ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-6654-0540-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICASSP43922.2022.9747665

ADAPTIVE GROUP TESTING WITH MISMATCHED MODELS

Mingzhou Fan[⋆]*, Byung-Jun Yoon*[⋆†]*, Francis J. Alexander*[†]*, Edward R. Dougherty*[⋆]*, Xiaoning Qian*[⋆†]

[⋆] Department of Electrical & Computer Engineering, Texas A&M University, College Station, TX
[†] Computational Science Initiative, Brookhaven National Laboratory, Upton, NY

ABSTRACT

Accurate detection of infected individuals is one of the critical steps in stopping any pandemic. When the underlying infection rate of the disease is low, testing people in groups, instead of testing each individual in the population, can be more efficient. In this work, we consider noisy adaptive group testing design with specific test sensitivity and specificity that select the optimal group given previous test results based on preselected utility function. As in prior studies on group testing, we model this problem as a sequential Bayesian Optimal Experimental Design (BOED) to adaptively design the groups for each test. We analyze the required number of group tests when using the updated posterior on the infection status and the corresponding Mutual Information (MI) as our utility function for selecting new groups. More importantly, we study how the potential bias on the ground-truth noise of group tests may affect the group testing sample complexity.

Index Terms— Group testing, Bayesian optimal experimental design (BOED), mismatched models,

1. INTRODUCTION

Originally proposed for blood testing in the World War II [1], group testing has been a powerful tool for detecting infected individuals in a large population, for example by polymerase chain reaction (PCR) tests for COVID-19 [2]. By mixing up the test samples (e.g. saliva or blood) of individuals in a group, the tester can determine whether there exists any infected individual in the group if no mistakes are made.

Current studies on group testing can be roughly divided into two categories – *non-adaptive* [3–7] and *adaptive* methods [3, 8–14] – based on whether the group to test is decided before the tests or adaptively given test results during the whole sequential procedure.

Non-adaptive group testing is solved in a two-stage fashion: designing testing groups and recovering the infection status based on the testing results. The very first paper about group testing [1] is in a non-adaptive way and derived the optimal group selection in a noiseless setup. A review of non-adaptive group testing and its applications can be found in [5].

Adaptive group testing updates the Bayesian models of infection status based on the previous testing results and designs a new group to test at each iteration. Works on adaptive group testing have examined its empirical performance in different set-ups. Model inference by lattice-based classification models [15] or sum-observation [3] has been explored. In [10], Loopy Belief Propagation (LBP) [16] and other approximation strategies [11] were adopted for scalable inference. All the aforementioned works design group tests based on the entropy-based utility function. Recently, other utility functions, including mutual information (MI) and expected area under the receiver operating characteristic curve (AUC), have been explored with a sequential Monte-Carlo (SMC) method in [9]. To the best of our knowledge, theoretical analysis for noisy adaptive group testing is limited. In [11], the sample complexity when using the entropy utility function was derived. However, computational of the sample complexity requires the ground-truth probability, which is typically unknown in practice.

In this paper, we consider adaptive group testing in the presence of uncertainty. We follow the model formulation in [9] to set a Bernoulli prior for the infection status. Group testing design is based on a mutual information utility function of the current posterior, iteratively updated given previous results. Based on a stopping criterion of conditional entropy, we derive a lower bound T_E of the required number of group tests. More importantly, we further analyze the sample complexity when we have possibly mismatched testing models. We prove that when the model parameters are mismatched with the ground truth, the lower bound increases as expected, $T'_E = (1+\alpha)T_E$ with a constant $\alpha > 0$, due to the optimal group selection based on the biased utility function. Such a theoretical analysis has not been discussed in the existing literature. We further confirm our analyses by simulation results.

2. GROUP TESTING

We note that capital letters denote random variables and vectors are with the bold font in this paper. The detailed deriva-

The work is supported in part by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Mathematical Multifaceted Integrated Capability Centers program under Award Q1 DE-SC0019303, and the National Science Foundation (NSF) Awards 1553281, 1812641, and 2119103.

tions can be found in [17].

