On the Inapproximability of the Discrete Witsenhausen Problem

Alex Olshevsky

***Abstract*—We consider a discrete version of the Witsenhausen problem where all random variables are bounded and take on integer values. Our main goal is to understand the complexity of computing good strategies given the distributions for the initial state and second-stage noise as inputs to the problem. Following Papadimitriou and Tsitsiklis, who showed that computing the optimal solution is NP-complete, we construct a sequence of problem instances with the initial state uniform over a set of size n and the noise uniform over a set of size at most n^2, such that finding a strategy whose cost is a multiplicative $n^{2-\epsilon}$ approximation to the optimal cost is NP-hard for any $\epsilon > 0$.**

***Index Terms*—Decentralized control, computational complexity.**

I. Introduction

WITSENHAUSEN'S seminal counterexample [2] demonstrated that linear strategies are not always in sequential stochastic control. The counterexample consists of a two-agent optimization problem with what has come to be known as a non-classical information pattern, in that it involves two agents acting in sequence, with the second agent having no knowledge of the information seen by the first agent. In the decades since [2], a considerable literature has sprung up analyzing control problems with non-classical information patterns [3]. Nevertheless, a complete analysis of the Witsenhausen's original counterexample is lacking, though considerable progress has been made in understanding the relation between optimal strategies and information patterns [4], [5], [6], [7], [8], [9].

The goal of this letter is to contribute to the literature which attempts to explain why Witsenhausen's problem is difficult. Our starting point is this letter [1], which considered a discrete version of the Witsenhausen counterexample where all the random variables and controls were restricted to be integers. This problem formulation can be obtained by quantizing the Witsenhausen problem and rescaling [10], [11]. Furthermore, the distribution of the initial state of the system and the noise were viewed as inputs; in Witsenhausen's original formulation, both of these were taken to be Gaussian. It was shown in [1] that computation of the optimal strategy for this version of the Witsenhausen problem is NP-complete.

While such NP-hardness results do not have any implications for Witsenhausen's original counterexample, in the generalized scenario where the initial state and noise have arbitrary distributions, they have a fairly powerful message. Indeed, let us consider what would count as a solution of the Witsenhausen problem in this more general scenario. Presumably, one would want a formula for the optimal strategy as a function of the initial state and noise distributions. However, such a formula would be quite useless if it could not be evaluated efficiently. Thus at the very least there should exist an efficient algorithm for the computation of the optimal strategy, and it is exactly this that [1] rules out.

Our goal in this letter is to strengthen the results of [1]. We seek to address the question of whether it is possible to find approximately optimal solutions to the Witsenhausen problem. It might initially seem that there are reasons to be hopeful. Indeed, the reduction in [1] reduces the Witsenhausen problem to a 3D matching problem, and, although 3D matching is NP-hard, a $4/3+\epsilon$ approximation algorithm is available for any $\epsilon > 0$ [12]. Moreover, constant factor approximation results were derived in [13] for a different, but finite dimensional problem formulation, albeit with Gaussian noises.

Unfortunately, our main result rules out the possibility of a favorable approximation with the discrete Witsenhausen problem with arbitrary initial state and noise distribution. We describe a family of examples, where the initial state is uniform over a set of n integers, and the noise is uniform over a set of at most n^2 integers, and it is NP-hard to find a strategy whose cost is upper bounded by $n^{2-\epsilon}$ times the cost of the optimal strategy, for any $\epsilon > 0$.

One might wonder if the multiplicative $n^{2-\epsilon}$ factor is the best one could do, i.e., if the problem might be even more difficult to approximate than that. In that direction, we show that if the initial distribution has support $\mathcal{X}$ and the noise distribution has support $\mathcal{Z}$, then it is always possible to approximate the optimal Witsenhausen strategy to within a multiplicative factor of $|\mathcal{X}|^3|\mathcal{Z}|^4$. Plugging in $|\mathcal{X}| = n$ and $\mathcal{Z} \leq n^2$ for the construction of the previous paragraph, we obtain that a multiplicative n^{11} approximation is possible in that case. This

Manuscript received February 12, 2019; revised April 10, 2019; accepted April 10, 2019. Date of publication April 18, 2019; date of current version April 30, 2019. This work was supported in part by NSF under Award 1740451, and in part by ARO under Award W911NF-18-1-0072. Recommended by Senior Editor J.-F. Zhang.

The author is with the Department of Electrical and Computer Engineering, Boston University, Boston, MA 02215 USA, and also with the Division of Systems Engineering, Boston University, Boston, MA 02215 USA (e-mail: alexols@bu.edu).

Digital Object Identifier 10.1109/LCSYS.2019.2911925

shows a limit to how much one could potentially improve the $n^{2-\epsilon}$-inapproximability result we described above.

