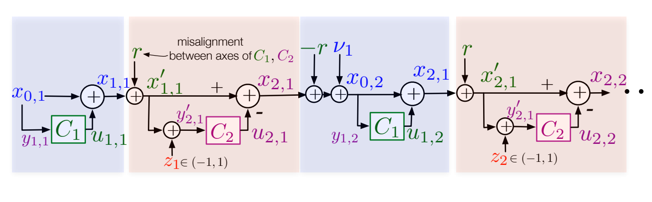

Fig. 5. A natural extension of the problem in Section II to a multi-shot case. The point that is worthy of note is that new perturbation variables ν_i is introduced that perturb the state evolution. Without these perturbation variables, the cost to go can be made zero with $n \geq 3$.

5) Problem statement: The problem is the natural multi-step extension of that in Section II, and is shown in Fig. 5. The agents can learn the misalignment only if they gain some knowledge from previous time-steps. We assume that a genie reveals the total costs at the end of each time-step in order to help them learn the misalignment.

While the values of the noise z_i and the perturbation ν_i can change with time arbitrarily, the coordinate-shift parameter r is assumed to be fixed. In order to denote time in the multi-shot case, the variables in single-shot case are appended with the index i to indicate the stage. For instance, $u_{2,i}$ denotes the control input of the second agent at time i.

6) Lower and upper bounds on the total cost:

Theorem 4 (Lower bound): For time horizon $n \geq 2$, the optimal total cost $\mathcal{J}^{(n)}$ is lower bounded as follows:

$$\mathcal{J}_{opt}^{(n)} \geq \max\left\{\min\{1,a^2\}, \inf_{P\geq 0} k^2P + \left(\left(\sqrt{\tfrac{2}{\pi e}}-\sqrt{P}\right)^+\right)^2\right\} +(n-1)\left(\inf_{P\geq 0} k^2P + \left(\left(\sqrt{\tfrac{2}{\pi e}}-\sqrt{P}\right)^+\right)^2\right)$$

Proof: The first term in the lower bound is the lower bound from Theorem 2. The second term is the lower bound on the sum of the costs at each stage of a single-shot Witsenhausen counterexample with the parameter r known at the two agents [13].[7]

For purposes of the lower bound, can we decouple the costs? If we provide the knowledge of $x_{2,i}$ (also called "side-information" in information theory literature [14]) to the agents at stage $i+1$, then the problem is effectively reset with no memory of the past, and the costs can only be lower (because the agents can choose to ignore this side information). We thus obtain a lower bound for each stage. ∎

The key part is the upper bound, which shows how the parameter r can be learned from the availability of attained costs at each time-step.

Theorem 5 (Upper bound): For time horizon $n \geq 2$, if $k^2 \geq 1$, $\mathcal{J}_{opt}^{(n)} \leq n$, if $k^2 < 1$, and $a \geq 1$, $\mathcal{J}_{opt}^{(n)} \leq 4k^2n+16$, and if $k^2 < 1$ and $a < 1$, $\mathcal{J}_{opt}^{(n)} \leq (n+3)k^2+5a^2$.

Proof: Case 1: $k^2 \geq 1$. We use zero-input strategy at all times, attaining a cost of at most 1 at each time-step.

Case 2: $k^2 < 1, a \geq 1$. At $t=1$, C_1 uses the control input to force the state to the nearest quantization point with a bin-size 4. C_2 then estimates the state to be the nearest state *smaller* than the observation $y_{2,1}$. Estimating to nearest smaller state ensures that the error (which is due to misalignment of axes) is always positive. The cost at this stage is, therefore, $4k^2+16$.

The positivity of error ensures that C_1 knows the misalignment to within $(mod\ 4)$ using the error. C_1 can thus align the *quantization bins* (though not the axes) at the second time-step. The cost for the remaining time is therefore $4k^2$ (for quantization bins of width 4).

