<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper published in *IEEE Transactions on Information Theory*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Vu, M T., Oechtering, T J., Skoglund, M. (2021)
Hypothesis Testing and Identification Systems
IEEE Transactions on Information Theory, 67(6): 3765-3780
<https://doi.org/10.1109/TIT.2021.3076497>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-297283>

Hypothesis Testing and Identification Systems

Minh Thanh Vu, Tobias J. Oechtering *Senior Member, IEEE*, Mikael Skoglund *Fellow, IEEE*

Abstract—We study hypothesis testing problems with fixed compression mappings and with user-dependent compression mappings to decide whether or not an observation sequence is related to one of the users in a database, which contains compressed versions of previously enrolled users' data. We first provide the optimal characterization of the exponent of the probability of the second type of error for the fixed compression mappings scenario when the number of users in the database grows exponentially. We then establish operational equivalence relations between the Wyner-Ahlsvede-Körner network, the single-user hypothesis testing problem, the multi-user hypothesis testing problem with user-dependent compression mappings and the identification systems both with user-dependent compression mappings. Finally they also show how an identification scheme can be turned into a multi-user hypothesis testing scheme with an explicit transfer of rate and error probability conditions and vice versa.

Index Terms—Mixture distribution, identification systems, information-spectrum method, strong converse, excess relative information, soft-covering, constructive transformation.

I. INTRODUCTION

Membership testing has not been actively considered in existing works on identification systems. It is often assumed that the observation sequence is related to the data inside the system. In this work we put our attention to this important problem. Assume a database that stores *compressed* versions of data sequences of M users $(x^n(m))_{m=1}^M$. An observation sequence y^n is provided to a processing center which has access to these compressed data sequences. The processing center performs a screening step and returns *Yes/No* when y^n is *related to one of the users / independent of all users* in the system. We call the first case hypothesis H_0 and the second case hypothesis H_1 . The case that $M = 1$ was studied in [1] and referred in this work as single-user testing against independence to differentiate it from our multi-user setting. We are interested in the error exponent of the second type of error in Stein's regime in our work.

Our problem is influenced by the identification system, cf. Fig. 1, investigated in [2], where the task is to identify the correct user based on the observation y^n and the information inside the database $(x^n(i))_{i=1}^M$. Therein, y^n is an output of an observation channel $P_{Y^n|X^n}$ where the input sequence is

Fig. 1: A simplified model for the identification system. In the *enrollment* phase, observations $(x^n(i))_{i=1}^M$ are mapped into stored indices by a mapping ϕ_n^{id} . In the *identification* phase, one user w is selected uniformly at random. The random selection is described by a random variable W . An observation sequence y^n , which is the output of the observation channel $P_{Y^n|X^n}$ with the input $x^n(w)$, is provided to the processing center. The processing center tries to recognize the user by producing an estimate index $\hat{w}$ based on y^n and the database.

selected uniformly at random from the database $(x^n(i))_{i=1}^M$. The compression of users' data is considered in [3], [4]. In contrast, we are not interested in identifying the true user even if the sequence y^n is indeed related to one user in the system. Instead, we only want to know if the sequence is related to the database.

We first study the hypothesis testing problem when only one compression mapping is used to enroll all users' data. Next, we generalize the setting by allowing that each user can have a different compression mapping. This generalization may arise when separate or distributed databases which use different compression techniques are merged together for the identification purpose. We also examine the related identification problem where each user can have their own compression mapping. To establish various optimal characterizations for these problems we study the code transformation between the Wyner-Ahlsvede-Körner network, the single-user hypothesis testing problem, the multi-user hypothesis testing problem with user-dependent compression mappings and the identification systems with user-dependent compression mappings.

We summarize our main results and contributions in the following.

- We characterize the error exponent of the hypothesis testing problem with a single compression mapping in weak converse sense. Our result complements the one given in [3] by showing the fundamental trade-off between the number of users in the database and the ability to reject a stranger. In the weak converse proof we use a soft-covering lemma to transform our multi-user setting to a

The work was partially supported by the Swedish Research Council (VR) under Grant 2016-03853 and Strategic Research Agenda Program, Information and Communication Technology – The Next Generation (SRA ICT – TNG) funded by the Swedish Government.

Minh Thanh Vu was with the Division of Information Science and Engineering, KTH-Royal Institute of Technology, Stockholm, Sweden.

Tobias J. Oechtering and Mikael Skoglund are with the Division of Information Science and Engineering, KTH-Royal Institute of Technology, Stockholm, Sweden.

This paper was presented in parts at ITW 2018 and at ISIT 2019.

single-user setting. To the best of our knowledge, this step is the first instance of novel use of a soft-covering lemma in a converse proof, which is otherwise mainly used in achievability proofs.

- We establish an operational equivalence between the single-user testing problem with the one of the Wyner-Ahlsvede-Körner (WAK) network [5], [6] via a code transformation argument in a general setting. The explicit code transformation arguments apply to any code rate. In particular, from this relation we establish the strong converse and exponentially strong converse equivalence between the single-user testing problem and the WAK network.
- Using another code transformation argument and the derived soft-covering lemma, we establish an operational equivalence between the single-user testing problem with our multi-user testing problem with user-dependent compression mappings in a general setting. As a consequence, under a uniformity condition we show the relations between the ϵ -achievable error exponents of these two problem which further implies the strong converse for the multi-user hypothesis testing with user-dependent compression mappings for the iid discrete case with uncompressed y^n .
- We also establish the exponentially strong converse equivalence between the multi-user hypothesis testing and the single-user hypothesis testing problem. As a consequence of the two established operational equivalence relations, we show the behavior of the error probabilities in the multi-user testing problem when the database contains more users than the identification capacity limit in the iid setting with uncompressed y^n .
- Finally, we establish an operational equivalence between the single-user testing problem and the identification system with user-dependent compression mappings in a general setting. Similarly, under a uniformity condition, we establish as consequences the relations between the maximum (second-order) ϵ -achievable error exponent in the single-user testing problem and the maximum (second-order) ϵ -achievable identification rate in the identification problem. These results further imply the strong converse and exponentially converse for the identification problem with user-dependent compression mappings in the iid discrete setting with uncompressed y^n .

Combining the two latter transformations it can be seen that an identification scheme can be explicitly turned into a multi-user hypothesis testing scheme and vice versa. In particular, when a fixed compression mapping is used in the identification problem, the same compressed database can be reused in the membership testing problem. This property is highly desirable in practice since one can add a new feature, herein rejecting a stranger, on top of an old feature (recognizing an existing user) without re-designing the whole system from the scratch.

Other hypothesis testing problems related to the identification problem include [7]–[9]. In [7] the hypothesis H_1 was tested against M other hypotheses in the binary setting where the focus was to minimize the overall identification error under

a specific decision rule. It was shown that when the rate of M is below a value, which is less than or equal to the uncompressed binary identification capacity, then the overall error probability goes to zero. In [8] the author considered the M -ary hypothesis testing problem with fixed M and studied the large deviation regime. In [9] the decision rule was based on a decoding metric using the hashed data and observation sequences at different lengths. The exponents of the probability of miss and the expected number of incorrect items on the list were provided for a fixed hashed function. Error exponent aspects of the probability of estimating the correct user in the identification systems have been studied in [10]–[12]. Recent developments on distributed hypothesis testing with privacy constraints and relay networks can be found in [13]–[17].

Multi-user hypothesis testing was also studied in [1]. In Remark 5 we compare their study with ours to highlight the differences. Additionally, operational equivalences between source coding problems and the hypothesis testing problems have appeared in the literature, cf. the works [18], [19]. The similarities and differences of these works with ours are stated in Remark 7.

The paper is organized as follows. We present the complete achievable error exponent in the weak converse sense of our multi-user hypothesis testing when the same compression mapping is used for all users in Section II. We also remark that the arguments used in the weak converse proof can be applied for other settings. In Section III we present the code transformation arguments between the single-user testing, the WAK network, the multi-user hypothesis testing and the identification systems as well as their consequences.

II. PRELIMINARIES

We begin with some notational conventions. Random variables, their realizations are denoted by uppercase, lowercase letters, respectively. Sets are denoted by calligraphic letters. The complement of a set $\mathcal{A}$ is denoted by $\mathcal{A}^c$. $\log$ is taken to the natural base. For a mapping ϕ_n , $|\phi_n^t|$ denotes the cardinality of its image. For a measure μ , $\mu^{\otimes n}$ denotes its n -fold product extension. Additionally, we use the $(\bar{\cdot})$ notation, e.g. $\bar{\alpha}_n$, $\bar{A}_n$, to emphasize that the single-user scenario is considered. The information density between two random variables Z and V that are jointly distributed according to P_{ZV} is defined when $P_{ZV} \ll P_Z \times P_V$, where $\ll$ denotes the absolute continuity relation, as

$$\iota_{ZV}(z; v) = \log \frac{dP_{ZV}}{d(P_Z \times P_V)}(z, v).$$

When P_{ZV} is clear from the context we omit the subscript.