Given a population of n individuals whose infection statuses are unknown and modeled by Bernoulli random variables X_i's, $X_i = 1$ if the i-th individual is infected, and otherwise $X_i = 0$. Denote the random vector representing our understanding on the population infection state by $\mathbf{X} = (X_1, X_2, \ldots, X_n)$, where $X_i \sim \text{Bern}(q_i)$ with the infection probability q_i for the ith individual. We are interested in designing group tests adaptively to discover the unknown infection state: $\mathbf{x_{true}} \in \{0,1\}^n$. We assume that the Bernoulli random variables X_i's are independent, i.e. for any $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \{0,1\}^n$, we denote $\pi_0(\mathbf{x}) = P(\mathbf{X} = \mathbf{x}) = \prod_i q_i^{x_i}(1-q_i)^{1-x_i}$. Here, $P(X_i = 1) = q_i$ can be identical or vary across individuals. For the j-th group test result $Y_j \in \{0,1\}$, $Y_j = 1$ if the j-th sample is tested positive, indicating that the combined sample contains the sample(s) from infected individual(s); and otherwise all the individuals in the jth group test are not infected. Group testing design is to choose a subset of individuals from the population as a group, denoted as a vector $\mathbf{g}_j = (g_{j,1}, g_{j,2}, \ldots, g_{j,n}) \in \{0,1\}^n$, and test the mixed samples from the individuals with the corresponding $g_{j,i} = 1$.

Existing testing assays have limitations and it is possible to have testing errors. As in [9], we assume that the group testing has the following sensitivity ($s_{\mathbf{g}_j}$) and specificity ($\sigma_{\mathbf{g}_j}$):

$$P(Y_j = 1|[\mathbf{g}_j, \mathbf{x}] = 1) = s_{\mathbf{g}_j},\ P(Y_j = 0|[\mathbf{g}_j, \mathbf{x}] = 0) = \sigma_{\mathbf{g}_j}, \tag{1}$$

where $[\mathbf{g}, \mathbf{x}] = \min(1, \mathbf{g}^T\mathbf{x}) \in \{0,1\}$. In what follows, we refer to $s_{\mathbf{g}_j}$ and $\sigma_{\mathbf{g}_j}$ as model parameters for adaptive group testing design. Assume that we have designed a batch of testing groups $\mathcal{G}_t = \{\mathbf{g}_{t_m+1}, \ldots, \mathbf{g}_{t_m+m}\}$ at stage t, where $t_m = (t-1)m$ and m is the batch size. Given their corresponding test results $\mathbf{Y}_t = \{Y_{t_m+1}, \ldots, Y_{t_m+m}\}$, we can compute:

$$Pr(\mathbf{Y}_t = \mathbf{y}_t|\mathbf{X} = \mathbf{x}) = \prod_{j=1}^{m} (Q^{(0)}_{t_m+j})^{(1-y_t^{(j)})} (Q^{(1)}_{t_m+j})^{y_t^{(j)}} \tag{2}$$

based on (1), where $Q^{(0)}_{t_m+j} = \sigma_{\mathbf{g}_{t_m+j}} - \rho_{\mathbf{g}_{t_m+j}}[\mathbf{g}_{t_m+j}, \mathbf{x}]$, $Q^{(1)}_{t_m+j} = 1 - \sigma_{\mathbf{g}_{t_m+j}} + \rho_{\mathbf{g}_{t_m+j}}[\mathbf{g}_{t_m+j}, \mathbf{x}]$. Here, $\rho_{\mathbf{g}_{t_m+j}} = s_{\mathbf{g}_{t_m+j}} + \sigma_{\mathbf{g}_{t_m+j}} - 1$. We can further infer the posterior of the population infection status by Bayes's rule: $\pi_t(\mathbf{x}) \propto \pi_0(\mathbf{x}) \prod_{k=1}^{t} Pr(\mathbf{Y}_k = \mathbf{y}_k|\mathbf{X} = \mathbf{x})$. For simplicity, from now on, we write the posterior of an event E given the previous t test results by: $P_t(E) = P(E|\mathbf{Y}_1 = \mathbf{y}_1, \mathbf{Y}_2 = \mathbf{y}_2, \ldots, \mathbf{Y}_t = \mathbf{y}_t)$.

3. ADAPTIVE GROUP TESTING

For adaptive group testing, we design a utility function $U_t(\mathcal{G}_t) = U(\mathcal{G}_t, \pi_t)$ to guide the iterative selection of a batch of groups to update the posterior π_t of infection status. More specifically, the task at each stage is to find a batch of groups $\mathcal{G}_t^*$ such that $\mathcal{G}_t^* \in \arg\max_{\mathcal{G}_t} U_t(\mathcal{G}_t)$.

