UNEQUAL ERROR PROTECTION QUERYING POLICIES FOR THE NOISY 20 QUESTIONS PROBLEM

Hye Won Chung[1], *Brian M. Sadler*[2], *Lizhong Zheng*[3], *Alfred O. Hero*[4]

[1]KAIST, [2]U.S. Army Research Laboratory, [3]MIT, [4]University of Michigan

ABSTRACT

We propose a non-adaptive unequal error protection (UEP) querying policy based on superposition coding for the noisy 20 questions problem. In this problem, a player wishes to successively refine an estimate of the value of a continuous random variable by posing binary queries and receiving noisy responses. When the queries are designed non-adaptively as a single block and the noisy responses are modeled as the outputs of a binary symmetric channel the 20 questions problem can be mapped to an equivalent problem of channel coding with UEP. A new non-adaptive querying strategy based on UEP superposition coding is introduced whose estimation error decreases with an exponential rate of convergence that is significantly better than that of the UEP repetition coding introduced by Variani *et al.* (2015). In fact, we show that the proposed non-adaptive UEP querying policy achieves the same order convergence rate as the adaptive policy.

Index Terms— Noisy 20 questions problem, estimation, superposition coding, unequal error protection.

1. INTRODUCTION

Consider a noisy 20 questions game between a player and an oracle. The objective of the player is to estimate the value of a continuous target variable $X \sim \text{unif}[0, 1]$. The player asks binary queries to the oracle who knows the value of X, and receives a noisy version of the oracle's correct answers transmitted through a binary symmetric channel with flipping probability $\epsilon \in (0, 1/2)$, denoted BSC(ϵ). The sequence of queries is designed by a controller that may either operate open-loop (non-adaptive 20 questions) or use feedback (adaptive 20 questions). The central question is: What is the optimal sequence of queries to estimate the value of X with a minimum estimation error for a fixed number of queries? This general setup of the noisy 20 questions game and the optimal query design problem is of broad interest, arising in various areas, including active learning [2, 3], optimal sensing [4] and experimental design [5, 6], with diverse applications. For example, a target localization problem in a sensor network [7] can be modeled as a noisy 20 questions game where a player (agency) aims to locate a target by receiving query responses from sensors probing the region of interest.

In the noisy 20 questions problem, estimates of the coefficients in dyadic expansion of the variable $X \approx 0.B_1B_2 \dots B_k$ may contain errors. Since the errors in the more significant bits (MSBs) cause a higher estimation error than do the errors in the less significant bits (LSBs), it is desirable to provide unequal error protection for MSBs vs. LSBs in order to minimize the estimation error with a limited number of queries.

One way to provide unequal error protection is repetition coding. In repetition coding, each bit is queried multiple times, with the number of repetitions varying in accordance with the desired level of unequal error protection. Such a UEP repetition coding approach to the noisy 20 questions problem was considered in [8]. It was shown that the mean squared error (MSE) of this approach decreases as order $O(e^{-c_1\sqrt{N}})$, $c_1 > 0$, where N is the number of queries. This square root of N rate is slower than order $O(e^{-c_2 N})$ rate with some $c_2 > 0$, achievable by the bisection-based adaptive 20 questions strategy [9] that corresponds to Horstein's coding scheme for a BSC(ϵ) with perfect feedback [10].

The main contribution of this paper is to provide a new non-adaptive querying strategy based on superposition coding [11] that can provide UEP for two levels of priority, i.e., a strictly better error protection for MSBs than that for LSBs, and can achieve better MSE convergence rate than that of repetition coding in [8]. We show that the proposed querying strategy achieves MSE that decreases exponentially in N, as contrasted to $\sqrt{N}$, matching the error rate of the adaptive 20 questions strategy [9]. Furthermore, this strategy achieves a better scale factor in the MSE exponent as compared to that of random block codes employing equal error protection.