The remainder of this letter is organized as follows. Section II contains technical background, including a formal definition of the Witsenhausen problem and the discretizations we described above. Section III contains a proof of the $n^{2-\epsilon}$-inapproximability result while Section IV contains a proof of the $|\mathcal{X}|^3|\mathcal{Z}|^4$ approximation result.

II. Background

We begin with an informal description of the Witsenhausen problem. Two agents attempt to stabilize a system by bringing its state close to zero in two time steps. The first agent observes the initial state X_0, which we assume to be a random variable with a known distribution. The first agent applies the control u_1, so that the state becomes $X_0 + u_1$. Now the second agent can only see a noisy version $X_0 + u_1 + Z$ of the state, where Z is some random variable with a known distribution. It applies a control u_2 which is therefore constrained to be a function of $X_0 + u_1 + Z$. The final cost depends only on the size of the control applied by the first agent as well as the final distance to the origin:

$$E\Big[u_1^2 + K(X_0 + u_1 + u_2)^2\Big],$$

where $K > 0$ is some constant. In particular, the control applied by the second agent is "free."

In the classical Witsenhausen counterexample [2], it is assumed that the initial state is $X_0 \sim N(0, \sigma_0^2)$ while the noise is $Z \sim N(0, 1)$, but in this letter we will consider arbitrary distributions for X_0, Z. We will find it convenient to reformulate the problem in a way that makes the inherent constraints explicit as follows. Given independent random variables X_0, Z we are looking for maps $T : \mathbb{R} \to \mathbb{R}$, $\delta : \mathbb{R} \to \mathbb{R}$ which minimize the cost function

$$E\Big[(T(X_0) - X_0)^2 + K(T(X_0) + \delta(T(X_0) + Z))^2\Big]. \tag{1}$$

Moreover, we will denote this quantity as $\Phi(p_{X_0}, p_Z, T, \delta)$ and refer to it as the "Witsenhausen cost." Furthermore, we will refer to $E[(T(X_0) - X_0)^2]$ as the "first-stage" or "transportation cost," while $E[[T(X_0) + \delta(T(X_0) + Z))]^2]$ will be referred to as the "second-stage" cost.

The discrete Witsenhausen problem, defined formally next, is simply the restriction of this problem to random variables and maps which take on integer values. For convenience, in the sequel we use $\mathbb{Z}$ to denote the set of integers.

Definition 1: Let X_0, Z be independent bounded random variables taking on integer values with probability mass functions p_{X_0}, p_Z. The Witsenhausen problem asks for maps $T : \mathbb{Z} \to \mathbb{Z}$ and $\delta : \mathbb{Z} \to \mathbb{Z}$ achieving the minimum in Eq. (1). We will use $\Phi^*(p_{X_0}, p_Z)$ to refer to the optimal cost as a function of the problem parameters.[1]

This problem was essentially introduced in [10]. It is not hard to see that an optimal solution exists: we can restrict our attention to a finite set of maps T, δ, as there is no need to consider maps which move some values in $\mathcal{X}$ too far. It is also standard that, given T, the corresponding δ can be found by solving a least squares problem.

Our goal in this letter is to prove the following theorem, which refines a result of [1] that the discrete Witsenhausen problem is NP-hard. We will adopt the convention of using $\mathcal{X}$ to refer to the support of X and $\mathcal{Z}$ to refer to the support of Z.

Theorem 1: Consider the discrete Witsenhausen problem restricted to problem instances where $|\mathcal{X}| = n$ and $|\mathcal{Z}| \leq n^2$. Unless $P = NP$, for any $\epsilon > 0$ there does not exist a polynomial-time algorithm which returns a number ϕ satisfying

$$\phi \leq n^{2-\epsilon}\Phi^*(p_{X_0}, p_Z)$$

such that $\phi = \Phi(p_{X_0}, p_Z, T, \delta)$ for some choice of maps T, δ.

We also show that the $n^{2-\epsilon}$ factor in the inapproximability result cannot be improved too much.

Theorem 2: There is a polynomial-time algorithm which returns T, δ satisfying

$$\Phi(p_{X_0}, p_Z, T, \delta) \leq |\mathcal{X}|^3|\mathcal{Z}|^4\Phi^*(p_{X_0}, p_Z)$$

Indeed, plugging in $|\mathcal{X}| = n, |\mathcal{Z}| = n^2$ into this last theorem, we obtain that in the setting described by Theorem 1, this provides an n^{11} multiplicative approximation.

III. Proof of Theorem 1

We now turn to a sequence of lemmas whose culmination will be the proof of Theorem 1. Our starting point is a definition which later on will be key to the way we will definite the initial state X_0.