3.1. Mutual Information based Utility Function

We adopt a mutual information utility function:

$$U_{MI}(\mathcal{G}_t, \pi_t) = I(\mathbf{X}; \mathbf{Y}_k|\mathbf{Y}_1, \ldots, \mathbf{Y}_{k-1}) \triangleq I_{P_{k-1}}(\mathbf{X}; \mathbf{Y}_k). \tag{3}$$

Denote $h(p) = -p\log_2 p - (1-p)\log_2(1-p)$ as the binary entropy and $H(\mathbf{X}|Y_1, Y_2, \ldots, Y_t)$ as $H_{P_t}(\mathbf{X})$. We have:

$$\begin{aligned} I_{P_{t-1}}(\mathbf{X}; \mathbf{Y}_t) &= H_{P_{t-1}}(\mathbf{Y}_t) \\ &\quad - \sum_{j=1}^{k} [h_{\sigma_{\mathbf{g}_{t_m+j}}} + \gamma_{\mathbf{g}_{t_m+j}} f_{\pi_t}(\mathbf{g}_{t_m+j})], \end{aligned} \tag{4}$$

where $h_\phi = h(\phi)$, $\gamma_{\mathbf{g}} = h_{s_{\mathbf{g}}} - h_{\sigma_{\mathbf{g}}}$, and

$$f_{\pi_t}(\mathbf{g}) = \sum_{\mathbf{x}} \pi_t(\mathbf{x})[\mathbf{g}, \mathbf{x}] = \sum_{\mathbf{x}:[\mathbf{g},\mathbf{x}]=1} \pi_t(\mathbf{x}), \tag{5}$$

representing the probability of having infected patient(s) in the chosen group.

Here, we consider the simplified set-ups with all the sensitivity and specificity being constant with respect to group selection, i.e. $\sigma_{\mathbf{g}_{t_m+j}} = \sigma$, and $s_{\mathbf{g}_{t_m+j}} = s$. The Mutual Information (4) can be written as $I_{P_{t-1}}(\mathbf{X}; \mathbf{Y}_t) = H_{P_{t-1}}(\mathbf{Y}_t) - \sum_{j=1}^{k} [h_\sigma + \gamma f_{\pi_t}(\mathbf{g}_{t_m+j})]$. Further assume that we test one group at each stage. We have:

$$I_{P_{t-1}}(\mathbf{X}; Y_t) = h(\rho f_{\pi_t}(\mathbf{g}_t) + 1 - \sigma) - h_\sigma - \gamma f_{\pi_t}(\mathbf{g}_t), \tag{6}$$

where $\rho f_{\pi_t}(\mathbf{g}) + 1 - \sigma = P_t(Y_t = 1)$. Note that we have replaced $\mathbf{Y}_t$ by Y_t and $\mathcal{G}_t$ by $\mathbf{g}_t$. Write $J(x) = h(\rho x + 1 - \sigma) - h_\sigma - \gamma x$, which is a concave function of x. Note that (6) can be written as $I_{P_{t-1}}(\mathbf{X}; Y_t) = J(f_{\pi_t}(\mathbf{g}_t))$, which is concave so it would be maximized when its derivative at $f_{\pi_t}(\mathbf{g}_t)$ is zero, leading to the closed-form optimal point of (6): $f^* = \frac{\sigma}{\rho} - \frac{\exp\frac{\gamma}{\rho}}{\rho(\exp\frac{\gamma}{\rho}+1)}$.

Note that $J(x)$ is fixed when the group testing sensitivity and specificity parameters are given. It is concave and we can find $f_{\pi_t}(\mathbf{g}_t)$ that optimizes $J(x)$. The design problem becomes $\mathbf{g}_t^* \in \arg\max_{\mathbf{g}_t} U_t(\mathbf{g}_t)$, and can be informally viewed as finding $\mathbf{g}_t^*$ to make $f_{\pi_t}(\mathbf{g}_t^*)$ as close as possible to f^*.