2. PROBLEM STATEMENT

To estimate an unknown random variable $X \sim \text{unif}[0, 1]$, a player asks an oracle whether X is located within some subregion $Q_i \subset [0, 1]$, either connected or non-connected. The oracle gives the correct answer $Z_i(X) = \mathbb{1}(X \in Q_i)$, and

E-mails: Hye Won Chung (hwchung@kaist.ac.kr), Brian M. Sadler (brian.m.sadler6.civ@mail.mil), Lizhong Zheng (lizhong@mit.edu), and Alfred O. Hero (hero@eecs.umich.edu). This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIT) (No. 2017R1E1A1A01076340) and by U.S. Army Research Office (ARO) grants W911NF-11-1-0391 and W911NF-15-1-0479. A longer version of this paper with full proofs is available [1].

the player receives a noisy version $Y_i \in \{0,1\}$ of the oracle's answer transmitted through a BSC(ϵ), which flips the oracle's answer with probability $\epsilon \in (0, 1/2)$. After receiving $(Y_1, \ldots, Y_N)$, the player calculates an estimate $\hat{X}_N$ of X. We consider two types of estimation errors. The first is the mean squared error (MSE) $\mathbb{E}[|X - \hat{X}_N|^2]$. The second is the quantized MSE $\mathbb{E}[c_{\mathsf{q}}(X, \hat{X}_N)]$ where the cost function $c_{\mathsf{q}}(X, \hat{X}_N)$ with 2^k levels is a stepwise function defined as

$$\begin{aligned} & c_{\mathsf{q}}(X, \hat{X}_N) = (d2^{-k})^2, \text{ when} \\ & d2^{-k} - 2^{-k}/2 < |X - \hat{X}_N| \leq d2^{-k} + 2^{-k}/2, \\ & \text{for } d \in \{0, \ldots, 2^k - 1\}, \end{aligned} \quad (1)$$

for $X, \hat{X}_N \in [0, 1]$. We consider the cost function (1) when the objective of the problem is to estimate the value of X up to the first k bits $(B_1, \ldots, B_k)$ in the dyadic expansion of X.

The problem of optimal query design can be mapped to an equivalent problem of channel coding. For non-adaptive block querying, we suppose that the player estimates X by querying about the first k bits in the dyadic expansion of $X \approx 0.B_1 \ldots B_k$ where $k = NR/\ln 2$ for a fixed rate $R > 0$. Discovering $(B_1, \ldots, B_k)$ is equivalent to finding the index $M = \sum_{i=1}^{k} B_i 2^{k-i} \in \{0, \ldots, 2^k - 1\}$ of the interval $I_M := [M2^{-k}, (M+1)2^{-k})$ containing the value of the target variable, $X \in I_M$. Here the region of interest $[0, 1]$ is uniformly quantized into 2^k disjoint sub-intervals $\{I_0, \ldots, I_{2^k - 1}\}$ of length 2^{-k}. Assume that each querying region Q_i is a union of some subset of quantized intervals $\{I_0, \ldots, I_{2^k-1}\}$. By considering the index $M \in \{0, \ldots, 2^k - 1\}$ as a message transmitted from the oracle to the player and the oracle's answers $(Z_1, \ldots, Z_N)$ to the block of queries $(Q_1, \ldots, Q_N)$ as a codeword, block querying can be mapped to an equivalent problem of block channel coding over a BSC(ϵ). When $\mathbf{z}^{(m)} = (z_1^{(m)}, \ldots, z_N^{(m)})$ is a length-N codeword for a message $m \in \{0, \ldots, 2^k - 1\}$, the associated i-th querying region Q_i becomes the union of the sub-intervals $\{I_{m'}\}$ for message m''s such that the i-th answer bit $z_i^{(m')}$ equals 1. Therefore, the encoder specifies a block of questions, and vice versa. After receiving N noisy answers $(Y_1, \ldots, Y_N)$, the player generates estimates $(\hat{B}_1, \ldots, \hat{B}_k)$ of $(B_1, \ldots, B_k)$ and $\hat{M} = \sum_{i=1}^{k} \hat{B}_i 2^{k-i}$ of the message M.