Case 3: $k^2 < 1, a < 1$. At $t=1$, the agents use the quantization-based signaling strategy shown in Fig. 3. The resulting cost is given by $k^2u_{1,1}^2 + r^2$, where $u_{1,1}$ is the input (of maximum amplitude $1+a$, and input cost at most $4k^2$) required to force x_0 to the nearest quantization point. The resulting squared error is at most a^2. At stage 2, the agents are told the attained cost at stage 1. Since C_1 knows $u_{1,1}$, it can calculate r^2, and thus knows r to within a set of two points, namely $+r$ or $-r$. At the second time-step, it uses a quantization codebook of bin-size 2, shifted by r (to compensate for the misalignment). The resulting total cost can be $k^2u_{1,2}^2$ or, if the axes are not aligned yet, $k^2u_{1,2}^2+4r^2$, which is at most k^2+4a^2. If the resulting cost is larger than $k^2u_{1,2}^2$, then C_1 can assume that the shift is $-r$ and from then on, use a quantization codebook of bin-size 2 shifted with $-r$. The cost is upper bounded by k^2 for every stage henceforth. ∎

We note that in *Case 2*, even though the strategy does not achieve a perfect alignment of the axes (but only to within modulo 4), it performs as well as a binning strategy with perfect alignment! Thus the context need not be learned completely in order for the signaling strategy to be effective.

Theorem 6 (Multi-shot case approximate-optimality): The ratio of upper bound in Theorem 5 and lower bound in Theorem 4 is smaller than 140 for all values of k^2, a and $n \geq 2$.

Proof:

Case 1: $k^2 \geq 1$. Upper bound is smaller than n (zero-input strategy). Lower bound is larger than $(n-1)\left(k^2P^* + \left(\left(\sqrt{\tfrac{2}{\pi e}}-\sqrt{P^*}\right)^+\right)^2\right)$, where P^* is the value of P that minimizes $k^2P + \left(\left(\sqrt{\tfrac{2}{\pi e}}-\sqrt{P}\right)^+\right)^2$. If $P^* \leq \frac{1}{2\pi e} = \frac{2}{4\pi e}$, the lower bound is larger than $(n-1)\frac{1}{2\pi e}$, and the ratio is smaller than $2 \times 2\pi e = 4\pi e < 35$. If $P^* > \frac{1}{2\pi e}$, the lower bound is larger than $(n-1)k^2\frac{1}{2\pi e} \geq \frac{n-1}{2\pi e}$. Thus the ratio is again smaller than $4\pi e < 35$.

Case 2: $k^2 < 1, a > 1$.

The upper bound is smaller than $4k^2n+16$. The lower bound is larger than $1 + (n-1)\left(k^2P^* + \left(\left(\sqrt{\tfrac{2}{\pi e}}-\sqrt{P^*}\right)^+\right)^2\right)$. Again, if $P^* \leq \frac{1}{2\pi e}$, the lower bound is larger than $1+\frac{n-1}{2\pi e}$, and thus the ratio is smaller than $\max\{4\times 2\times 2\pi e, 16\} = 16\pi e < 140$. If $P^* > \frac{1}{2\pi e}$, the lower bound is larger than $1+\frac{(n-1)k^2}{2\pi e}$, and the ratio is again smaller than 140.

Case 3: $k^2 < 1, a < 1$. The upper bound is smaller than $(n+3)k^2+5a^2$. The lower bound is larger than $a^2+(n-1)\left(k^2P^* + \left(\left(\sqrt{\tfrac{2}{\pi e}}-\sqrt{P^*}\right)^+\right)^2\right)$. By the same argument as in *Case 2*, the ratio is smaller than $\max\{4\times 2\pi e, 5\} < 70 < 140$. ∎

[7] Our lower bound allows for learning the parameter r after just one time-step. However, because C_1 can only know the magnitude of the error after the first time-step, this is clearly not possible. This is one evidence why the derived bounds are loose.

IV. Discussions and conclusions

We discussed how decentralized agents can use a common context in order to coordinate, and that a context can be obtained by repeated play. Interestingly, the agents do not have to obtain complete alignment of axes in order to signal to each other: for quantization-based strategies, alignment of quantization bins is sufficient (*i.e.* the origins of the axes of the two agents can differ by a factor of the quantization bin-size).