3.2. Stopping Criteria

In previous studies, either the budget [9] or the maximum probability of infection status, i.e. $\max_{\mathbf{x}} P_{t-1}(\mathbf{x})$ [10, 11], have been used as stopping criteria. The former cannot help us analyze the asymptotic performance, while the latter can be difficult to analyze when using estimation methods like LBP. We use Conditional Entropy (CE) as the stopping criterion:

$$H_{P_t}(\mathbf{X}) \leq \delta H(\mathbf{X}), \tag{7}$$

where $0 \leq \delta \leq 1$ and $H(\mathbf{X})$ is the entropy of the prior and is fixed once the prior is given. With the mutual information definition, we have $H_{P_t}(\mathbf{X}) = H(\mathbf{X}) - \sum_{k=1}^{t} I_{P_{k-1}}(\mathbf{X}; Y_k)$. If we only search for one group to query at each stage,

$$H_{P_t}(\mathbf{X}) = H(\mathbf{X}) - \sum_{k=1}^{t} J(f_{\pi_k}(\mathbf{g}_k)). \tag{8}$$

The stopping criterion becomes

$$\sum_{k=1}^{t} J(f_{\pi_k}(\mathbf{g}_k)) \geq (1-\delta)H(\mathbf{X}). \tag{9}$$

With this, we are able to interpret the stopping criterion from an information theoretic perspective. Furthermore, it is straightforward to compute once each $f_{\pi_t}(\mathbf{g}_t)$ is given by adaptive group selection.

3.3. Sample Complexity

It worth noting that, by the nature of deriving $f(\mathbf{g})$ in (5), we are not guaranteed to reach f^* every adaptive group testing iteration. In other words, there are always gaps between the achieved $f(\mathbf{g})$ by the designed group test and the optimal f^*. Therefore we cannot treat the information gain at each iteration as a constant. Besides, it can be difficult to analyze how close $f_{\pi_t}(\mathbf{g})$ can approach f^* as π_t adapts over the iterations.

Here, we assume that $f_{\pi_t}(\mathbf{g}_t^*) \in \arg\max_{f_{\pi_t}(\mathbf{g}_t)} J(f_{\pi_t}(\mathbf{g}_t))$ can be approximately modeled by Gaussian random variables: $F_t \sim N(f^*, \nu_F^2)$. Thus, we can transform the stopping criterion (9) into $\sum_{k=1}^{t} J(F_t) \geq (1-\delta)H(\mathbf{X})$.

Note that $h(x) \geq -4(x-0.5)^2 + 1$. Therefore, $J(x) \geq J^{(4)}(x)$, where $J^{(A)}(x) = -A\rho^2 x^2 - [2A(0.5-\sigma)\rho - \gamma]x - A(0.5-\sigma)^2 + 1 - h_\sigma$. We can derive the following theorem.

Theorem 1. *If $T \geq T_E^{(A)}$, we have*

$$\begin{aligned} &Pr(\sum_{k=1}^{T} J^{(A)}(F) \geq (1-\delta)H(\boldsymbol{X})) \\ &\quad \geq \frac{[E_F^{(A)}T - (1-\delta)H(\boldsymbol{X})]^2}{V_F^{(A)}T + [E_F^{(A)}T - (1-\delta)H(\boldsymbol{X})]^2}, \end{aligned} \tag{10}$$

where $T_E^{(A)} = \frac{(1-\delta)H(\boldsymbol{X})}{E_F^{(A)}}$, $E_F^{(A)} = -A\rho^2\nu_F^2 - A(0.5-\sigma)^2 + 1 - h_\sigma - A\rho^2(f^*)^2 - B_A f^*$, $V_F^{(A)} = 2A^2\rho^4\nu_F^4 + (B_A + 2A\rho^2 f^*)^2\nu_F^2$, and $B_A = 2A(0.5-\sigma)\rho - \gamma$.

We now give the condition for $H_{P_T}(\mathbf{X}) \leq \delta H(\mathbf{X})$ in the following proposition.

Proposition 1. *If $T \geq T_E^{(4)}$, we have*

$$\begin{aligned} &Pr(H_{P_T}(\boldsymbol{X}) \leq \delta H(\boldsymbol{X})) \\ &\quad \geq 1 - \frac{V_F^{(4)}T}{V_F^{(4)}T + [(T - T_E^{(4)})E_F^{(4)}]^2}. \end{aligned} \tag{11}$$

Based on this theorem, we have shown that the probability of meeting the stopping criterion is in the rate of $1 - o(T^{-1})$ when $T \geq T_E^{(4)}$. $T_E^{(4)}$ takes the form $\frac{(1-\delta)H(\mathbf{X})}{E_F^{(A)}}$, which is related to the population size n and the infection rate q_i. Given an i.i.d. Bernoulli prior, $T_E^{(4)}$ becomes $n\frac{(1-\delta)h(q)}{E_F^{(A)}}$, which is proportional to the number of patients. In general the variance $\nu \ll 1$ is small, which leads to $(E_F^{(4)})^2 \gg V_F^{(4)}$ so that $Pr(H_{P_T}(\mathbf{X}) \leq \delta H(\mathbf{X}))$ can be close to 1 very quickly as soon as $T \geq T_E^{(4)}$.