For the finite resolution estimator defined by $\hat{X}_{N,\mathsf{finite}} := \hat{M}2^{-k} + 2^{-k}/2$, the MSE $\mathbb{E}[|X - \hat{X}_{N,\mathsf{finite}}|^2]$ can be written as a sum of the quantized MSE $\mathbb{E}[c_{\mathsf{q}}(X, \hat{X}_{N,\mathsf{finite}})]$ and the error from finite resolution,

$$\mathbb{E}[|X - \hat{X}_{N,\mathsf{finite}}|^2] = \mathbb{E}[c_{\mathsf{q}}(X, \hat{X}_{N,\mathsf{finite}})] + c2^{-2k}, \quad (2)$$

for some constant $0 < c \leq 1/4$. Assuming that the resolution k scales as $k = NR/\ln 2$ in the number N of queries for some fixed positive rate $R > 0$, we denote by $E^*_{\mathsf{MSE,policy}}(R)$ and $E^*_{\mathsf{q,policy}}(R)$ the best achievable exponentially decreasing rates of the MSE and of the quantized MSE in N at a fixed rate R, respectively, for some policy, i.e.,

$$E^*_{\mathsf{MSE,policy}}(R) := \liminf_{N \to \infty} \frac{-\ln \mathbb{E}[|X - \hat{X}_{N,\mathsf{finite}}|^2]}{N}, \quad (3)$$

$$E^*_{\mathsf{q,policy}}(R) := \liminf_{N \to \infty} \frac{-\ln \mathbb{E}[c_{\mathsf{q}}(X, \hat{X}_{N,\mathsf{finite}})]}{N}. \quad (4)$$

Then the equality in (2) implies that as $N \to \infty$, the exponential convergence rate of the MSE $\mathbb{E}[|X - \hat{X}_{N,\mathsf{finite}}|^2]$ in N is dominated by the minimum between the exponentially decreasing rate of the quantized MSE and 2R, i.e.,

$$E^*_{\mathsf{MSE,policy}}(R) = \min\{E^*_{\mathsf{q,policy}}(R), 2R\}. \quad (5)$$

In this paper, we analyze the performance of the querying policy by first calculating the best achievable quantized-MSE exponent $E^*_{\mathsf{q,policy}}(R)$ at a fixed rate $R > 0$ for querying resolution of $k = NR/\ln 2$ bits, and then calculating the corresponding MSE exponent $E^*_{\mathsf{MSE,policy}}(R)$ by using (5).

Note that with the finite-resolution estimator $\hat{X}_{N,\mathsf{finite}} = \hat{M}2^{-k} + 2^{-k}/2$, the quantized MSE can be bounded above as

$$\mathbb{E}[c_{\mathsf{q}}(X, \hat{X}_{N,\mathsf{finite}})] \leq \sum_{i=1}^{k} \Pr(\hat{B}_i \neq B_i) 2^{-2(i-1)}, \quad (6)$$

by using $\mathbb{E}[|X - \hat{X}_N|^2 \,|\, \hat{B}_i \neq B_i, \hat{B}_1^{i-1} = B_1^{i-1}] \leq 2^{-2(i-1)}$. This upper bound shows how each bit-error probability contributes to the estimation error. As the bit position i increases corresponding to lower significance, the weights on the bit error probabilities decrease exponentially in i. In order to minimize the upper bound on the the quantized MSE for a fixed number of querying N, we need to design a querying strategy (or the associated channel coding) that can provide unequal error protection depending on the bit positions. [1]

3. PREVIOUS APPROACHES

We review two well-known non-adaptive querying policies.

3.1. Non-adaptive UEP Repetition Policy

For the non-adaptive UEP repetition policy [8], each raw bit B_i of $M = (B_1, B_2, \ldots, B_k)$ is repeatedly queried N_i times, where the total number of queries is restricted to $\sum_{i=1}^{k} N_i = N$. When the answer bit B_i is transmitted N_i times through a BSC(ϵ), a simple majority voting algorithm is the maximum likelihood (ML) decoder for B_i achieving $\Pr(\hat{B}_i \neq B_i) \leq e^{-N_i \cdot D_{\mathsf{B}}(1/2 \| \epsilon)}$. By assigning different numbers of repetitions $(N_1, N_2, \ldots N_k)$ for each information bit B_i, we can provide