Definition 2: An integer set S is called a Sidon set of order p if all the sums

$$s_1 + s_2 + \cdots + s_p$$

with $s_1, \ldots, s_p \in S$ and $s_1 \leq s_2 \leq \cdots \leq s_p$ are distinct.

For example, $S = \{1, 2, 4\}$ is a Sidon set of order 2 because the pairwise sums of elements from this set are distinct; but $S = \{1, 2, 3, 4\}$ is not a Sidon set of order 2 because $3 + 3 = 4 + 2$.

It is well-known that Sidon sets of arbitrary order exist and can be easily constructed. We will need the following variation of this fact, which is very similar to a lemma from [1].

Lemma 1: There exists a Sidon set S of order 4 with $|S| = k$ satisfying $S \subset \{1, 2, \ldots, 20k^8\}$. Moreover, it is possible to construct S in polynomial time in k.

Proof: We prove this lemma by induction. When $k = 1$, we can simply choose $S = \{1\}$. Now suppose we have a Sidon set $S = \{s_1, s_2, \ldots, s_k\}$, with s_i being distinct positive integers, and $\max_{i=1,\ldots,k} s_i \leq 20k^8$. We construct a Sidon set of size $k+1$ by choosing a positive integer $s_{k+1} \leq 20(k+1)^8$ to add to S.

To ensure this works, we need that

$$s_a + s_b + s_c + s_d \neq s_e + s_f + s_g + s_{k+1}$$

[1] A common convention in the literature is to specify the distribution of $Y = T(X_0) + Z$ conditioned on $T(X_0)$, but since X_0 and Z are independent here, it is easier in this case to simply specify the distribution of Z.

for all possible choices $a \leq b \leq c \leq d, e \leq f \leq g$ from $a, b, c, d, e, f, g \in \{1, \ldots, k+1\}$. In other words, we need to have

$$s_{k+1} \neq s_a + s_b + s_c + s_d - s_e - s_f - s_g.$$

There are at most $(k+1)^7$ choices of a, b, c, d, e, f, g, and each inequality produces at most one value for s_{k+1} to avoid. Indeed, it is possible for an inequality to produce no value for s_{k+1} to avoid, for example if s_{k+1} cancels from both sides. However, if this does not happen, then the corresponding choice of a, b, c, d, e, f, g results in one value for s_{k+1} to avoid.

It follows that this produces a Sidon set of order 4 if we just place s_{k+1} outside this set of at most $(k+1)^7$ values. We will place s_{k+1} in the range $(20k^8, 20(k+1)^8]$, which we can always do by the pigeonhole principle, since

$$20(k+1)^8 - 20k^8 > 20 \cdot 8k^7 > (2k)^7 \geq (k+1)^7.$$

Finally, we can construct S via the above procedure which clearly takes polynomial time. ∎

The bounds of the lemma are rather loose and it is possible to improve them, but they suffice for our purposes.

Our next step is discuss a variation on the notion of the chromatic number which we will need. Although the following discussion seems unrelated to Sidon sets, we will bring the two concepts together in our NP-hardness proof later on.

The chromatic number of a graph is the minimum number of colors needed to color the vertices so that no adjacent vertices share the same color; we will use $\kappa(G)$ to denote the chromatic number of the graph G. We will need to use a certain notion which we call the l_2-chromatic number, which as far as we know has not been previously considered, and we motivate this notion with the following discussion.

We may formulate the search for the chromatic number of the graph $G = (\{1, \ldots, n\}, E)$ as minimizing

$$\Phi(\gamma) := \max_{i=1,\ldots,n} \gamma(i) - \min_{i=1,\ldots,n} \gamma(i),$$

over all functions $\gamma : V \to \mathbb{Z}$ satisfying $\gamma(i) \neq \gamma(j)$ for all $(i, j) \in E$. Indeed, the objective $\Phi(\gamma)$ is precisely the number of colors needed minus one.

The quantity $\Phi(\gamma)$ may be thought of as a measure of dispersion. This motivated the introduction of the l_2-chromatic number, which uses a slightly different measure of dispersion: the variance of the distribution of $\gamma(i)$ about zero.

Definition 3: Given an undirected graph $G = (\{1, \ldots, n\}, E)$, the l_2-chromatic number asks for a function $\gamma : V \to \mathbb{Z}$ satisfying $\gamma(i) \neq \gamma(j)$ for all $(i, j) \in E$ and minimizing

$$\gamma^* := \frac{1}{n} \sum_{i=1}^{n} \gamma^2(i).$$

We remark that the graph G should not have any self-loops, for otherwise the constraint $\gamma(i) \neq \gamma(j)$ for all edges (i, j) in G is impossible to satisfy.