4. MISMATCHED MODEL

Now suppose there is a mismatch between the assumed group test model parameters and the true parameters. Specifically, we consider the impact of using incorrect sensitivity s' and specificity σ' for group testing, instead of the true sensitivity s and the true specificity σ, which are unknown in practice. In each iteration, we would optimize the 'mismatched' utility function: $I'_{P'_{t-1}}(\mathbf{X}; \mathbf{Y}_t) = H_{P'_{t-1}}(\mathbf{Y}_t) - \sum_{j=1}^{k}[h'_{\sigma_{\mathbf{g}_{t_m+j}}} + \gamma'_{\mathbf{g}_{t_m+j}} f_{\pi'_t}(\mathbf{g}_{t_m+j})]$ and select the group such that $\mathcal{G}' \in \arg\max_{\mathcal{G}} I'_{P'_{t-1}}(\mathbf{X}; \mathbf{Y}_t)$, where π'_t is the mismatched posterior updated with the mismatched parameters. With the same setup in Section 3, the selection at each iteration is $\mathbf{g}' \in \arg\max_{\mathbf{g}} J'(f_{\pi'_t}(\mathbf{g}))$.

The mismatched selection target of $f_{\pi'_t}(\mathbf{g})$ would be $f' = \frac{\sigma'}{\rho'} - \frac{\exp\frac{\gamma'}{\rho'}}{\rho'(\exp\frac{\gamma'}{\rho'}+1)}$. The actual information gain, however, should be calculated with the true parameters, $I_{P_{t-1}}(\mathbf{X}; \mathbf{Y}_t) = J(f_{\pi_t}(\mathbf{g}'_t))$. Notice that here the true posterior π_t needs to be updated with true parameters and have the 'true' understanding on the infection status.

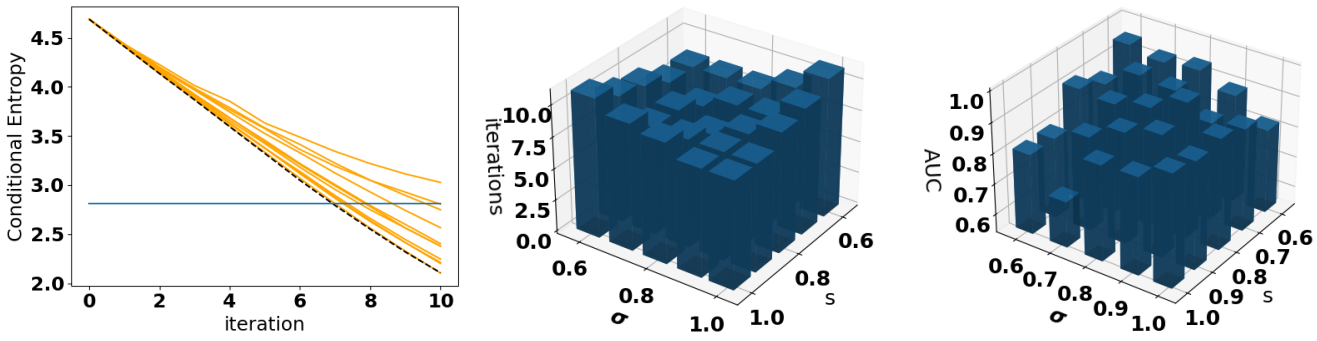
Similar to the derivation in the previous section, we consider $f'_{\pi_t}(\mathbf{g}_t) = F'_t \sim N(f', \nu_{F'}^2)$. When there is model mismatch, the variance would be much larger. Similar to Proposition 1, we have:

Proposition 2. *If $T \geq (1+\alpha^{(4)})T_E^{(4)}$, we have*

$$\begin{aligned} &Pr(H_{P_T}(\boldsymbol{X}) \leq \delta H(\boldsymbol{X})) \\ &\quad \geq 1 - \frac{TV_{F'}^{(4)}}{TV_{F'}^{(4)} + [(T - (1+\alpha^{(4)})T_E^{(4)})E_{F'}^{(4)}]^2} \end{aligned} \tag{12}$$

for mismatched models, where
$\alpha^{(A)} = \frac{A\rho^2(f')^2 + B_A f' - A\rho^2(f^*)^2 - B_A f^*}{E_{F'}^{(A)}}$.