[1] Notation: Let the bold face $\mathbf{z}$ or z_1^N denote the length-N binary sequence $(z_1 z_2 \ldots z_N)$ where z_t is the t-th bit of $\mathbf{z}$. The entropy of a binary random variable X distributed as Bernoulli(α) is denoted $H_{\mathsf{B}}(\alpha)$. The Kullback-Leibler divergence between two Bernoulli distributions Bernoulli(α) and Bernoulli(β) is denoted $D_{\mathsf{B}}(\alpha \| \beta)$.

unequal error protection. In [8], it was shown that the minimal MSE achievable by UEP repetition coding with the optimal choice of $(N_1, \dots, N_k)$ and k decreases exponentially in $\sqrt{N}$ but not faster than that

$$c_1 e^{-c_2\sqrt{N}} \le \min_{(N_1,\dots,N_k),k} \mathbb{E}[|X - \hat{X}_{N,\mathsf{finite}}|^2] \le c_3 e^{-c_4\sqrt{N}}, \tag{7}$$

for some positive constants $c_1, c_2, c_3, c_4 > 0$. This implies that the best achievable quantized-MSE exponent of repetition coding is $E^*_{\mathsf{q,repetition}}(R) = 0$ at any positive rate $R > 0$.

3.2. Non-adaptive Random Coding Policy

We also consider a non-adaptive block-querying strategy based on random block coding [12]. The encoding map $f : \{0, \dots, 2^k - 1\} \to \{0, 1\}^N$ of the random block codes of rate $R = k \ln 2/N$ independently generates length-N codewords $\mathbf{z}^{(m)} = (z_1^{(m)}, \dots, z_N^{(m)}) := f(m)$ each of which is composed of i.i.d. symbols following the Bernoulli(1/2) distribution. This is equivalent to independently choosing a querying region at each round as the union of a subset of intervals $\{[m'2^{-k}, (m'+1)2^{-k}) : m' \in \{0, \dots, 2^k - 1\}\}$, each of which is randomly included in the querying region with probability 1/2. With this policy, it is guaranteed that the quantized MSE of random block coding decreases exponentially in N with exponent

$$E^*_{\mathsf{q,rc}}(R) = E_{\mathsf{r}}(R) \tag{8}$$

for

$$E_{\mathsf{r}}(R) = \begin{cases} E_0(1/2, \epsilon) - R, & 0 \le R < R_{\mathsf{crit}}(\epsilon), \\ D_{\mathsf{B}}(\gamma_{\mathsf{GV}}(R) \| \epsilon), & R_{\mathsf{crit}}(\epsilon) \le R \le C, \end{cases} \tag{9}$$

where $E_0(1/2, \epsilon) = -\ln(1/2 + \sqrt{\epsilon(1-\epsilon)})$, $R_{\mathsf{crit}}(\epsilon) = D_{\mathsf{B}}(\gamma_{\mathsf{crit}}(\epsilon) \| 1/2)$ with $\gamma_{\mathsf{crit}}(\epsilon) = \frac{\sqrt{\epsilon}}{\sqrt{\epsilon} + \sqrt{1-\epsilon}}$, $C = H_{\mathsf{B}}(1/2) - H_{\mathsf{B}}(\epsilon)$, and $\gamma_{\mathsf{GV}}(R)$ is the normalized Gilbert-Varshamov distance, defined such that $D_{\mathsf{B}}(\gamma_{\mathsf{GV}}(R) \| 1/2) = R$.

4. NON-ADAPTIVE BLOCK QUERYING BASED ON SUPERPOSITION CODING

In this section, we propose a new non-adaptive block querying policy based on superposition coding [11]. By using superposition coding, we design a querying strategy that provides a better error protection for MSBs than for LSBs in the dyadic expansion of the target variable $X \approx 0.B_1 B_2 \dots B_k$. By unequally distributing a fixed amount of querying resource to the MSBs and LSBs of the target variable, this querying strategy achieves better MSE convergence rates than that of random block coding, which distributes the querying resource equally to $(B_1, B_2, \dots, B_k)$.