We next use the concept of Sidon sets to give a way to construct a discrete Witsenhausen problem starting from a graph. The ensuring sequence of lemmas will show that computation of the l_2-chromatic number on that graph will then be equivalent to computation of the optimal Witsenhausen strategy.

Definition 4: Given a graph $G = (\{1, \ldots, n\}, E)$ and an integer B, construct an instance of the discrete Witsenhausen as follows:

- Let $\{y_1, \ldots, y_n\}$ be Sidon set of order 4 with n elements, and set $x_i = l y_i$, where $l = 4(\lceil n^{1.5} \rceil + 1)$. Generate X_0 to be uniform over $x_1, \ldots, x_n$.
- Generate Z to be uniform over all the pairs $(x_i - x_j)/2$, $(i, j) \in E$.
- Set K to be any number strictly bigger than n^5.

Observe that the graph G enters the definition of the corresponding discrete Witsenhausen problem solely through the distribution of Z. Observe further that the support of Z always symmetric about the origin since $(i, j) \in E$ whenever $(j, i) \in E$. Note also that the support of the random variable X_0, i.e., the set $\{x_1, \ldots, x_n\}$, is a Sidon set of order 4 with n elements (because it is obtained via scaling each element of a Sidon set by the same factor l). Finally, note that this construction may be performed in polynomial time in n as a consequence of Lemma 1 which tells us that the set $\{y_1, \ldots, y_n\}$ may be constructed in polynomial time; that all remaining operations take polynomial time is obvious.

The equivalence of l_2-chromatic number on the original graph and the cost of the optimal Witsenhausen strategy in this construction is established in the following two lemmas.

Lemma 2: Suppose $0 \leq B \leq n^2$. If the discrete Witsenhausen problem constructed in Definition 4 has a solution with cost at most B, then the l_2-chromatic number of the graph G is at most B.

Proof: Let T, δ be maps which achieve a cost of at most B in the resulting Witsenhausen problem. We will define

$$\gamma(i) = T(x_i) - x_i,$$

and argue that this choice of γ works. The key observation is that, if the discrete Witsenhausen problem constructed in this way has a cost at most B, then it has zero second-stage cost, i.e., we must have with probability one that

$$T(X_0) = \delta(T(X_0) + Z). \tag{2}$$

This follows because of the way K was chosen. Formally, observe that if Eq. (2) fails with positive probability, then, because X_0, Z were constructed to be uniform over $\mathcal{X}$ and $\mathcal{Z}$, it fails with probability at least $(1/|\mathcal{X}|)(1/|\mathcal{Z}|) \geq 1/n^3$. Moreover, when Eq. (2) fails, then because both the left-hand side and the right-hand side of this equation are integer, it follows they differ by at least one. Thus, in that case the expectation of the Witsenhausen cost of Eq. (1) is at least $(1/n^3)K \cdot 1 > B$. This is a contradiction. We have thus shown that Eq. (2) holds with probability one.

In particular, this means that for all possible $x_i, x_j \in \mathcal{X}$ such that $T(x_i) \neq T(x_j)$, and all possible $z_a, z_b \in \mathcal{Z}$, we must have

$$T(x_i) + z_a \neq T\big(x_j\big) + z_b. \tag{3}$$

Indeed, if Eq. (3) fails, then it is immediate that a zero second-stage cost cannot be obtained.

We now claim that, due to the way X_0 was defined in Definition 4, we can conclude that actually $T(x_i) \neq T(x_j)$ for all pairs $i, j = 1, \ldots, n$, so that the conclusion of the previous paragraph is actually applicable to all pairs i, j. Indeed, suppose $T(x_i) = T(x_j)$ for some pair i, j. Since $|x_i - x_j| > 4n^{1.5}$, we have that either $|T(x_i) - x_i| > 2n^{1.5}$ or $|T(x_j) - x_j| > 2n^{1.5}$. Either one of these will imply the first-stage transportation cost is strictly bigger than n^2 and thus strictly bigger than B.

Putting the last two paragraphs together, we have that for all realizations $x_i, x_j, \in \mathcal{X}, z_i, z_j \in \mathcal{Z}$, we have that

$$T(x_i) + z_a \neq T\big(x_j\big) + z_b.$$

In particular,

$$T(x_i) - T\big(x_j\big) \neq z_b - z_a,$$

or

$$T(x_i) - x_i - \big(T\big(x_j\big) - x_j\big) \neq z_b - z_a - x_i + x_j.$$

But if (i, j) is an edge in G, then the right-hand side of this equations equals zero when

$$z_b = \frac{x_i - x_j}{2}, z_a = -z_b,$$

and these are both in $\mathcal{Z}$. So we conclude that if i and j are neighbors in G, then

$$T(x_i) - x_i - \big(T\big(x_j\big) - x_j\big) \neq 0$$

or $\gamma(i) \neq \gamma(j)$. Thus $\gamma(i)$ satisfies the constraint in the definition of the l_2-chromatic number (i.e., Definition 3).