Here $\alpha^{(4)}$ represents the change of sample complexity because of model bias on f' to f^*. We want to point out that $(E_{F'}^{(4)})^2 \gg \nu_{F'} \gg \nu_F$ holds, so the probability can still be close to 1 once $T \geq (1+\alpha^{(4)})T_E^{(4)}$. The main influence of biased model is the difference of f' and f^*. Also $\alpha^{(4)} = 0$ if $f' = f^*$, so that we can observe that the performance is similar when $f' = f^*$ in the experiments.



(a) Average CE (b) Iterations needed ($\delta = 0.6$) (c) Average AUC after 8 tests

Fig. 1: Simulation Results. 1a: The average CE over test iterations for each setup, the matched model is illustrated as black dashed line. 1b: Required iterations when setting $\delta = 0.6$. 1c: Average AUC after 8 active group testing iterations.

5. EXPERIMENTAL RESULTS

We perform simulations to confirm the derived bounds of the required number of group testing iterations in this section.

5.1. Experimental Settings

We investigate how the group testing performance changes with different model parameter settings. We have simulated results for 24 combinations of mismatched parameters (biased) $\sigma', s' \in \{0.6, 0.7, 0.8, 0.9, 0.99\}$ together with the ground-truth group testing parameters (unbiased), $\sigma' = \sigma = 0.8$ and $s' = s = 0.8$. To allow the exhaustive search to achieve the best achieved group design, we have simulated 1,000 runs with a population of ten individuals, among whom one individual is randomly selected to be infected. The prior π_0 is set as the independent Bernoulli for each individual and the probability of each individual being infected is 0.1, $X_i \sim \text{Bern}(0.1)$. We perform adaptive group testing as described in previous sections for each simulation run and take the average conditional entropy and Area Under the receiver operating characteristic Curve (AUC) [18] over 1,000 runs for each iteration for performance evaluation.

5.2. Conditional Entropy

In this set of experiments, the ground-truth entropy for performed simulations is $H(\mathbf{X}) = -nh(p) = 10h(0.1)$. We have plotted the average Conditional Entropy defined in (8) over group testing iterations in Figure 1a. The dashed curve is the performance based on the ground-truth model parameters, which outperforms the ones based on the utility function with mismatched models as one would expect.

We then show the average number of the required group tests in Figures 1b when $\delta = 0.6$, respectively. The horizontal line in Figure 1a shows the value of $\delta H(\mathbf{X})$ when $\delta = 0.6$. We can see that when $f' = f^*$, we have similar required test numbers. More importantly, with mismatched models, more group testing iterations are required to achieve the desired conditional entropy level.

5.3. AUC

To further confirm the infection detection performance, we compute the AUC based on the marginal likelihood as the criterion to evaluate the performance of our updated posterior given corresponding group test results across adaptive group testing iterations. The marginal likelihood for each of the individuals, indexed by i, can be computed as $m_t(i) = P_{t-1}(X_i = 1) = \sum_{[\mathbf{g}_i, \mathbf{x}]=1} \pi_t(\mathbf{x})$, where X_i is the infection status of the i-th individual, $\mathbf{g}_i$ is a group that only contains i-th individual, i.e. one-hot coding of the i-th individual. With that, the AUC of the infection marginal likelihood m_t can be written as $\mathbf{AUC}(m_t) = \frac{\sum_{i^+ \in POS} \sum_{i^- \in NEG} \mathbb{I}(m_t(i^+) > m_t(i^-))}{|POS||NEG|}$, where $\mathbb{I}$ is the indicator function: $\mathbb{I}(E) = 1$ if E is true.

As illustrated in Figure 1c, although not directly optimized with respect to AUC, the AUC values for each setup can be improved, reflected by the change of conditional entropy, during the adaptive group testing iterations. It does fluctuate and may not have a strictly monotonic relationship with the accuracy of infection detection, probably due to the inherent gap between the optimal f^* and the best achievable result.