We first partition the information bits $(B_1, \dots, B_k)$ into two sub-groups, a group containing the first $k_1 < k$ bits of X $(B_1, \dots, B_{k_1})$ and the other group containing the remaining $k_2 := k - k_1$ bits of X $(B_{k_1+1}, \dots, B_{k_1+k_2})$. The group of MSBs $(B_1, \dots, B_{k_1})$ determines the more important partial message $M_1 \in \{0, \dots, 2^{k_1} - 1\}$, while the group of LSBs $(B_{k_1+1}, \dots, B_{k_1+k_2})$ determines the less important partial message $M_2 \in \{0, \dots, 2^{k_2} - 1\}$. Denote the rates of M_1 (MSBs) and of M_2 (LSBs) by $R_1 = (k_1 \ln 2)/N$ and $R_2 = (k_2 \ln 2)/N$, respectively. Superposition codes are constructed by superimposing two types of random block codes generated by different distributions. The first type of random block codes of length N and rate R_1 is composed of e^{NR_1} binary length-N codewords, $\{\mathbf{u}^{(m_1)}\}$, $m_1 \in \{0, \dots, e^{NR_1} - 1\}$, which encode the more important partial message m_1 (MSBs). The symbols of every codeword are chosen independently at random with Bernoulli(1/2) distribution. We call these partial codewords "cloud centers" in the output space $\{0,1\}^N$. The second type of random block codes of length N and rate R_2 is composed of codewords $\{\mathbf{v}^{(m_2)}\}$, $m_2 \in \{0, \dots, e^{NR_2} - 1\}$, which encode the less important partial message m_2 (LSBs). Every symbol of every codeword in $\{\mathbf{v}^{(m_2)}\}$ is i.i.d. with Bernoulli(α) distribution for a fixed $\alpha \in (0, 1/2)$. The superposition codes $\mathcal{C}_s$ of rate $R = R_1 + R_2$ are composed of $\{\mathbf{z}^{(m_1,m_2)}\}$ for messages $(m_1, m_2) \in \{0, \dots, e^{NR_1} - 1\} \times \{0, \dots, e^{NR_2} - 1\}$, where $\mathbf{z}^{(m_1,m_2)} = \mathbf{u}^{(m_1)} \oplus \mathbf{v}^{(m_2)}$ and $\oplus$ is the bitwise XOR. The codewords $\{\mathbf{z}^{(m_1,m_2)}\}$ for a fixed m_1 are called "satellite codewords" for the respective cloud center $\mathbf{u}^{(m_1)}$.

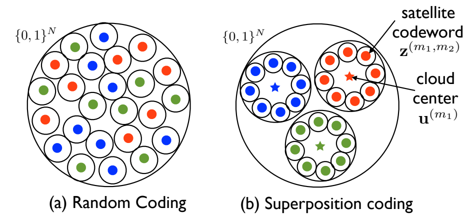

Fig. 1. The distributions of codewords (each color dot) in the output space $\{0,1\}^N$ for random block coding and for superposition coding with two levels of error protection. To better protect the color information of the codewords, which represents the MSBs of the message of the codewords, the same color codewords should be clustered together as those of superposition coding.

Fig. 1 illustrates the distribution of codewords with superposition coding as compared to that of random block coding. In the figure, the partial message M_1 (MSBs) is represented by the color of the codeword. Codewords with the same color have the same partial message M_1 (MSBs), while their M_2's (LSBs) are different. For the random block coding policy, codewords of the same color are uniformly distributed in $\{0,1\}^N$. When a noise vector corrupts the transmitted codeword beyond the correct decoding region, the decoded codeword may not have the same color as that of the transmitted codeword, since the codewords are uniformly distributed. In contrast, in the proposed superposition coding policy the

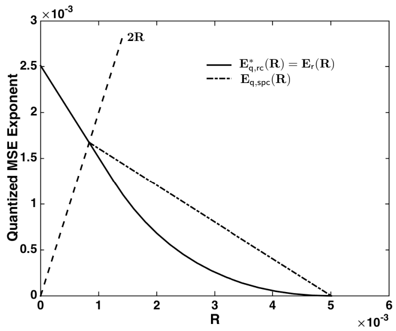

Fig. 2. A plot of $E^*_{\mathsf{q,rc}}(R) = E_{\mathsf{r}}(R)$, $E_{\mathsf{q,spc}}(R)$, and $2R$ for a BSC(ϵ) with $\epsilon = 0.45$ where $E^*_{\mathsf{q,rc}}(R)$ is the best achievable quantized-MSE exponent with random block coding and $E_{\mathsf{q,spc}}(R)$ is a lower bound on the best achievable quantized-MSE exponent $E^*_{\mathsf{q,spc}}(R)$ with superposition coding. For any $R \in (E_0(1/2,\epsilon)/3, /C)$, there is a gain in the achievable quantized-MSE exponent from superposition coding than that of random coding.