Finally, we observe that

$$\gamma^* \leq \frac{1}{n}\sum_{i=1}^{n} \gamma(i)^2 = \frac{1}{n}\sum_{i=1}^{n} (T(x_i) - x_i)^2,$$

and because the second-stage cost is zero, this is equal to the expected Witsenhausen cost, which is at most B by assumption. ∎

Note that Lemma 2 did not use that the support of X_0 is a Sidon set. The next lemma, which is just the converse of Lemma 2, will use this fact.

Lemma 3: Suppose $1 \leq B \leq n^2$. If the l_2-chromatic number of G is at most B, then the discrete Witsenhausen problem constructed in Definition 4 has a solution of cost at most B.

Before we give a proof of this lemma, we require the following fact.

Lemma 4: $(x_i - x_j)/2 \in \mathcal{Z}$ if and only if (i, j) is an edge in G.

Proof: One direction is one immediate from Definition 4. On the other hand, suppose $(x_i - x_j)/2 \in \mathcal{Z}$. This means there exist neighbors a, b in G such that

$$\frac{x_i - x_j}{2} = \frac{x_a - x_b}{2}$$

or

$$x_i + x_i + x_b + x_b = x_a + x_a + x_j + x_j$$

Since $x_a \neq x_b$ and $\{x_1, \ldots, x_n\}$ is a Sidon set of order four, this implies that $x_i = x_a, x_b = x_j$. Thus i and j are neighbors. ∎

Proof of Lemma 3: Paralleling the proof of Lemma 2, we define

$$T(x_i) = x_i + \gamma(i),$$

where γ is the coloring that achieves l_2-chromatic number at most B. For integers $x' \notin \{x_1, \ldots, x_n\}$, we can define $T(x')$ arbitrarily, as it does not affect the Witsenhausen cost. We will show that, with this choice, the second-stage cost is zero. Once this is shown, the proof will be complete as the l_2-chromatic number $(1/n)\sum_i \gamma^2(i)$ is just the transportation cost.

To argue that the second stage cost is zero, we proceed by contradiction. The second stage cost is not zero if and only if there exist $x_i, x_j \in \mathcal{X}, z_a, z_b \in \mathcal{Z}$ with $T(x_i) \neq T(x_j)$ such that

$$T(x_i) + z_a = T\big(x_j\big) + z_b \tag{4}$$

But, as in Lemma 2, we cannot have $x_i \neq x_j$ with $T(x_i) = T(x_j)$; indeed, by the same argument as Lemma 2, this implies that $|T(x_i) - x_i| > 2n^{1.5}$, which now contradicts the fact that γ achieves l_2-chromatic number at most $B \leq n^2$. So the second stage cost is zero if and only if there exist $x_i, x_j \in \mathcal{X}, x_i \neq x_j, z_a, z_b \in \mathcal{Z}$ such that Eq. (4) is satisfied. Now observe we can write Eq. (4)

$$x_i + \gamma(i) + z_a = x_j + \gamma(j) + z_b$$

or

$$x_i + z_a - x_j - z_b = \gamma(j) - \gamma(i). \tag{5}$$

Now the way Z was constructed in Definition 4 means that there exist neighbors c, d and neighbors e, f such that

$$z_a = \frac{x_c - x_d}{2}, \quad z_b = \frac{x_e - x_f}{2}.$$

Plugging this into Eq. (5) and doubling both sides,

$$2x_i + x_c - x_d - 2x_j - x_e + x_f = 2\big[\gamma(j) - \gamma(i)\big]$$

or

$$\big(x_i + x_i + x_c + x_f\big) - \big(x_j + x_j + x_d + x_e\big) = 2(\gamma(j) - \gamma(i)) \tag{6}$$

Now we consider two possibilities both of which lead to a contradiction. The left-hand side of Eq. (6) is either zero or nonzero.

If it is zero, then since $x_i \neq x_j$, and $\{x_1, \ldots, x_n\}$ being a Sidon set of order 4, we must have

$$x_i = x_d = x_e \quad \text{and} \quad x_j = x_c = x_f.$$

But this means that $(x_j - x_i)/2 = (x_c - x_d)/2 \in \mathcal{Z}$ so that by Lemma 3 we have that i and j are neighbors. But since the left-hand side of Eq. (6) is zero, we have that $\gamma(j) = \gamma(i)$ for a pair of neighbors i, j, a contradiction.