We assume that the database consists of M users with $\lim_{n \rightarrow \infty} \frac{1}{n} \log M = R$. i.e., the number of users grows with the block length n at rate R . We also denote this relation in the sequel by $M \doteq e^{nR}$. In this section we mainly consider the discrete scenario where alphabets $\mathcal{X}$ and $\mathcal{Y}$ are finite. For each i the corresponding data sequence $x^n(i)$ is generated iid

Fig. 2: Screening to find out whether the observation is related to one user inside the system.

from the distribution P_X . Under H_0 the joint distribution of the sequence y^n and sequences $(x^n(m))_{m=1}^M$ is given by

$$P_{H_0} = \sum_{i=1}^M \frac{1}{M} P_{Y^n X^n(i)} \times \prod_{k=1, k \neq i}^M P_{X^n(k)}, \quad (1)$$

i.e., the sequence y^n is related to one randomly chosen user in the system, where $P_{Y^n X^n(i)} = P_{Y^n}^{\otimes n}$. The joint distribution under H_1 is given by

$$P_{H_1} = P_{Y^n} \times \prod_{i=1}^M P_{X^n(i)}, \quad (2)$$

where herein $P_{Y^n} = P_Y^{\otimes n}$, i.e., the sequence y^n is not related to the information in the database. Note that under both hypotheses the users' data sequences are mutually independent. Additionally, the processes governed by P_{H_0} and P_{H_1} are *non-stationary* since $M \doteq e^{nR}$ grows exponentially with n ¹. We can also view P_{H_0} as the result of mixing M general random processes uniformly where the distributions at instance n are given by $P_{Y^n X^n(i)} \times \prod_{k=1, k \neq i}^M P_{X^n(k)}$, $i \in [1 : M]$.

We first consider the case that a single compression mapping is used to enroll all users' data into the database. We illustrate the setting in Fig. 2. We state in the following the formal definition of a corresponding testing scheme.

Definition 1. A testing scheme consists of a compression mapping ϕ_n^t , which enrolls the users' data sequences of length n into the database according to

$$\phi_n^t: \mathcal{X}^n \rightarrow \mathcal{M}_1, \quad (3)$$

and a decision mapping ψ_n^t , which outputs whether H_0 or H_1 is deemed true

$$\psi_n^t: \mathcal{Y}^n \times \mathcal{M}_1^M \rightarrow \{0, 1\}. \quad (4)$$

Note that the compression mapping for our membership testing scheme ϕ_n^t could be different from the compression mapping ϕ_n^{id} in Fig. 1, i.e., the testing scheme might induce additional storage space. In Section III we show that an identification scheme $(\phi_n^{\text{id}}, \psi_n^{\text{id}})$ can be turned into a testing scheme (ϕ_n^t, ψ_n^t) and vice versa such that $\phi_n^t = \phi_n^{\text{id}}$, i.e., no additional database for the pre-scan is needed.

For brevity in the following we abbreviate the ensemble of user sequences $(X^n(i))_{i=1}^M$ as $\mathbf{X}^n$ and the ensemble of enrollments $(\phi_n^t(X^n(i)))_{i=1}^M$ as $\phi_n^t(\mathbf{X}^n)$. The bold lower case

notations $\mathbf{x}^n$ and $\phi_n^t(\mathbf{x}^n)$ are used to denote the corresponding realizations. For a given compression mapping ϕ_n^t the induced distributions under both hypotheses are denoted by

$$H_0: P_{Y^n \phi_n^t(\mathbf{X}^n)}, \quad H_1: P_{Y^n} \times P_{\phi_n^t(\mathbf{X}^n)}. \quad (5)$$

The acceptance region of hypothesis H_0 is defined as

$$\mathcal{A}_n = \{(y^n, \phi_n^t(\mathbf{x}^n)) \mid \psi_n^t(y^n, \phi_n^t(\mathbf{x}^n)) = 0\}. \quad (6)$$

An error of the first (second) type occurs when y^n is related to one unknown user (independent from all users) in the system but the testing scheme declares otherwise. Accordingly, the probability of first and second type of error are given respectively as

$$\alpha_n = P_{Y^n \phi_n^t(\mathbf{X}^n)}(\mathcal{A}_n^c), \quad \beta_n = P_{Y^n} \times P_{\phi_n^t(\mathbf{X}^n)}(\mathcal{A}_n). \quad (7)$$

Definition 2. An error exponent E of type II is achievable given (R, R_c) with $M \doteq e^{nR}$ if there exist a sequence of testing schemes (ϕ_n^t, ψ_n^t) such that

$$\lim_{n \rightarrow \infty} \alpha_n = 0, \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\phi_n^t| \leq R_c, \\ \liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\beta_n} \geq E. \quad (8)$$

We define

$$E_f^*(R, R_c) = \sup\{E \mid E \text{ is achievable given } (R, R_c)\}. \quad (9)$$

Note that the case $M = 1$ in our setting corresponds to the testing against independence setting studied in [1]. We refer to it in the following as the single-user scenario.

Given a compression rate R_c we define the following function

$$R_{\max}(R_c) = \max_{\substack{U-X-Y, |U| \leq |\mathcal{X}|+1, \\ I(X;U) \leq R_c}} I(Y;U). \quad (10)$$

Different interpretations of $R_{\max}(R_c)$ appear in previous works. In [1] $R_{\max}(R_c)$ is the maximum error exponent of type II for the single-user testing against independence problem. In [3] $R_{\max}(R_c)$ characterizes the number of users that can be supported at a given compression rate R_c in the identification systems in Fig. 1 with vanishing probability of identification error.

Our first result characterizes the error exponent $E_f^*(R, R_c)$.

Theorem 1. For $E_f^*(R, R_c)$ defined in Definition 2 we have

$$E_f^*(R, R_c) = \max\{R_{\max}(R_c) - R, 0\}. \quad (11)$$

Proof: The achievability part straightforwardly follows from the single-user scenario. An outline is presented here for completeness. Let $(\bar{\phi}_n^t, \bar{\psi}_n^t)$ be an arbitrary but given single-user hypothesis testing scheme with probabilities of errors $\bar{\alpha}_n$ and $\bar{\beta}_n$. Let $\bar{\mathcal{A}}_n$ be the corresponding acceptance region for the single-user scenario. For each user $i \in [1 : M]$ the compression mapping $\bar{\phi}_n^t$ maps the data sequence $x^n(i)$ into an index that is stored in the database. We define the acceptance region for the multi-user case as follows

$$\mathcal{A}_n = \{(y^n, \bar{\phi}_n^t(\mathbf{x}^n)) \mid (y^n, \bar{\phi}_n^t(x^n(i))) \in \bar{\mathcal{A}}_n \text{ for some } i\}.$$

¹Under P_{H_0} the process is also *non-ergodic*.

The probabilities of the first and second type of errors can be bounded accordingly as

$$\alpha_n \leq \bar{\alpha}_n, \beta_n \leq \bar{\beta}_n M. \quad (12)$$

We omit the details. The existence of testing schemes that achieve the exponent $R_{\max}(R_c) - \gamma$ for the single-user scenario for any $\gamma > 0$ is well-known, cf. [1], [20]. Thus $E_f^*(R, R_c) \geq R_{\max}(R_c) - R$ for $R < R_{\max}(R_c)$ follows from (12) since $M \doteq e^{nR}$ holds.

We now present the converse for Theorem 1. Since the distributions in our setting are non-stationary, the standard weak converse proof using divergence cannot be straightforwardly extended. We will use in the following the information-spectrum method plus a covering lemma to obtain the weak converse result.

We first consider the case $R < R_{\max}(R_c)$. Suppose that $E > 0$ is an achievable error exponent, i.e., there exists a sequence of compression mappings (ϕ_n^t) and a sequence of decision mappings (ψ_n^t) such that all the conditions in Definition 2 hold. Then we can view (ψ_n^t) as decision mappings for the simple hypothesis testing problem $H_0 : P_{Y^n \phi_n^t}(\mathbf{X}^n)$ versus $H_1 : P_Y^n \times P_{\phi_n^t}(\mathbf{X}^n)$. Define two general sources $\mathfrak{C}^1 = \{C_n^1\}_{n=1}^\infty$, $\mathfrak{C}^2 = \{C_n^2\}_{n=1}^\infty$ in which for each n , $C_n^1 \sim P_{Y^n \phi_n^t}(\mathbf{X}^n)$ and $C_n^2 \sim P_{Y^n} \times P_{\phi_n^t}(\mathbf{X}^n)$ hold. By [18, Theorem 4.1.1] we have

$$E \leq \underline{D}(\mathfrak{C}^1 \parallel \mathfrak{C}^2), \quad (13)$$

where the right-hand side is the *spectral inf-divergence*, which is defined as for two general sources $\mathbf{Z} = \{Z_i\}_{i=1}^\infty$ and $\mathbf{Z}' = \{Z'_i\}_{i=1}^\infty$ such that at each n , $P_{Z^n} \ll P_{Z'^n}$ holds, as

$$\underline{D}(\mathbf{Z} \parallel \mathbf{Z}') = \sup \left\{ \alpha \mid \lim_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} \log \frac{dP_{Z^n}}{dP_{Z'^n}}(Z^n) < \alpha \right\} = 0 \right\}.$$

Using a soft-covering lemma from Corollary 1 presented in the next section with $\hat{E} = \underline{D}(\mathfrak{C}^1 \parallel \mathfrak{C}^2) - \gamma$ for any $\gamma > 0$ small enough such that $\hat{E} > 0$, $\phi_{kn}^t = \phi_n^t$ for all $k \in [1 : M]$, and ϕ_n^y be the identity mapping, we obtain

$$\lim_{n \rightarrow \infty} \Pr \{ \iota(\bar{Y}^n; \phi_n^t(\bar{X}^n)) < n(R + \hat{E} - \gamma) \} = 0, \quad (14)$$

where $(\bar{Y}^n, \bar{X}^n) \sim P_{XY}^{\otimes n}$. This implies that

$$R + \hat{E} - \gamma \leq \underline{D}(\bar{\mathfrak{C}}^1 \parallel \bar{\mathfrak{C}}^2), \quad (15)$$

in which $\bar{\mathfrak{C}}^1 = \{\bar{C}_n^1\}_{n=1}^\infty$, and $\bar{\mathfrak{C}}^2 = \{\bar{C}_n^2\}_{n=1}^\infty$ are two general source where for each n , $\bar{C}_n^1 \sim P_{\bar{Y}^n \phi_n^t(\bar{X}^n)}$ and $\bar{C}_n^2 \sim P_{\bar{Y}^n} \times P_{\phi_n^t(\bar{X}^n)}$ hold. From [18, Theorem 3.5.2] the right-hand side of (15) can be upper-bounded further by the following

$$\underline{D}(\bar{\mathfrak{C}}^1 \parallel \bar{\mathfrak{C}}^2) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} I(\bar{Y}^n; \phi_n^t(\bar{X}^n)). \quad (16)$$