While here we have assumed that the agents are provided with the payoffs after each time-step, often the information gained from a game-play could be different. A team of robots put on the ground may have no common reference frame for their positions to begin with, but could develop one with trial and error in repeated play. For instance, if one robot can observe another, it can adjust its axes in accordance with the other's errors in order to get the two aligned. This case can be modeled by revealing the estimation error to C_1, instead of the total cost formulation considered in this paper. However, our results here easily extend to this case as well.

Alternatively, we can consider a problem where the agents obtain *no information* after each time step. Is it possible to obtain the context at all in such cases? Consider Gaussian observation noise, instead of bounded adversarial noise considered thus far. If the quantization bin-sizes are sufficiently large, independent observations of quantization bins can be obtained over a long time. This can help zero down on the quantization bins' limits, reducing the misalignment so that coordination is useful. However, the misalignment can never be made zero, and thus there is a tradeoff between the misalignment and the immediate costs that is reminiscent of tradeoffs between probing (information gathering) and control quality in system identification.

Other forms of misalignment are also possible. For communication problems, misalignment in phase typically arises due to shift of axes, but misalignment in fade parameter can arise due to an unknown *scaling* of axes. While either form of misalignment can hurt, misalignment in scaling can hurt rather severely: if the scale parameter evolves with time fast enough that it cannot be learned, the channel capacity is known to scale only as $\log\log(SNR)$ [15], instead of $\log(SNR)$ when the fade parameter is known exactly. It would be interesting to explore what effect a scaling misalignment, and/or a slowly time-evolving misalignment, has on problems in decentralized control.

Acknowledgments

Cedric Langbort was supported in part by the U.S. Air Force Office of Scientific Research (AFOSR) under grant number MURI FA9550-10-1-0573. Pulkit Grover was supported in part by NSF grant CNS-0932410 and in part by the Interconnect Focus Center of the Semiconductor Research Corporation.

References

[1] S. Haykin, *Digital communications*. Wiley, 1988.

[2] Y. C. Ho, M. P. Kastner, and E. Wong, "Teams, signaling, and information theory," *IEEE Trans. Autom. Control*, vol. 23, no. 2, pp. 305–312, Apr. 1978.

[3] R. Bansal and T. Başar, "Stochastic teams with nonclassical information revisited: When is an affine control optimal?" *IEEE Trans. Autom. Control*, vol. 32, pp. 554–559, Jun. 1987.

[4] Y. Bar-Shalom and E. Tse, "Dual effect, certainty equivalence, and separation in stochastic control," *IEEE Trans. Autom. Control*, vol. 19, no. 5, pp. 494 – 500, Oct. 1974.

[5] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[6] P. Grover, "Actions can speak more clearly than words," Ph.D. dissertation, UC Berkeley, Berkeley, CA, 2010.

[7] P. Grover and A. Sahai, "Vector Witsenhausen counterexample as assisted interference suppression," *Special issue on Information Processing and Decision Making in Distributed Control Systems of the International Journal on Systems, Control and Communications (IJSCC)*, vol. 2, pp. 197–237, 2010.

[8] P. Grover, S. Y. Park, and A. Sahai, "The finite-dimensional Witsenhausen counterexample," Submitted to IEEE Transactions on Automatic Control, 2010. [Online]. Available: http://arxiv.org/abs/1003.0514

[9] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, Jan. 1968.

[10] C. H. Papadimitriou and J. N. Tsitsiklis, "Intractable problems in control theory," *SIAM Journal on Control and Optimization*, vol. 24, no. 4, pp. 639–654, 1986.

[11] C. Daskalakis, P. Goldberg, and C. Papadimitriou, "The complexity of computing a nash equilibrium," in *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, 2006, pp. 71–78.

[12] B. Juba and M. Sudan, "Universal semantic communication I," in *Proceedings of the 40th annual ACM symposium on Theory of computing*. ACM, 2008, pp. 123–132.

[13] P. Grover and A. Sahai, "Is Witsenhausen's counterexample a relevant toy?" *Proceedings of the 49th IEEE Conference on Decision and Control (CDC), Altanta, Georgia, USA*, Dec. 2010.

[14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 1st ed. New York: Wiley, 1991.

[15] A. Lapidoth and P. Narayan, "Reliable Communication Under Channel Uncertainty," *IEEE Trans. Inf. Theory*, vol. 44, no. 10, pp. 2148–2177, 1998.