On the other hand, if the left-hand side of Eq. (6) is nonzero, then, since every x_i is a multiple of l by construction (recall Definition 4), the same left-hand side must have absolute value at least l. It follows that

$$|\gamma(j) - \gamma(i)| \geq \frac{l}{2} > 2n^{1.5},$$

where the strict inequality used the definition of l. Thus at least one of $|\gamma(i)|, |\gamma(j)|$ is strictly bigger than $n^{1.5}$. But this contradicts that the l_2-chromatic number is at most $B \leq n^2$. This concludes the proof. ∎

We now turn to an analysis of the l_2-chromatic number. We begin with a lemma which shows that the l_2-chromatic number is not very far from the ordinary chromatic number. Recall that we use the notation $\kappa(G)$ for the ordinary chromatic number of G.

Lemma 5:

$$\kappa^2 \geq \gamma^* \geq \frac{1}{n}\frac{(\kappa-2)^3}{12}.$$

Proof: For the first inequality, simply consider taking $\gamma(i)$ to be the color of vertex i, represented by an integer in the set $\{1, \ldots, \kappa\}$, using a coloring that minimizes the number of colors used.

For the second inequality, consider the optimal γ in the definition of l_2-coloring. Let us translate the γ so that the smallest interval I containing its range is symmetric about the origin, i.e., it equals either $[-a, a]$ or $[-a, a+1]$. Observe that every element in I is used, i.e., every element in I equals $\gamma(i)$ for some i, for else it would be possible to obtain a γ with smaller l_2 chromatic number. This implies that

$$\gamma^* \geq 2\frac{1}{n}\left(1^2 + \cdots + a^2\right) \geq \frac{2}{3}\frac{a^3}{n}.$$

On the other hand, the chromatic number is at most $2(a+1)$. Thus

$$\kappa \leq 2a + 2 \leq 2\left((3/2)n\gamma^*\right)^{1/3} + 2$$

or

$$(\kappa - 2)^3 \leq 12n\gamma^*,$$

which is a rearrangement of the second inequality. ∎

Lemma 6: Unless $P = NP$, for any $\epsilon > 0$, there exists no polynomial time algorithm which, given an undirected graph on n vertices, returns a number between γ^* and $n^{2-\epsilon}\gamma^*$.

Proof: It is possible to define the notion of a *fractional chromatic number* of a graph G, denoted by $\chi_f(G)$. We avoid giving a definition here because we only need to use the following two facts about it:

- In [14, Th. 1.2], it was shown that, for any $\epsilon > 0$, it is NP-hard to distinguish between graphs G on n vertices with fractional chromatic number of n^{ϵ} from graphs with fractional chromatic number of $n^{1-\epsilon}$.
- In [15], it was shown that the fractional chromatic number is a logarithmic approximation to the chromatic number, i.e.,

$$\frac{\kappa(G)}{1+\log n} \leq \chi_f(G) \leq \kappa(G),$$

where, recall, $\kappa(G)$ is the ordinary chromatic number; for more details, see the discussion in [16, Sec. 3.3].

As remarked in [16], these two facts imply that it is NP-hard to distinguish between graphs of chromatic number $n^{\epsilon}(1+\log n)$ and graphs with chromatic number $n^{1-\epsilon}$.

Algorithm 1 Approximation Algorithm for the Discrete Witsenhausen Problem

1: Input: distributions of X_0, Z
2: **for** $k = 0, \ldots, n$ **do**
3: Set T^k be a map that map that fixes $x_1, \ldots, x_k$ and maps $x_{k+1}, \ldots, x_n$ to values ensuring there are no collisions except between $x_1, \ldots, x_k$.
4: Choose δ^k to be the optimal second-stage map given T^k.
5: **end for**
6: Choose the pair among $(T^k, \delta^k), k = 1, \ldots, n$ with lowest Witsenhausen cost.

Applying Lemma 5, it follows that it is NP-hard to distinguish between graphs with $\gamma^* \leq n^{2\epsilon}(1+\log n)^2$ and graphs with $\gamma^* \geq \frac{1}{12n}(n^{1-\epsilon}-2)^3$. We conclude that, for any $\epsilon > 0$, it is NP-hard to approximate γ^* within a multiplicative factor of less than

$$\frac{(n^{1-\epsilon}-2)^3}{12n^{1+2\epsilon}(1+\log n)^2}.$$

Because this quantity can be lower bounded by $n^{2-O(\epsilon)}$, this completes the proof. ∎

Finally, we are now able to provide a proof of our main result.