Let (n_k) be a sub-sequence of indices such that $\frac{1}{n_k} I(\bar{Y}^{n_k}; \phi_{n_k}^t(\bar{X}^{n_k}))$ converges to the right-hand side of (16). With n_k sufficiently large we have $\frac{1}{n_k} \log |\phi_{n_k}^t| \leq R_c + \gamma$ and

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} I(\bar{Y}^n; \phi_n^t(\bar{X}^n)) &\leq \frac{1}{n_k} I(\bar{Y}^{n_k}; \phi_{n_k}^t(\bar{X}^{n_k})) + \gamma \\ &\leq R_{\max}(R_c + \gamma) + \gamma. \end{aligned} \quad (17)$$

The last inequality follows from the corresponding entropy characterization of $R_{\max}(R_c)$ provided in [1]. This step can also be carried out using the standard single-letterization approach. In summary we have

$$\begin{aligned} E - 3\gamma &\leq R_{\max}(R_c + \gamma) - R, \\ \xrightarrow{\gamma \rightarrow 0} E &\leq R_{\max}(R_c) - R. \end{aligned} \quad (18)$$

Since the right-hand side of the above inequality is positive, our reasoning is valid. This implies that $E_f^*(R, R_c) \leq R_{\max}(R_c) - R$, as E is arbitrary.

Now, assume that $R \geq R_{\max}(R_c)$. Let (ϕ_n^t, ψ_n^t) be a testing scheme such that $\lim_{n \rightarrow \infty} \alpha_n = 0$. Suppose further that (14) is still valid, i.e., $\underline{D}(\mathfrak{C}^1 \parallel \mathfrak{C}^2) > 0$ holds. Then by taking $\gamma \rightarrow 0$ we see that (15) is contradicted. Therefore $\underline{D}(\mathfrak{C}^1 \parallel \mathfrak{C}^2) = 0$ for all sequences (ϕ_n^t) . Therefore $E_f^*(R, R_c) = 0$ in this case. ■

Remark 1. We note that the proof of Theorem 1 remains valid when P_{XY} is jointly Gaussian. The only step that needs to be changed is (17) when an explicit calculation based on the entropy power inequality is needed. We leave the details to the interested reader.

Remark 2. Note that in the case of no compression our setting is an instance of hypothesis testing for the mixed source problem. When M is a constant, we obtain a similar result as in [18, Example 4.1.1]. Namely the optimal exponent is given by $E_f^* = I(X; Y)$ and does not depend on M . Therefore, Theorem 1 states that allowing the mixing coefficients, herein $1/M$ in P_{H_0} , to depend on n can lead to a non-trivial reduction in the error exponent of type II.

III. EQUIVALENCE RELATIONS AND THEIR CONSEQUENCES

This section is devoted to establishing operational equivalence, i.e. the correspondence between the sets of codes, between different settings: Wyner-Ahlsvede-Körner (WAK) network, single-user testing against independence, multi-user hypothesis testing and identification systems. These equivalence relations are established via simple code transformations. We summarize the results of this section in Fig. 3, in which a constructive transformation indicates an explicit construction from a starting node to an end node whereas a non-constructive transformation signifies the existence of the correspondence between these nodes.

First we show a simple code transformation argument between the WAK setting and the single-user HT in a general setting where the process $(\bar{Y}_i)_{i=-\infty}^\infty$ is *stationary* and *ergodic*. This transformation implies the strong converse, exponentially strong converse equivalence statements between these two settings.

Secondly, we show another simple code transformation between the single-user HT and the multi-user HT with user-dependent compression mappings in a general setting. The transformation allows us to establish the relations between the corresponding ϵ -achievable error exponents, hence proving the strong converse of the multi-user HT with user-dependent compression mappings in the iid discrete case. As another

Fig. 3: Code transformations across iid settings when no compression of the sequence y^n is involved, and the same compression mapping is used in the multi-user hypothesis testing and the identification setting. Solid and dashed lines indicate constructive and non-constructive transformations. Note that some transformations holds under more general assumptions.

consequence, an exponential strong converse equivalence between the single-user HT and the multi-user HT with user-dependent compression mappings is proven. We also establish the behaviors of the probabilities of errors in the multi-user HT problem in the iid discrete case when $R > R_{\max}(R_c)$.

Lastly, we derive a code transformation between the single-user HT and the identification system with user-dependent compression mappings in a general setting. As consequences of this transformation, we establish the relations between the maximum (second-order) ϵ -achievable error exponent of the single-HT and the maximum (second-order) ϵ -achievable identification rate in the identification problem. In particular these results implies the strong converse for the identification system with user-dependent compression mappings in the iid discrete scenario. As another consequence, an exponential strong converse equivalence between the single-user HT and the identification system with user-dependent compression mappings is provided. Additionally, we use the transformation to investigate the problem of redesigning the identification mapping in an under-capacity identification system.

Specifically, as a consequence of the operational equivalence it can be seen, from the red and blue curves in Fig. 3, that a code for an identification system can be used for the membership testing purpose and vice versa. When a fixed compression mapping is used in the identification problem, the result implies that we do not need to use a second database for screening users. This property is highly desirable in practice since one can add a new feature, herein membership testing, on top of an old feature (recognizing an existing user) without re-designing the whole system from the scratch.

For simplicity we drop superscripts $(\cdot)^{\text{id}}$ and $(\cdot)^t$ at most of places.

A. Equivalence between single-user HT and WAK problems

Let $(\bar{X}_i, \bar{Y}_i)_{i=-\infty}^{\infty}$ be a joint stochastic process such that the process $(\bar{Y}_i)_{i=-\infty}^{\infty}$ is *stationary* and *ergodic* on a finite or countable infinite alphabet $\mathcal{Y}$ with a finite *entropy rate* $\bar{H}_Y$. For a given $\gamma > 0$ and $n \geq 1$, we define the weakly typical set $\mathcal{A}_\gamma^n$ as follows

$$\mathcal{A}_\gamma^n = \left\{ y^n \mid \left| -\frac{1}{n} \log P_{\bar{Y}^n}(y^n) - \bar{H}_Y \right| < \gamma \right\}. \quad (19)$$

Then by the Shannon-McMillan-Breiman Theorem [21, Theorem 16.8.1] we have

$$\lim_{n \rightarrow \infty} \Pr\{\bar{Y}^n \in \mathcal{A}_\gamma^n\} = 1. \quad (20)$$

Assume that the source in the WAK problem is given by $\bar{X}^n \bar{Y}^n \sim P_{\bar{X}^n \bar{Y}^n}$. We briefly recap the definitions of a WAK-code as follows. A WAK-code consists of a pair of encoding mappings (ϕ_n, ϕ_n^y) and a decoding function ψ_n which are defined as

$$\begin{aligned} \phi_n: \mathcal{X}^n &\rightarrow \mathcal{M}, \quad \phi_n^y: \mathcal{Y}^n \rightarrow \mathcal{M}_y, \\ \psi_n: \mathcal{M} \times \mathcal{M}_y &\rightarrow \mathcal{Y}^n. \end{aligned} \quad (21)$$

The WAK problem aims to control $\Pr\{\bar{Y}^n \neq \hat{\bar{Y}}^n\}$, where $\hat{\bar{Y}}^n = \psi_n(\phi_n(\bar{X}^n), \phi_n^y(\bar{Y}^n))$. We depict the WAK setting in Fig. 4.

Fig. 4: Illustration for the WAK setting.

The single-user HT setting aims to differentiate whether the observations (x^n, y^n) are generated from hypothesis $H_0: P_{\bar{X}^n \bar{Y}^n}$ or hypothesis $H_1: P_{\bar{X}^n} \times P_{\bar{Y}^n}$. A single-user HT scheme consists of an encoding mapping $\bar{\phi}_n$ and a *stochastic* decision mapping $\bar{\psi}_n$, which are defined as follows

$$\bar{\phi}_n: \mathcal{X}^n \rightarrow \mathcal{M}, \quad \bar{\psi}_n: \mathcal{M} \times \mathcal{Y}^n \rightarrow \{0, 1\}, \quad (22)$$

in which H_0 is chosen with probability $P_{H_0|\bar{\phi}_n(x^n), y^n}$. The probabilities of error of type I and II are given as

$$\begin{aligned} \bar{\alpha}_n &= \sum_{y^n, m} P_{\bar{Y}^n \bar{\phi}_n(\bar{X}^n)}(y^n, m) (1 - P_{H_0|y^n, m}), \\ \bar{\beta}_n &= \sum_{y^n, m} P_{\bar{Y}^n} \times P_{\bar{\phi}_n(\bar{X}^n)}(y^n, m) P_{H_0|y^n, m}. \end{aligned} \quad (23)$$

The single-user HT setting is depicted in Fig. 5.

Fig. 5: Single-user hypothesis testing with one-sided compression.

We present in the following an equivalent relation between a WAK-code and a single-user HT scheme.