6. CONCLUSIONS

In this paper, we proved that the probability of meeting the stopping criterion based on conditional entropy is in the rate of $1 - o(T^{-1})$. More importantly, we have shown that a mismatch in the group testing model would lead to a multiplicative constant $1 + \alpha$ ($\alpha > 0$), determined by the difference between f^* and f', to the required number of group tests. Our simulation study shows that the adaptive group testing can be efficient in infection detection based on the mutual information utility. Adaptive design with the correct group testing model outperforms the ones with mismatched models. The performance evaluation by AUC has shown to be related to the conditional entropy though not strictly monotonic.

References

[1] Robert Dorfman, "The detection of defective members of large populations," *The Annals of Mathematical Statistics*, vol. 14, no. 4, pp. 436 – 440, 1943.

[2] Christian Gollier and Olivier Gossner, "Group testing against covid-19," EconPol Policy Brief 24, Munich, 2020.

[3] Weidong Han, Purnima Rajan, Peter I. Frazier, and Bruno M. Jedynak, "Bayesian group testing under sum observations: A parallelizable two-approximation for entropy loss," *IEEE Transactions on Information Theory*, vol. 63, no. 2, pp. 915–933, 2017.

[4] Aiyi Liu, Chunling Liu, Zhiwei Zhang, and Paul S. Albert, "Optimality of group testing in the presence of misclassification," *Biometrika*, vol. 99, no. 1, pp. 245–251, Dec. 2011.

[5] Ding-Zhu Du and Frank K. Hwang, *Combinatorial Group Testing and Its Applications*, WORLD SCIENTIFIC, Nov. 1993.

[6] George K. Atia and Venkatesh Saligrama, "Boolean compressed sensing and noisy group testing," *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1880–1901, 2012.

[7] Milton Sobel and R. M. Elashoff, "Group testing with a new goal, estimation," *Biometrika*, vol. 62, no. 1, pp. 181–193, 04 1975.

[8] Leonardo Baldassini, Oliver Johnson, and Matthew Aldridge, "The capacity of adaptive group testing," in *2013 IEEE International Symposium on Information Theory*, 2013, pp. 2676–2680.

[9] Marco Cuturi, Olivier Teboul, Quentin Berthet, Arnaud Doucet, and Jean-Philippe Vert, "Noisy adaptive group testing using Bayesian sequential experimental design," 2020.

[10] Ayaka Sakata, "Active pooling design in group testing based on Bayesian posterior prediction," *Phys. Rev. E*, vol. 103, pp. 022110, Feb 2021.

[11] Yechao Bai, Qingsi Wang, Chun Lo, Mingyan Liu, Jerome P. Lynch, and Xinggan Zhang, "Adaptive Bayesian group testing: Algorithms and performance," *Signal Processing*, vol. 156, pp. 191–207, 2019.

[12] Christopher S. McMahan, Joshua M. Tebbs, Timothy E. Hanson, and Christopher R. Bilder, "Bayesian regression for group testing data," *Biometrics*, vol. 73, no. 4, pp. 1443–1452, Apr. 2017.

[13] Jacqueline M. Hughes-Oliver and William H. Swallow, "A two-stage adaptive group-testing procedure for estimating small proportions," *Journal of the American Statistical Association*, vol. 89, no. 427, pp. 982–993, Sept. 1994.

[14] Jonathan Scarlett, "Noisy adaptive group testing: Bounds and algorithms," *IEEE Transactions on Information Theory*, vol. 65, no. 6, pp. 3646–3661, 2019.

[15] Curtis Tatsuoka and Weicong Chen, "Bayesian group testing with dilution effects," *medRxiv*, 2021.

[16] Kevin P Murphy, Yair Weiss, and Michael I Jordan, "Loopy belief propagation for approximate inference: An empirical study," in *PROCEEDINGS OF UNCERTAINTY IN AI,*, 1999, pp. 467–475.

[17] Mingzhou Fan, Byung-Jun Yoon, Francis J. Alexander, Edward R. Dougherty, and Xiaoning Qian, "Adaptive group testing with mismatched models," `https://arxiv.org/abs/2110.02265`, 2021.

[18] Andrew P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern Recognition*, vol. 30, no. 7, pp. 1145–1159, July 1997.