Proof of Theorem 1: Consider a graph G with l_2-chromatic number of B. Since every vertex can be colored by a different color, we have that $B \leq n^2$. Consider the discrete Witsenhausen problem constructed in Definition 4: putting together Lemma 2 and Lemma 3, we obtain that its optimal solution has cost B. Now observing that by Lemma 6, it is NP-hard to approximate B to within a multiplicative factor of $n^{2-\epsilon}$ completes the proof. ∎

IV. Proof of Theorem 2

We now describe an algorithm for the Witsenhausen problem whose approximation ratio is polynomial in $|\mathcal{X}|$ and $|\mathcal{Z}|$. We begin with an informal discussion intended to motivate our approach. Parallelling our arguments in the previous section, we'll adopt the convention of saying that i and j "collide" if

$$T(x_i) + z_a = T\left(x_j\right) + z_b, \tag{7}$$

for some $z_a, z_b \in \mathcal{Z}$.

Our approach is simple: we "interpolate" between the optimal solution when $K = 0$ (which results in $T(x_i) = x_i$) and $K \to +\infty$ (which results in a T that avoids any collisions) by fixing the k elements in $\mathcal{X}$ with the highest probabilities, and moving all the other entries in $\mathcal{X}$ to avoid collisions. We do this for all $k = 1, \ldots, n$ where $n = |\mathcal{X}|$ and choose the best result.

We outline the approach in the algorithm box below, where we use the convention that p_i is the probability of $X_0 = x_i$ and

$$p_1 \geq p_2 \geq \cdots \geq p_n.$$

It is easy to see that this is a polynomial-time algorithm. Indeed, step 6 can easily be done in polynomial time: the

cost of each pair T^i, δ^i is a sum over $|\mathcal{X}||\mathcal{Z}|$ values. Second, step 4 can also be done without difficulty, since the selection of the best second-stage map given the transportation map is an ordinary least-squares estimation problem. The following lemma discusses how to do step 3 and implicitly gives an upper bound on the transportation cost of the T^i chosen in that step.

Lemma 7: Step 3 can be done in polynomial time with $|T^k(x_j) - x_j| \leq |\mathcal{X}||\mathcal{Z}|^2$ for all j.

Proof: Starting with $j = k+1$, we sequntially set $T(x_j)$ to be the closest value to x_j that does not yield a collision; when we have set $T(x_n)$, we are done. When we consider x_m, looking at Eq. (7), we have to avoid

$$T(x_m) = T\big(x_j\big) + z_b - z_a, \quad j < m, \ \ z_a, z_b \in \mathcal{Z},$$

which rules out at most $(m-1)|\mathcal{Z}|^2$ different values. It follows that we can always assign $T(x_m)$ so that $|T(x_m) - x_m| \leq |\mathcal{X}||\mathcal{Z}|^2$. Moreover, each step of this procedure requires examining at most $|\mathcal{X}||\mathcal{Z}|^2$ possibilities, and the number of steps is at most $|\mathcal{X}|$, so this procedure is polynomial time. ■

We can now proceed to the proof of Theorem 2. Our first step is to introduce some notation. We let $\Phi_1(p_{X_0}, p_Z, T, \delta)$ to be the first-stage (transporation) cost when X_0, Z, T, δ are the random variables and maps in the discrete Witsenhausen problem. Likewise, we will use $\Phi_2(p_{X_0}, p_Z, T, \delta)$ to denote the second-stage cost. Occasionally, we will omit to write the δ in this notation, and it should be understood that δ is then selected to be the optimal choice for the given T.

Proof of Theorem 2: We claim that Algorithm 1 with the selection procedure of Lemma 7 returns a solution with cost $|\mathcal{X}|^3|\mathcal{Z}|^4\Phi^*$ where Φ^* is the optimal Witsenhausen cost.

Indeed, consider the optimal strategy T^*, δ^*. Let l be the smallest index such that $T^*(x_l) \neq x_l$ (we can assume such an index exists, because otherwise Algorithm 1 finds the optimal solution when $k = n$ and there is nothing to prove). The transport cost incurred by T^* is at least p_l.

Now consider the (T^k, δ^k) when $k = l$. The transport cost incurred by T^l is upper bounded by $(|\mathcal{X}|p_l)(|\mathcal{X}||\mathcal{Z}|^2)^2$ because the probability of not landing at a fixed point is at most $|\mathcal{X}|p_l$, in which case one moves by at most $|\mathcal{X}||\mathcal{Z}|^2$ as a consequence of Lemma 7. Thus the transport cost incurred by T^l is at most $p_l|\mathcal{X}|^3|\mathcal{Z}|^4$. Putting the last two paragraphs together,

$$\Phi_1\Big(p_{X_0}, p_Z, T^l, \delta^l\Big) \leq |\mathcal{X}|^3|\mathcal{Z}|^4\Phi_1^*\big(p_{X_0}, p_Z\big) \tag{8}$$

We now consider the second-stage cost of T^l, δ^l. By construction whenever one of $(x_{l+1}, \ldots, x_n)$ is generated, the second-stage cost is zero. Defining p' to be the distribution proportional to $(p_1, p_2, \ldots, p_{l-1})$, this means that

$$\Phi_2\Big(p_{X_0}, p_Z, T^l, \delta^l\Big) = (p_1 + \cdots p_{l-1})\Phi_2\big(p', p_Z, I\big). \tag{9}$$

where we use I for the identity map and we used that T^l fixes $x_1, \ldots, x_{l-1}$.