Theorem 2. Fix an arbitrary $\gamma > 0$. Given a WAK-code $(\phi_n, \phi_n^y, \psi_n)$, we can construct a single-user hypothesis testing scheme $(\bar{\phi}_n, \bar{\psi}_n)$ such that the corresponding error probabilities of type I and II, $(\bar{\alpha}_n, \bar{\beta}_n)$ defined as in (23), are upper bounded by

$$\begin{aligned} \bar{\alpha}_n &\leq \Pr\{\bar{Y}^n \neq \hat{\bar{Y}}^n\} + \Pr\{\bar{Y}^n \notin \mathcal{A}_\gamma^n\}, \\ \bar{\beta}_n &\leq e^{-n(\bar{H}_Y - \gamma)} |\mathcal{M}_y|, \end{aligned} \quad (24)$$

where $\mathcal{A}_\gamma^n$ is defined as in (19). Conversely, given a testing scheme $(\bar{\phi}_n, \bar{\psi}_n)$ for the single-user hypothesis testing problem with the corresponding error probabilities $(\bar{\alpha}_n, \bar{\beta}_n)$, there exists a WAK-code $(\bar{\phi}_n, \bar{\phi}_n^y, \bar{\psi}_n)$ such that

$$\Pr\{\hat{Y}^n \neq \bar{Y}^n\} \leq \bar{\alpha}_n + e^{n\hat{E}}\bar{\beta}_n + \Pr\{\bar{Y}^n \notin \mathcal{A}_\gamma^n\} + \frac{e^{n(\bar{H}_Y + \gamma - \hat{E})}}{|\mathcal{M}_Y|}, \quad (25)$$

where $\hat{E}$ is a free parameter that satisfies the condition $\bar{H}_Y > \hat{E} - \gamma$.

Proof: WAK $\Rightarrow$ Single-user HT: For a given $m \in \mathcal{M}$, define the following correctly decodable set of the WAK-code

$$\mathcal{D}_m = \{y^n \mid y^n = \psi_n(m, \phi_n^y(y^n)), y^n \in \mathcal{A}_\gamma^n\}. \quad (26)$$

Then, it is clear that for all $m \in \mathcal{M}$ we have $|\mathcal{D}_m| \leq |\mathcal{M}_Y|$ as ϕ_n^y can only take at most $|\mathcal{M}_Y|$ values. A decision region for the single-user HT, based on $\mathcal{Y}^n \times \mathcal{M}$, is defined as

$$\bar{\mathcal{A}}_n = \bigcup_m (\mathcal{D}_m \times \{m\}) \subset \mathcal{Y}^n \times \mathcal{M}. \quad (27)$$

The validity of $\bar{\mathcal{A}}_n$, i.e., the existence of a decision mapping $\bar{\psi}_n$, follows from the fact that we have full access to the sequence y^n when making a decision. We use the mapping $\bar{\phi}_n$ as the compression mapping for $\mathcal{X}^n$ in the single-user HT problem. From (27) the probability of type I of error is bounded by

$$\begin{aligned} \bar{\alpha}_n &= P_{\bar{Y}^n | \bar{\phi}_n(\bar{X}^n)}(\bar{\mathcal{A}}_n^c) \\ &\leq \Pr\{\hat{Y}^n \neq \bar{Y}^n\} + \Pr\{\bar{Y}^n \notin \mathcal{A}_\gamma^n\}, \end{aligned} \quad (28)$$

and the probability of type II of error is bounded above by

$$\begin{aligned} \bar{\beta}_n &= P_{\bar{Y}^n} \times P_{\bar{\phi}_n(\bar{X}^n)}(\bar{\mathcal{A}}_n) = \sum_m P_{\bar{\phi}_n(\bar{X}^n)}(m) P_{\bar{Y}^n}(\mathcal{D}_m) \\ &\leq \sum_m P_{\bar{\phi}_n(\bar{X}^n)}(m) |\mathcal{D}_m| e^{-n(\bar{H}_Y - \gamma)} \\ &\leq e^{-n(\bar{H}_Y - \gamma)} |\mathcal{M}_Y|. \end{aligned} \quad (29)$$

WAK $\Leftarrow$ Single-user HT: Given a single-user HT testing scheme $(\bar{\phi}_n, \bar{\psi}_n)$, for each $m \in \mathcal{M}$ we define the set

$$\mathcal{D}_n(m) = \left\{ y^n \mid \frac{1}{n} \log \frac{P_{\bar{Y}^n | \bar{\phi}_n(\bar{X}^n)}(y^n | m)}{P_{\bar{Y}^n}(y^n)} > \hat{E} \right\} \cap \mathcal{A}_\gamma^n. \quad (30)$$

$\mathcal{D}_n(m)$ plays the role of the conditionally typical set in the standard proof of the WAK setting, cf. [21], [22]. We use the mapping $\bar{\phi}_n$ as the compression mapping for $\mathcal{X}^n$ in the WAK problem. From the definition of $\mathcal{D}_n(m)$ we obtain

$$\begin{aligned} &\Pr\{\bar{Y}^n \notin \mathcal{D}_n(\bar{\phi}_n(\bar{X}^n))\} \\ &\stackrel{(*)}{\leq} \bar{\alpha}_n + e^{n\hat{E}}\bar{\beta}_n + \Pr\{\bar{Y}^n \notin \mathcal{A}_\gamma^n\}, \end{aligned} \quad (31)$$

where $(*)$ follows from [23, Lemma 12.2]. Furthermore, we have $|\mathcal{D}_n(m)| \leq e^{n(\bar{H}_Y + \gamma - \hat{E})}$ for all $m \in \mathcal{M}$, since

$$\begin{aligned} 1 &\geq P_{\bar{Y}^n | \bar{\phi}_n(\bar{X}^n)}(\mathcal{D}_n(m) | m) \\ &= \sum_{y^n \in \mathcal{D}_n(m)} P_{\bar{Y}^n | \bar{\phi}_n(\bar{X}^n)}(y^n | m) \geq \sum_{y^n \in \mathcal{D}_n(m)} P_{\bar{Y}^n}(y^n) e^{n\hat{E}} \end{aligned}$$

$$\geq |\mathcal{D}_n(m)| e^{-n(\bar{H}_Y + \gamma)} e^{n\hat{E}}. \quad (32)$$

Let $m_Y, m_y \in [1 : \mathcal{M}_Y]$, be a uniformly random bin index of y^n and $\mathcal{B}(m_Y)$ be the set of all such y^n . The decoder decides that $\hat{y}^n$ is the reconstructed sequence if it is the unique sequence such that $\hat{y}^n \in \mathcal{B}(m_Y) \cap \mathcal{D}_n(m)$, where m and m_Y are sent messages from Encoder 1 and 2. It then follows that

$$\begin{aligned} &\Pr\{\hat{Y}^n \neq \bar{Y}^n\} \\ &\leq \Pr\{\bar{Y}^n \notin \mathcal{B}(M_Y) \cap \mathcal{D}_n(\bar{\phi}_n(\bar{X}^n))\} \\ &\quad + \Pr\{\exists \tilde{y}^n \neq \bar{Y}^n, \tilde{y}^n \in \mathcal{D}_n(\bar{\phi}_n(\bar{X}^n)) \cap \mathcal{B}(M_Y)\} \\ &\stackrel{(a)}{\leq} \Pr\{\bar{Y}^n \notin \mathcal{D}_n(\bar{\phi}_n(\bar{X}^n))\} \\ &\quad + \Pr\{\exists \tilde{y}^n \neq \bar{Y}^n, \tilde{y}^n \in \mathcal{D}_n(\bar{\phi}_n(\bar{X}^n)) \cap \mathcal{B}(M_Y)\} \\ &\stackrel{(b)}{\leq} \alpha_n + e^{n\hat{E}}\beta_n + \Pr\{\bar{Y}^n \notin \mathcal{A}_\gamma^n\} + \frac{e^{n(\bar{H}_Y + \gamma - \hat{E})}}{|\mathcal{M}_Y|}. \end{aligned} \quad (33)$$

(a) is valid since $\bar{Y}^n \in \mathcal{B}(M_Y)$. The inequality (b) holds since each $\tilde{y}^n$ is assigned independently to a bin with probability $1/|\mathcal{M}_Y|$ and the number of such $\tilde{y}^n$ is bounded by $e^{n(\bar{H}_Y + \gamma - \hat{E})}$, cf. (32). The existence of deterministic mappings $\bar{\phi}_n^y$ and $\bar{\psi}_n$ follows immediately. ■

Note that the non-constructivity of the transformation from the single-user HT to the WAK problem, cf. Fig. 3, comes from the random binning argument.