same color codewords are concentrated together. Therefore, even if the channel noise corrupts the transmitted codeword, the color information will have higher probability of being correctly decoded. As the parameter α of the superposition coding policy decreases from 1/2 to 0, the satellite codewords become more and more concentrated around the cloud centers. This allows a better error protection for M_1 (MSBs), but at the cost of worse error protection of M_2 (LSBs) since it becomes harder to distinguish between the same color codewords as they get closer to each other.

Note that the quantized MSE can be bounded above by

$$\begin{aligned}&\mathbb{E}[c_{\mathsf{q}}(X, \hat{X}_{N,\mathsf{finite}})] \leq \\ &\Pr(\hat{M}_1 \neq M_1) + \Pr(\hat{M}_2 \neq M_2 | \hat{M}_1 = M_1) e^{-2NR_1}.\end{aligned} \tag{10}$$

Thus, the quantized MSE exponent $E^*_{\mathsf{q,spc}}(R)$ in (4) of superposition coding optimized over (R_1, R_2, α) is larger than

$$\begin{aligned}&E^*_{\mathsf{q,spc}}(R) \geq \\ &\max_{(R_1,R_2,\alpha)} \min\{E^*_{\mathsf{MSBs}}(R_1,R_2,\alpha), E^*_{\mathsf{LSBs}}(R_2,\alpha) + 2R_1\}\end{aligned}$$

where $E^*_{\mathsf{MSBs}}(R_1,R_2,\alpha) = \liminf_{N\to\infty} \frac{-\ln \Pr(\hat{M}_1 \neq M_1)}{N}$ and $E^*_{\mathsf{LSBs}}(R_2,\alpha) = \liminf_{N\to\infty} \frac{-\ln \Pr(\hat{M}_2 \neq M_2 | \hat{M}_1 = M_1)}{N}$. With the proposed non-adaptive policy based on superposition coding, we prove a strictly positive gain in the quantized MSE exponent $E^*_{\mathsf{q,spc}}(R)$ compared to the best achievable quantized MSE exponent $E^*_{\mathsf{q,rc}}(R)$ in (8) of random block coding.

Theorem 1 *For a very noisy BSC(ϵ) with $\epsilon = 0.5 - \delta$ for a sufficiently small $\delta > 0$, the best achievable quantized-MSE exponent $E^*_{\mathsf{q,spc}}(R)$ of superposition coding is strictly larger than the best achievable quantized-MSE exponent $E^*_{\mathsf{q,rc}}(R)$ of random block coding for every rate $R \in (E_0(1/2,\epsilon)/3, C)$ where $E_0(1/2,\epsilon) = -\ln(1/2 + \sqrt{\epsilon(1-\epsilon)}) \approx C/6$ and $C = H_{\mathsf{B}}(1/2) - H_{\mathsf{B}}(\epsilon)$.*

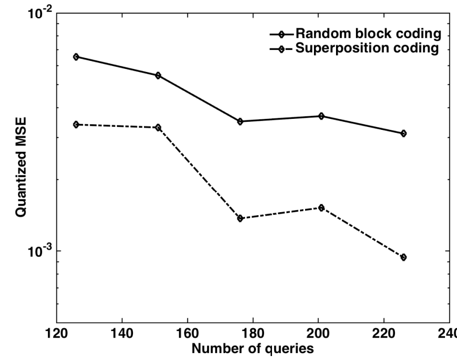

Fig. 3. Monte Carlo simulation (3000 runs) for Quantized-MSE $\mathbb{E}[c_{\mathsf{q}}(X, \hat{X}_{N,\mathsf{finite}})]$ of the querying policies based on superposition coding (dash-dot line) and of random block coding (solid line) for different number of queries, where the rates (R_1, R_2) of the partial messages (M_1, M_2) are fixed as $(R_1, R_2) = (0.5(C - R_2), 0.9C_2(\alpha))$ for capacity C of the BSC(ϵ) and for the maximum achievable rate $C_2(\alpha)$ of M_2. We set $\epsilon = 0.3$ and $\alpha = 0.1$. From the simulations, we checked that the quantized MSE from random block coding is larger on average and also its empirical distribution has heavier tail compared to that of superposition coding.