On the other hand, consider the second stage cost under T^*, δ^*. Let A be the event that $X_0 \in \{x_1, \ldots, x_{l-1}\}$. The second-stage cost cannot be increased if the first agent transmits to the second agent whether A has occurred or not. Thus

$$\Phi_2\big(p_{X_0}, p_Z, T^*, \delta^*\big) \geq (p_1 + \cdots p_{l-1})\Phi_2\big(p', p_Z, I\big). \tag{10}$$

Finally, comparing Eq. (9) and Eq. (10) we obtain $\Phi_2(p_{X_0}, p_Z, T^*, \delta^*) \geq \Phi_2(p_{X_0}, p_Z, T^l, \delta^l)$. Putting this together with Eq. (8) completes the proof. ■

Acknowledgment

The author would like to thank Dr. S. Kopparty and Dr. A. Bhangale for suggesting the reduction between the l_2-chromatic number and the ordinary chromatic number. The author would also like to acknowledge Dr. M. Agarwal for multiple discussions of this problem.

References

[1] C. H. Papadimitriou and J. N. Tsitsiklis, "Intractable problems in control theory," *SIAM J Control Optim.*, vol. 24, no. 4, pp. 639–654, 1986.
[2] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM J Control*, vol. 6, no. 1, pp. 131–147, 1968.
[3] S. Yüksel and T. Basar, *Stochastic Networked Control Systems*. New York, NY, USA: Birkhäuser, 2013.
[4] T. Basar, "Variations on the theme of the Witsenhausen counterexample," in *Proc. 47th IEEE Conf. Decis. Control (CDC)*, 2008, pp. 1614–1619.
[5] C. A. Uribe, T. Keviczky, and J. H. van Schuppen, "Computing optimal control laws for finite stochastic systems with non-classical information patterns," in *Proc. IEEE Amer. Control Conf. (ACC)*, 2014, pp. 5742–5747.
[6] A. Gupta, S. Yüksel, T. Başar, and C. Langbort, "On the existence of optimal policies for a class of static and sequential dynamic teams," *SIAM J Control Optim.*, vol. 53, no. 3, pp. 1681–1712, 2015.
[7] A. A. Kulkarni and T. P. Coleman, "An optimizer's approach to stochastic control problems with nonclassical information structures," *IEEE Trans. Autom. Control*, vol. 60, no. 4, pp. 937–949, Apr. 2015.
[8] S. T. Jose and A. A. Kulkarni, "A linear programming relaxation for stochastic control problems with non-classical information patterns," in *Proc. IEEE 54th Annu. Decis. Control Conf. (CDC)*, 2015, pp. 5743–5748.
[9] A. Nayyar, A. Mahajan, and D. Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," *IEEE Trans. Autom. Control*, vol. 58, no. 7, pp. 1644–1658, Jul. 2013.
[10] Y. Ho and T. Chang, "Another look at the nonclassical information structure problem," *IEEE Trans. Autom. Control*, vol. 25, no. 3, pp. 537–540, Jun. 1980.
[11] N. Saldi, S. Yüksel, and T. Linder, "Finite model approximations and asymptotic optimality of quantized policies in decentralized stochastic control," *IEEE Trans. Autom. Control*, vol. 62, no. 5, pp. 2360–2373, May 2017.
[12] M. Cygan, "Improved approximation for 3-dimensional matching via bounded pathwidth local search," in *Proc. IEEE 54th Annu. Symp. Found. Comput. Sci. (FOCS)*, 2013, pp. 509–518.
[13] P. Grover, A. Sahai, and S. Y. Park, "The finite-dimensional Witsenhausen counterexample," in *Proc. 7th Int. Symp. Model. Optim. Mobile Ad Hoc Wireless Netw. (WiOPT)*, 2009, pp. 1–10.
[14] D. Zuckerman, "Linear degree extractors and the inapproximability of max clique and chromatic number," in *Proc. 38th Annu. ACM Symp. Theory Comput.*, 2006, pp. 681–690.
[15] L. Lovász, "On the ratio of optimal integral and fractional covers," *Disc. Math.*, vol. 13, no. 4, pp. 383–390, 1975.
[16] U. Feige and J. Kilian, "Zero knowledge and the chromatic number," *J. Comput. Syst. Sci.*, vol. 57, no. 2, pp. 187–199, 1998.