Remark 3. Assume that in the WAK problem, we allow ψ_n to be stochastic with the corresponding decoding conditional probability $P_{\hat{Y}^n | \phi_n(\bar{X}^n), \phi_n^y(\bar{Y}^n)}$. Then we can define a stochastic decision mapping $\bar{\psi}_n$ for the single-user hypothesis testing as

$$\begin{aligned} &\bar{\psi}_n(y^n, \phi_n(x^n)) \\ &= 1 - \mathbf{1}\{y^n = \psi_n(\phi_n^y(y^n), \phi_n(x^n)), y^n \in \mathcal{A}_\gamma^n\}, \end{aligned} \quad (34)$$

with the corresponding conditional decision probability $P_{H_0 | \phi_n(x^n), y^n} = P_{\hat{Y}^n | \phi_n(\bar{X}^n), \bar{\phi}_n^y(\bar{Y}^n)}(y^n | \phi_n(x^n), \phi_n^y(y^n)) \times \mathbf{1}\{y^n \in \mathcal{A}_\gamma^n\}$. Then similarly we can see that the false alarm probability is upper bounded by

$$\begin{aligned} \bar{\alpha}_n &= \Pr\{\bar{\psi}_n(\bar{Y}^n, \phi_n(\bar{X}^n)) = 1\} \\ &\leq \Pr\{\hat{Y}^n \neq \bar{Y}^n\} + \Pr\{\bar{Y}^n \notin \mathcal{A}_\gamma^n\}. \end{aligned} \quad (35)$$

The miss detection probability is bounded as

$$\begin{aligned} \bar{\beta}_n &= \sum_{y^n \in \mathcal{A}_\gamma^n} \sum_{m \in \mathcal{M}} P_{\bar{Y}^n}(y^n) P_{\bar{\phi}_n(\bar{X}^n)}(m) \\ &\quad \times P_{\hat{Y}^n | \phi_n(\bar{X}^n), \bar{\phi}_n^y(\bar{Y}^n)}(y^n | m, \phi_n^y(y^n)) \\ &\leq e^{-n(\bar{H}_Y - \gamma)} \sum_{m \in \mathcal{M}} P_{\bar{\phi}_n(\bar{X}^n)}(m) \\ &\quad \times \sum_{m_Y \in \mathcal{M}_Y} \sum_{y^n \in \phi_n^{m_Y^{-1}}(m_Y)} P_{\hat{Y}^n | \phi_n(\bar{X}^n), \bar{\phi}_n^y(\bar{Y}^n)}(y^n | m, m_Y) \\ &\leq e^{-n(\bar{H}_Y - \gamma)} \sum_{m \in \mathcal{M}} P_{\bar{\phi}_n(\bar{X}^n)}(m) \sum_{m_Y \in \mathcal{M}_Y} 1 \\ &\leq e^{-n(\bar{H}_Y - \gamma)} |\mathcal{M}_Y|. \end{aligned} \quad (36)$$

Therefore the conclusion of Theorem 2 still holds in case that the decoding mapping for the WAK problem is stochastic.

As a consequence of Theorem 2 in the following we establish the connection between the ϵ -achievable regions. For that purpose we need several definitions. Fix an $\epsilon \in [0, 1)$. Let $\mathcal{R}_{\text{WAK}, \epsilon}$ be the closure of all (R_c, R_y) such that there exists a WAK-code $(\phi_n, \phi_n^y, \psi_n)$ which satisfies

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\phi_n| &\leq R_c, \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\phi_n^y| \leq R_y, \\ \limsup_{n \rightarrow \infty} \Pr\{\hat{Y}^n \neq \bar{Y}^n\} &\leq \epsilon. \end{aligned} \quad (37)$$

Define $R_{y, \epsilon}^*(R_c) = \inf\{R_y \mid (R_c, R_y) \in \mathcal{R}_{\text{WAK}, \epsilon}\}$. As it can be seen from the definition of $\mathcal{A}_\gamma^n$ in (19), cf. [18, Example 1.6.1], we have $(R_c, \bar{H}_Y) \in \mathcal{R}_{\text{WAK}, \epsilon}$, $\forall \epsilon \in [0, 1)$. This implies that $R_{y, \epsilon}^*(R_c)$ is finite for all $\epsilon \in [0, 1)$. Additionally, let us define the following. Let $\mathcal{R}_{\text{HT}, \epsilon}$ be the closure of all (R_c, E) such that there exists a single-user testing scheme $(\bar{\phi}_n, \bar{\psi}_n)$ which satisfies

$$\begin{aligned} \limsup_{n \rightarrow \infty} \bar{\alpha}_n &\leq \epsilon, \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\bar{\phi}_n| \leq R_c, \\ \liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\bar{\beta}_n} &\geq E. \end{aligned} \quad (38)$$

Define $E_\epsilon^*(R_c) = \sup\{E \mid (R_c, E) \in \mathcal{R}_{\text{HT}, \epsilon}\}$. It can be seen² that $E_\epsilon^*(R_c) \leq \bar{H}_Y$. Additionally, $E_\epsilon^*(R_c) \geq 0$ also holds since $(R_c, 0) \in \mathcal{R}_{\text{HT}, \epsilon}$, $\forall \epsilon \in [0, 1)$. As a direct consequence, the following result can be inferred from Theorem 2.

Theorem 3. For all $R_c \geq 0$ and for all $\epsilon \in [0, 1)$ we have

$$E_\epsilon^*(R_c) + R_{y, \epsilon}^*(R_c) = \bar{H}_Y. \quad (39)$$

In particular Theorem 3 implies that a strong converse for the WAK problem implies a strong converse for the single-user HT problem and vice versa. In the joint iid discrete case strong converse proofs for the WAK problem with a deterministic decoding mapping have been given in [24] and [25], while the strong converse for the single-user hypothesis testing problem with a stochastic decision mapping has been given in [26]. By Remark 3 and Theorem 3 we see that the strong converse also holds for the WAK problem in the joint iid discrete scenario when a stochastic decoding mapping is allowed.

To establish Theorem 3 we show that for a given near optimal ϵ -achievable WAK-code, we can construct a near optimal ϵ -achievable single-user testing scheme and conversely for a given near optimal ϵ -achievable single-user testing scheme, there exists a correspondingly near optimal ϵ -achievable WAK-code. The details is given in the following.

Proof: Given a $\gamma > 0$ and $\epsilon > 0$, there exists a WAK-code $(\phi_n, \phi_n^y, \psi_n)$ which satisfies all the conditions in (37) for the rate pair $(R_c + \gamma, R_{y, \epsilon}^*(R_c) + \gamma)$. This implies that for all sufficiently large n we have $|\mathcal{M}_y| \leq e^{n(R_{y, \epsilon}^*(R_c) + 2\gamma)}$. By the

first part of Theorem 2, the corresponding single-user testing scheme satisfies

$$\limsup_{n \rightarrow \infty} \bar{\alpha}_n \leq \epsilon, \quad \liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\bar{\beta}_n} \geq \bar{H}_Y - 3\gamma - R_{y, \epsilon}^*(R_c).$$

This implies that $(R_c + \gamma, \bar{H}_Y - 3\gamma - R_{y, \epsilon}^*(R_c)) \in \mathcal{R}_{\text{HT}, \epsilon}$. Since $\gamma > 0$ is arbitrary and $\mathcal{R}_{\text{HT}, \epsilon}$ is closed by definition, we obtain $E_\epsilon^*(R_c) \geq \bar{H}_Y - R_{y, \epsilon}^*(R_c)$.

Conversely, given $\gamma > 0$ and $\epsilon > 0$ there exists a single-user testing scheme $(\bar{\phi}_n, \bar{\psi}_n)$ which satisfies all the conditions in (38) for the pair $(R_c + \gamma, E_\epsilon^*(R_c) - \gamma)$. Then for all sufficiently large n we have $\bar{\beta}_n \leq e^{-n(E_\epsilon^*(R_c) - 2\gamma)}$. Choosing $\hat{E} = E_\epsilon^*(R_c) - 3\gamma$ and $|\mathcal{M}_y| = e^{n(\bar{H}_Y + 5\gamma - E_\epsilon^*(R_c))}$ we obtain by the second part of Theorem 2 that

$$\limsup_{n \rightarrow \infty} \Pr\{\hat{Y}^n \neq \bar{Y}^n\} \leq \epsilon, \quad (40)$$

which implies that $(R_c + \gamma, \bar{H}_Y + 5\gamma - E_\epsilon^*(R_c)) \in \mathcal{R}_{\text{WAK}, \epsilon}$. Since γ is arbitrary and $\mathcal{R}_{\text{WAK}, \epsilon}$ is also closed by definition, therefore we obtain $R_{y, \epsilon}^*(R_c) \leq \bar{H}_Y - E_\epsilon^*(R_c)$. ■

Another important consequence of Theorem 2, which states an equivalence of exponentially strong converse statements, is given as follows.

Theorem 4. Given an arbitrary number $E_{\max}(R_c)$ such that $E_{\max}(R_c) \in [0, \bar{H}_Y)$ holds. Assume further that $(\bar{Y}_i)$ is an iid process on a finite alphabet. The following statements are equivalent:

- 1) For any code $(\phi_n, \phi_n^y, \psi_n)$ which satisfies $\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\phi_n| \leq R_c$ and $\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\phi_n^y| \leq R_y$ in the WAK problem, if $R_y < \bar{H}_Y - E_{\max}(R_c)$, then $\Pr\{\hat{Y}^n \neq \bar{Y}^n\} \rightarrow 1$ exponentially fast at a positive convergence rate.
- 2) For any single-user HT scheme $(\bar{\phi}_n, \bar{\psi}_n)$ with $\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\bar{\phi}_n| \leq R_c$ and $\liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\bar{\beta}_n} \geq \bar{E}$, if $\bar{E} > E_{\max}(R_c)$, then $\bar{\alpha}_n \rightarrow 1$ exponentially fast at a positive convergence rate.

In the joint iid discrete scenario with $E_{\max}(R_c) = R_{\max}(R_c)$, a proof of the first statement is presented in [27], [28]. It can also be obtained independently based on the techniques in [29, Section VII].