A detailed proof of this theorem can be found in the longer version of this paper [1]. In Fig 2, we provide a plot of $E_{\mathsf{q,spc}}(R)$ (dash-dot line), which is a lower bound on $E^*_{\mathsf{q,spc}}(R)$, and $E^*_{\mathsf{q,rc}}(R)$ (solid line), which is the best achievable quantized-MSE exponent using the random block coding policy. Also plotted is the line $2R$ (dashed line). We can observe the gain in the quantized-MSE exponent from superposition coding in $R \in (E_0(1/2,\epsilon)/3, C)$. Moreover, in this high rate regime, the MSE exponent in (5) is dominated by the quantized MSE exponent. Therefore, the MSE exponent of the proposed superposition coding policy is also strictly larger than that of the random block coding policy in this high rate regime.

Note that our theorem is stated for a very noisy channel (crossover probability $\epsilon \approx 1/2$). In Fig. 3, we compare Monte Carlo simulation for the quantized MSE $\mathbb{E}[c_q(X, \hat{X}_{N,\mathsf{finite}})]$ of the proposed superposition coding policy (dash-dot line) and that of the random block coding policy (sold line). This shows that our proposed policy outperforms that of random block coding even in moderate noise regimes.

5. CONCLUSION

The problem of optimal query design was considered with the goal of estimating the value of a target variable. We proposed a new non-adaptive block-querying policy based on superposition coding that can provide unequal error protection to MSBs vs. LSBs. The proposed policy achieves the same order convergence rate as the adaptive policy, with the MSE exponent larger than that of the random block coding policy.

6. REFERENCES

[1] Hye Won Chung, Brian M Sadler, Lizhong Zheng, and Alfred O Hero, "Unequal error protection querying policies for the noisy 20 questions problem," *arXiv preprint arXiv:1606.09233*, 2016.

[2] David JC MacKay, "Information-based objective functions for active data selection," *Neural computation*, vol. 4, no. 4, pp. 590–604, 1992.

[3] Burr Settles, "Active learning literature survey," *University of Wisconsin, Madison*, vol. 52, no. 55-66, pp. 11, 2010.

[4] Rui Castro and Robert Nowak, "Active learning and sampling," in *Foundations and Applications of Sensor Management*, pp. 177–200. Springer, 2008.

[5] Dennis V Lindley, "On a measure of the information provided by an experiment," *The Annals of Mathematical Statistics*, pp. 986–1005, 1956.

[6] Valerii Vadimovich Fedorov, *Theory of optimal experiments*, Elsevier, 1972.

[7] Theodoros Tsiligkaridis, Brian M Sadler, and Alfred O Hero, "Collaborative 20 questions for target localization," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2233–2252, 2014.

[8] Ehsan Variani, Kamel Lahouez, Avner Bar-Hen, and Bruno Jedynak, "Non-adaptive policies for 20 questions target localization," in *Proc. IEEE International Symposium on Information Theory Proceedings (ISIT)*. IEEE, 2015, pp. 775 – 778.

[9] Marat Valievich Burnashev and Kamil Shamilevich Zigangirov, "An interval estimation problem for controlled observations," *Problemy Peredachi Informatsii*, vol. 10, no. 3, pp. 51–61, 1974.

[10] Michael Horstein, "Sequential transmission using noiseless feedback," *IEEE Transactions on Information Theory*, vol. 9, no. 3, pp. 136–143, 1963.

[11] Thomas M Cover, "Broadcast channels," *IEEE Transactions on Information Theory*, vol. 18, no. 1, pp. 2–14, 1972.

[12] Claude Elwood Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948.