Proof: Assume that the first statement holds. It suffices to show the second statement when $\bar{E} < \bar{H}_Y$. Let $(\bar{\phi}_n, \bar{\psi}_n)$ be a single-user hypothesis testing scheme such that $\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\bar{\phi}_n| \leq R_c$ and $\liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\bar{\beta}_n} \geq \bar{E}$ where $\bar{E} > E_{\max}(R_c)$. Let $\gamma > 0$ be small enough such that $\bar{E} - 4\gamma > E_{\max}(R_c)$. By the second part of Theorem 2, there exists a WAK-code $(\phi_n, \phi_n^y, \psi_n)$ such that with $\hat{E} = \bar{E} - 2\gamma$ and $|\mathcal{M}_y| = e^{n(\bar{H}_Y + 4\gamma - \hat{E})}$ we have

$$\Pr\{\hat{Y}^n \neq \bar{Y}^n\} \leq \bar{\alpha}_n + 2e^{-n\gamma} + \Pr\{\bar{Y}^n \notin \mathcal{A}_\gamma^n\}, \quad (41)$$

for all sufficiently large n . Since the weakly typical set $\mathcal{A}_\gamma^n$ includes the strongly typical set $\mathcal{T}_\epsilon^n$ for a fixed, positive, and small enough ϵ , the last term goes to 0 exponentially with a convergence rate of at least $2\epsilon^2$. Since $\bar{H}_Y + 4\gamma - \bar{E} < \bar{H}_Y - E_{\max}(R_c)$ the assumption implies that $\Pr\{\bar{Y}^n = \hat{Y}^n\}$ goes to 0 exponentially at a rate of $\eta > 0$, we then have $\bar{\alpha}_n \rightarrow 1$ exponentially at a positive rate.

²Assume that $\bar{E}$ is an ϵ -achievable error exponent with a corresponding single-user testing scheme $(\bar{\phi}_n, \bar{\psi}_n)$. Define $(\bar{\phi}(\bar{\mathbf{X}}), \bar{\mathbf{Y}}) = \{(\bar{\phi}_n(\bar{X}^n), \bar{Y}^n)\}_{n=1}^\infty$. Not only do we have $\bar{I}(\bar{\phi}(\bar{\mathbf{X}}); \bar{\mathbf{Y}}) \leq \min\{R_c + \gamma, \log |\mathcal{Y}|\}$ where $\gamma > 0$ is arbitrary, but also $\bar{I}(\bar{\phi}(\bar{\mathbf{X}}); \bar{\mathbf{Y}}) \leq \bar{H}(\bar{\mathbf{Y}})$ holds, where the latter is the spectral-sup entropy rate of the process $\bar{\mathbf{Y}}$. Then from the inequality $\bar{\alpha}_n + e^{n\bar{E}}\bar{\beta}_n \geq \Pr\{\bar{Y}^n \in \bar{\phi}_n(\bar{X}^n)\} \leq n\bar{E}$ we have $\bar{E} \leq \bar{I}(\bar{\phi}(\bar{\mathbf{X}}); \bar{\mathbf{Y}})$, the spectral-sup mutual information rate, since the right-hand side goes to 1 when $\bar{E} = \bar{I}(\bar{\phi}(\bar{\mathbf{X}}); \bar{\mathbf{Y}}) + \gamma$. As $(\bar{Y}_i)$ is a stationary ergodic process we have $\bar{H}(\bar{\mathbf{Y}}) = \bar{H}_Y$, which implies that $\bar{E} \leq \bar{H}_Y$.

Conversely, assume that the second statement holds. Let $(\phi_n, \phi_n^y, \psi_n)$ be a WAK-code such that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\phi_n| \leq R_c, \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\phi_n^y| \leq R_y, \quad (42)$$

where $R_y < \bar{H}_Y - E_{\max}(R_c)$. Let $\gamma > 0$ be small enough such that $R_y + 2\gamma < \bar{H}_Y - E_{\max}(R_c)$. By the first part of Theorem 2, the constructed testing scheme satisfies

$$\begin{aligned} \bar{\alpha}_n &\leq \Pr\{\bar{Y}^n \neq \hat{Y}^n\} + \Pr\{\bar{Y}^n \notin \mathcal{A}_\gamma^n\} \\ \bar{\beta}_n &\leq e^{-n(\bar{H}_Y - 2\gamma - R_y)}, \end{aligned} \quad (43)$$

for all sufficiently large n . Since $\bar{H}_Y - 2\gamma - R_y > E_{\max}(R_c)$ holds the corresponding false alarm probability $\bar{\alpha}_n$ hence goes to 1 exponentially at a rate of $\xi > 0$, or $\Pr\{\bar{Y}^n \neq \hat{Y}^n\} \rightarrow 1$ exponentially at a rate of $\min\{\xi, 2\epsilon^2\}$. ■

B. Equivalence between single-user HT and multi-user HT

In this subsection, we consider the relation between the multi-user hypothesis testing problem and the single-user hypothesis testing problem under a more general setup as follows.

Two hypotheses in the single-user HT problem are $H_0: P_{\bar{X}^n \bar{Y}^n}$, and $H_1: P_{\bar{X}^n} \times P_{\bar{Y}^n}$, where $P_{\bar{X}^n \bar{Y}^n}$ is a general distribution on a product space $\mathcal{X}_n \times \mathcal{Y}_n$. Note that we do not assume either $P_{\bar{X}^n \bar{Y}^n} = P_{\bar{X}^n}^{\otimes n} \times P_{\bar{Y}^n}^{\otimes n}$, or $\mathcal{X}_n$ and $\mathcal{Y}_n$ are finite. Suppose further that $\mathcal{X}_n$ and $\mathcal{Y}_n$ are Polish spaces. With abuse of terminology and notation we will redefine some terms and notations in the subsequent development. A single-user testing scheme consists of two compression mappings $(\bar{\phi}_n^x, \bar{\phi}_n^y)$ and a stochastic decision mapping $\bar{\psi}_n$ where

$$\begin{aligned} \bar{\phi}_n^x: \mathcal{X}_n &\rightarrow \mathcal{M}, \quad \bar{\phi}_n^y: \mathcal{Y}_n \rightarrow \mathcal{M}_y, \\ \bar{\psi}_n: \mathcal{M} \times \mathcal{M}_y &\rightarrow \{0, 1\}. \end{aligned} \quad (44)$$

The probabilities of type I and II errors $\bar{\alpha}_n$ and $\bar{\beta}_n$ can also be determined similarly as in (23). For the joint iid, discrete case with a *deterministic* decision mapping, this setup was discussed briefly in [1], for which a single-letter characterization for the optimal achievable error exponent of type II of error is an open question. The single-user hypothesis testing setting is illustrated in Fig. 6.

Fig. 6: Single-user hypothesis testing with two-sided compression.

The underlying distributions of the multi-user hypothesis testing problem are given similarly as in Eqs. (1) and (2) in which the joint distributions $P_{Y^n X^n(i)}$ are given by $P_{Y^n X^n(i)} = P_{\bar{Y}^n \bar{X}^n}$ for all $i \in [1 : M]$. We consider the case that each user has its own compression mapping ϕ_{kn} , for $k \in [1 : M]$. This setting for instance represents the case where the hypothesis testing database $(\phi_{kn}(X^n(k)))$ is a

merger of different, distributed sub-databases, which employ different compression mechanisms. For simplicity we denote by $\phi_n(\mathbf{X}^n)$ the ensemble $(\phi_{1n}(X^n(1)), \dots, \phi_{Mn}(X^n(M)))$, by $\phi_n(\mathbf{x}^n)$ a corresponding realization of $\phi_n(\mathbf{X}^n)$ and by ϕ_n the multiset of mappings $(\phi_{kn})_{k=1}^M$. Furthermore the observation sequence y^n is also compressed before sending to the decision center via a mapping $\phi_n^y: \mathcal{Y}^n \rightarrow \mathcal{M}_y$. These mappings induce the following distributions for our hypothesis testing problem.

$$\begin{aligned} H_0: P_{\phi_n^y(Y^n) \phi_n(\mathbf{X}^n)} \\ &= \frac{1}{M} \sum_{k=1}^M P_{\phi_n^y(Y^n) \phi_{kn}(X^n(k))} \times \prod_{j \neq k} P_{\phi_{jn}(X^n(j))}, \\ H_1: P_{\phi_n^y(Y^n)} \times P_{\phi_n(\mathbf{X}^n)} &= P_{\phi_n^y(Y^n)} \times \prod_{k=1}^M P_{\phi_{kn}(X^n(k))}. \end{aligned}$$

We use a *stochastic* decision mapping $\psi_n \mapsto \{0, 1\}$ with the corresponding $P_{H_0 | \phi_n^y(y^n), \phi_n(\mathbf{x}^n)} \mapsto [0, 1]$ be the probability that H_0 is chosen given $(\phi_n^y(y^n), \phi_n(\mathbf{x}^n))$. Then the false alarm and miss detection probabilities $\alpha_{n, \text{var}}, \beta_{n, \text{var}}$ can be determined similarly as in (23). The multi-user hypothesis testing setting is depicted in Fig. 7.

Fig. 7: Multi-user hypothesis testing with user-dependent compression mappings ϕ_n and a compression mapping for the observation sequence ϕ_n^y .

To establish the equivalence between the multi-user hypothesis testing problem and the single-user hypothesis testing problem we need the following support lemma. It resembles the soft covering using excess relative information metric in [30, Theorem 24].

Lemma 1. For each $i \in [1 : M]$ define the following distribution $Q_i = P_{ZV_i} \times \prod_{j \neq i} P_{V_j}$. Furthermore, for each i let $(Z^i, V_1^i, \dots, V_M^i)$ be a tuple of random variables that follows the law Q_i . For brevity we denote an outcome in the corresponding sample space $(z, v_1, \dots, v_m)$ by $(z, \mathbf{v})$. Then we define two distributions of interest as follows

$$Q_{H_0} = \frac{1}{M} \sum_{i=1}^M Q_i, \quad \text{and} \quad Q_{H_1} = P_Z \times \prod_{k=1}^M P_{V_k}. \quad (45)$$

Assume that for all $k \in [1 : M]$, $P_{ZV_k} \ll P_Z \times P_{V_k}$ holds. Given positive numbers η, σ such that $\eta > \sigma$ we have

$$Q_{H_0} \left(\left\{ (z, \mathbf{v}) \mid \frac{dQ_{H_0}}{dQ_{H_1}}(z, \mathbf{v}) > \eta \right\} \right)$$

$$\leq \frac{1}{M} \sum_{k=1}^M \Pr \left\{ \frac{dP_{ZV_k}}{d(P_Z \times P_{V_k})}(Z^k, V_k^k) > M\sigma \right\} + \frac{1}{\eta - \sigma}.$$

For a given tuple $\mathbf{v}$, the distribution $\frac{1}{M} \sum_{i=1}^M P_{Z|V_i=v_i}$ can be seen as the output distribution of a general channel that depends on which input is picked. Lemma 1 therefore can still be seen as a soft-covering lemma. The proof is given in the following.

Proof: First we have

$$\begin{aligned} Q_{H_0} \left(\{(z, \mathbf{v}) \mid \frac{dQ_{H_0}}{dQ_{H_1}}(z, \mathbf{v}) > \eta\} \right) \\ = \frac{1}{M} \sum_{i=1}^M Q_i \left(\{(z, \mathbf{v}) \mid \frac{dQ_{H_0}}{dQ_{H_1}}(z, \mathbf{v}) > \eta\} \right). \end{aligned} \quad (46)$$

Since

$$\frac{dQ_{H_0}}{dQ_{H_1}}(z, \mathbf{v}) = \frac{1}{M} \sum_{i=1}^M \frac{dP_{ZV_i}}{d(P_Z \times P_{V_i})}(z, v_i), \quad Q_{H_1}\text{-a.s.}, \quad (47)$$

we have

$$\begin{aligned} Q_i \left(\{(z, \mathbf{v}) \mid \frac{dQ_{H_0}}{dQ_{H_1}}(z, \mathbf{v}) > \eta\} \right) \\ = Q_i \left(\{(z, \mathbf{v}) \mid \frac{1}{M} \sum_{k=1}^M \frac{dP_{ZV_k}}{d(P_Z \times P_{V_k})}(z, v_k) > \eta\} \right) \\ \leq \Pr \left\{ \frac{dP_{ZV_i}}{d(P_Z \times P_{V_i})}(Z^i, V_i^i) > M\sigma \right\} \\ + \Pr \left\{ \frac{1}{M} \sum_{k \neq i} \frac{dP_{ZV_k}}{d(P_Z \times P_{V_k})}(Z^i, V_k^i) > \eta - \sigma \right\}. \end{aligned} \quad (48)$$

Since for $k \neq i$ we have $(Z^i, V_k^i) \sim P_Z \times P_{V_k}$, the following equality is valid

$$\mathbb{E} \left[\frac{dP_{ZV_k}}{d(P_Z \times P_{V_k})}(Z^i, V_k^i) \right] = 1. \quad (49)$$

The second term can be upper bounded by Markov's inequality as

$$\begin{aligned} \Pr \left\{ \frac{1}{M} \sum_{k \neq i} \frac{dP_{ZV_k}}{d(P_Z \times P_{V_k})}(Z^i, V_k^i) > \eta - \sigma \right\} \\ \leq \frac{1}{\eta - \sigma} \frac{1}{M} \sum_{k \neq i} \mathbb{E} \left[\frac{dP_{ZV_k}}{d(P_Z \times P_{V_k})}(Z^i, V_k^i) \right] \leq \frac{1}{\eta - \sigma}. \end{aligned} \quad (50)$$

Hence everything is done. $\blacksquare$

Remark 4. By closely following the proof of [30, Theorem 24] the inequality stated in Lemma 1 can be generalized into

$$\begin{aligned} Q_{H_0} \left(\{(z, \mathbf{v}) \mid \frac{dQ_{H_0}}{dQ_{H_1}}(z, \mathbf{v}) > \eta\} \right) \\ \leq \frac{1}{M} \sum_{k=1}^M \left[\Pr \left\{ \frac{dP_{ZV_k}}{d(P_Z \times P_{V_k})}(Z^k, V_k^k) > M\sigma \right\} \right. \\ \left. + \frac{1}{\nu} \Pr \left\{ \frac{dP_{ZV_k}}{d(P_Z \times P_{V_k})}(Z^k, V_k^k) \geq M \exp(-\tau) \right\} \right] \\ + \frac{\exp(-\tau)}{(\eta - \sigma - \nu - 1)^2}, \end{aligned} \quad (51)$$

where $\nu > 0$ such that $\eta - \sigma - \nu - 1 > 0$ and $\tau \in \mathbb{R}$ is arbitrary. This generalization is more than what is needed in our setting, hence we choose to present in Lemma 1 a simplified version instead.

In the subsequent development we assume that the mappings (ϕ_n, ϕ_n^y) are chosen such that the following condition holds³

$$\begin{aligned} P_{\phi_n^y(\bar{Y}^n) \phi_{kn}(\bar{X}^n)} \\ \ll P_{\phi_n^y(\bar{Y}^n)} \times P_{\phi_{kn}(\bar{X}^n)}, \quad \forall k \in [1 : M]. \end{aligned} \quad (52)$$

By choosing $\eta = e^{n\hat{E}}$ and $\sigma = \eta/2$, Lemma 1 implies the following result.

Corollary 1. For a given doubly indexed sequence (ϕ_{kn}) , a given sequence (ϕ_n^y) , and a given positive number $\hat{E}$ the following holds

$$\begin{aligned} \Pr \{ \iota_{\phi_n^y(Y^n) \phi_n(\mathbf{X}^n)}(\phi_n^y(Y^n); \phi_n(\mathbf{X}^n)) > n\hat{E} \} \\ \leq \frac{1}{M} \sum_{k=1}^M \Pr \{ \iota_{\phi_n^y(\bar{Y}^n) \phi_{kn}(\bar{X}^n)}(\phi_n^y(\bar{Y}^n); \phi_{kn}(\bar{X}^n)) \\ > \log M + n\hat{E} - \log 2 \} + 2 \exp(-n\hat{E}), \end{aligned} \quad (53)$$

where $(\phi_n^y(Y^n), \phi_n(\mathbf{X}^n)) \sim P_{\phi_n^y(Y^n) \phi_n(\mathbf{X}^n)}$ and $(\bar{Y}^n, \bar{X}^n) \sim P_{\bar{Y}^n \bar{X}^n}$. We denote the left-hand side in (53) as $L_{n, \text{var}}$.

With the help of Corollary 1 we establish the equivalence between the single-user and the multi-user hypothesis testing problem in the following theorem.

Theorem 5. Fix $\hat{E} > 0$. Given a single-user testing scheme $(\bar{\phi}_n, \bar{\phi}_n^y, \bar{\psi}_n)$ with the corresponding error probabilities $(\bar{\alpha}_n, \bar{\beta}_n)$, we can construct a multi-user testing scheme $(\phi_n, \phi_n^y, \psi_n)$ such that the multi-user error probabilities are bounded by

$$\alpha_n \leq \bar{\alpha}_n, \quad \beta_n \leq M\bar{\beta}_n. \quad (54)$$

Conversely, given a multi-user testing scheme $(\phi_n, \phi_n^y, \psi_n)$ with the corresponding probabilities of errors (α_n, β_n) , we can construct a single-user testing scheme $(\bar{\phi}_n, \bar{\phi}_n^y, \bar{\psi}_n)$ such that for each n , $\bar{\phi}_n \in \phi_n$ holds, and the corresponding probabilities of errors are given by

$$\begin{aligned} \bar{\alpha}_n &\leq \alpha_n + e^{n\hat{E}} \beta_n + 2 \exp(-n\hat{E}) \\ \bar{\beta}_n &\leq e^{-n\hat{E}} \frac{2}{M}. \end{aligned} \quad (55)$$

Proof: Single-user HT $\Rightarrow$ Multi-user HT: Given the single-user testing scheme $(\bar{\phi}_n, \bar{\phi}_n^y, \bar{\psi}_n)$, we use the mapping $\bar{\phi}_n$ to compress user's information inside the database and the mapping $\bar{\phi}_n^y$ to compress the observation sequence in the multi-user hypothesis testing problem, i.e.,

$$\phi_{kn} = \bar{\phi}_n, \quad \forall k = 1, \dots, M, \quad \phi_n^y = \bar{\phi}_n^y. \quad (56)$$

³Condition (52) excludes cases, for example when ϕ_n^y, ϕ_n are the identity mappings and $P_{\bar{Y}^n | \bar{X}^n = x^n} = \delta_{x^n}$ as well as $P_{\bar{X}^n} \ll \lambda^{\otimes n}$ hold, where δ_{x^n} and $\lambda^{\otimes n}$ are the Dirac and the (product) Lebesgue measures. If $\phi_n^y(\bar{Y}^n)$ takes values on a countable alphabet then (52) is satisfied for arbitrary ϕ_n . Similarly if all $\phi_{kn}(\bar{X}^n)$ take values on countable alphabets then (52) holds.

A decision mapping for the multi-user hypothesis testing problem is given as follows:

$$\begin{aligned} \psi_n(\phi_n^y(y^n), \phi_n(\mathbf{x})) &= 0, \text{ if } \exists k \in [1 : M], \\ \text{such that } \bar{\psi}_n(\bar{\phi}_n^y(y^n), \bar{\phi}_n(x^n(k))) &= 0. \end{aligned} \quad (57)$$

The rest of the proof of the direct part is similar to the one of Theorem 1 and hence omitted.

Single-user HT $\Leftarrow$ Multi-user HT: Given a multi-user testing scheme $(\phi_n, \phi_n^y, \psi_n)$ for each n and $k \in [1 : M]$, define the following set

$$\begin{aligned} \bar{\mathcal{A}}_{kn} &= \{(y^n, x^n) \mid \iota_{\phi_n^y(Y^n)\phi_{kn}(\bar{X}^n)}(\phi_n^y(y^n); \phi_{kn}(x^n)) \\ &> \log M + n\hat{E} - \log 2\}. \end{aligned} \quad (58)$$

Let $k^*(n)$ be such that

$$P_{\bar{Y}^n \bar{X}^n}(\bar{\mathcal{A}}_{k^*(n)}^c) = \min_{k \in [1:M]} P_{\bar{Y}^n \bar{X}^n}(\bar{\mathcal{A}}_{kn}^c) \quad (59)$$

We define a sequence of compression mappings $(\bar{\phi}_n)$ for the single-user hypothesis testing problem by $\bar{\phi}_n = \phi_{k^*(n)}$, $n = 1, 2, \dots$. We also use the mapping ϕ_n^y to compress the sequence y^n in the single-user hypothesis testing problem. The decision region for the single-user hypothesis testing problem is defined to be $\bar{\mathcal{A}}_n = \bar{\mathcal{A}}_{k^*(n)}$. By [23, Lemma 12.2] and Corollary 1 we obtain

$$\begin{aligned} \alpha_n + e^{n\hat{E}} \beta_n &\geq 1 - L_{n,\text{var}} \geq P_{\bar{Y}^n \bar{X}^n}(\bar{\mathcal{A}}_n^c) - 2 \exp(-n\hat{E}) \\ &= \bar{\alpha}_n - 2 \exp(-n\hat{E}). \end{aligned} \quad (60)$$

By the change of measure we also have $\bar{\beta}_n \leq 2e^{-n\hat{E}}/M$. ■

In the following we present two consequences of Theorem 5. In the first result we establish the relation between the ϵ -achievable error exponent of the multi-user hypothesis testing problem and the one of the single-user hypothesis testing problem. To enable the characterization in this setting we assume a *uniformity condition* that all the mappings ϕ_{kn} admit the same compression rate. We state in the following the corresponding ϵ -achievable definition.

Definition 3. An error exponent E of type II is ϵ -achievable for a given triple (R, R_c, R_y) with $M \doteq e^{nR}$ if there exist a doubly indexed sequence (ϕ_{kn}) of enrollment mappings, a sequence of compression mappings for the observation sequence (ϕ_n^y) and a sequence of decision mappings (ψ_n) such that

$$\begin{aligned} \limsup_{n \rightarrow \infty} \max_{k \in [1:M]} \frac{1}{n} \log |\phi_{kn}| &\leq R_c, \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\phi_n^y| \leq R_y, \\ \limsup_{n \rightarrow \infty} \alpha_{n,\text{var}} &\leq \epsilon, \quad \liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\beta_{n,\text{var}}} \geq E. \end{aligned} \quad (61)$$

Let $E_{\epsilon,\text{var}}^*(R, R_c, R_y)$ be the supremum of all ϵ -achievable error exponent for a given triple (R, R_c, R_y) . If we restrict the collections of enrollment mappings to be such that for all $k \in [1 : M]$ we have $\phi_{kn} = \phi_n$ for some mapping ϕ_n then we denote the corresponding maximum error exponent of type II by $E_{\epsilon,\text{f}}^*(R, R_c, R_y)$.

When the compression mapping for the observation sequence ϕ_n^y is restricted to be $\phi_n^y = \text{id}$, the identity mapping, we denote the corresponding quantities by $E_{\epsilon,\text{var}}^*(R, R_c)$, and

$E_{\epsilon,\text{f}}^*(R, R_c)$. Similarly, in the single-user case, we use the notation $E_{\epsilon,\text{s}}^*(R_c, R_y)$ to denote the supremum ϵ -achievable error exponent of type II for a given compression rate pair (R_c, R_y) . It can be seen that for a given compression rate R_c , $E_{\epsilon,\text{s}}^*(R_c, R_y)$ is finite, cf. Footnote 2. Then we have the following theorem which generalizes Theorem 1.

Theorem 6. When $\{P_{\bar{Y}^n \bar{X}^n}\}_{n=1}^\infty$ is a sequence of general distributions, we have for all $\epsilon \in [0, 1)$, and all $(R, R_c, R_y) \in \mathbb{R}_+^3$,

$$\begin{aligned} E_{\epsilon,\text{var}}^*(R, R_c, R_y) &= E_{\epsilon,\text{f}}^*(R, R_c, R_y) \\ &= \max\{E_{\epsilon,\text{s}}^*(R_c, R_y) - R, 0\}. \end{aligned} \quad (62)$$

In particular, when $P_{\bar{Y}^n \bar{X}^n} = P_{YX}^{\otimes n}$ where P_{XY} is a joint distribution on a finite alphabet $\mathcal{X} \times \mathcal{Y}$, and $\phi_n^y = \text{id}$, then for all $\epsilon \in [0, 1)$, and $(R_c, R) \in \mathbb{R}_+^2$ we have

$$\begin{aligned} E_{\epsilon,\text{var}}^*(R, R_c) &= E_{\epsilon,\text{f}}^*(R, R_c) \\ &= \max\{R_{\max}(R_c) - R, 0\}. \end{aligned} \quad (63)$$

Proof: Assume that $R < E_{\epsilon,\text{s}}^*(R_c, R_y)$ holds. By the first part of Theorem 5 we have $E_{\epsilon,\text{f}}^*(R, R_c, R_y) \geq E_{\epsilon,\text{s}}^*(R_c, R_y) - R$. Therefore for all (R, R_c, R_y) we have

$$E_{\epsilon,\text{f}}^*(R, R_c, R_y) \geq \max\{E_{\epsilon,\text{s}}^*(R_c, R_y) - R, 0\}. \quad (64)$$

Assume for now that E is an ϵ -achievable error exponent of type II for the multi-user testing problem with user-dependent compression mappings, then there exist a multi-user testing scheme $(\phi_n, \phi_n^y, \psi_n)$ such that all the conditions in (61) are fulfilled. Applying the converse part of Theorem 5 with $\hat{E} = E - 2\gamma$ we obtain a single-user testing scheme $(\bar{\phi}_n, \bar{\phi}_n^y, \bar{\psi}_n)$ such that

$$\begin{aligned} \bar{\phi}_n &\in \phi_n, \quad \bar{\alpha}_n \leq \alpha_n + e^{-n\gamma} + 2 \exp(-n(E - 2\gamma)), \\ \bar{\beta}_n &\leq \exp(-n(E - 4\gamma + R)), \end{aligned} \quad (65)$$

for all sufficiently large n . Consider first the case $R < E_{\epsilon,\text{s}}^*(R_c, R_y)$. Assume further that $E > 0$, which is plausible if $E_{\epsilon,\text{var}}^*(R, R_c, R_y) > 0$, then with $\gamma > 0$ sufficiently small we obtain

$$\begin{aligned} \limsup_{n \rightarrow \infty} \bar{\alpha}_n &\leq \epsilon, \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\bar{\phi}_n| \leq R_c, \\ \liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\bar{\beta}_n} &\geq E - 4\gamma + R. \end{aligned} \quad (66)$$

Therefore we have $E - 4\gamma + R \leq E_{\epsilon,\text{s}}^*(R_c, R_y)$. Accordingly, we must have $E_{\epsilon,\text{var}}^*(R, R_c, R_y) \leq E_{\epsilon,\text{s}}^*(R_c, R_y) - R$ for this case, which implies that $E_{\epsilon,\text{var}}^*(R, R_c, R_y) = E_{\epsilon,\text{f}}^*(R, R_c, R_y) = E_{\epsilon,\text{s}}^*(R_c, R_y) - R$ when $R < E_{\epsilon,\text{s}}^*(R_c, R_y)$.

Next, consider the case that $R \geq E_{\epsilon,\text{s}}^*(R_c, R_y)$. Assume that $E_{\epsilon,\text{var}}^*(R, R_c, R_y) > 0$ still holds. Then again we have $E - 3\gamma + R \leq E_{\epsilon,\text{s}}^*(R_c, R_y)$. By taking $\gamma \rightarrow 0$ we see that the last inequality is not valid for $0 < E < E_{\epsilon,\text{var}}^*(R, R_c, R_y)$. Therefore $E_{\epsilon,\text{var}}^*(R, R_c, R_y) = 0$ in this case. ■

The following result shows an equivalence for the exponentially strong converse between our multi-user testing with user-dependent compression mapping problem and the single-user testing against independence problem.

Theorem 7. For an arbitrarily given positive number $E_{\max}(R_c, R_y)$, assume that $R < E_{\max}(R_c, R_y)$ then the two following statements are equivalent.

- 1) For any single-user HT scheme $(\bar{\phi}_n, \bar{\phi}_n^y, \bar{\psi}_n)$ with $\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\bar{\phi}_n| \leq R_c$, $\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\bar{\phi}_n^y| \leq R_y$, and $\liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\bar{\beta}_n} \geq \bar{E}$, if $\bar{E} > E_{\max}(R_c, R_y)$, then $\bar{\alpha}_n \rightarrow 1$ exponentially fast at a positive convergence rate.
- 2) For any multi-user HT scheme $(\phi_n, \phi_n^y, \psi_n)$ with $\limsup_{n \rightarrow \infty} \max_{k \in [1:M]} \frac{1}{n} \log |\phi_{kn}| \leq R_c$, $\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\phi_{kn}^y| \leq R_y$, and $\liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\beta_n} \geq E$, if $E > E_{\max}(R_c, R_y) - R$, then $\alpha_n \rightarrow 1$ exponentially fast at a positive convergence rate.

If we assume further that both $\phi_n^y = \text{id}$, $\bar{\phi}_n^y = \text{id}$, and $P_{\bar{X}^n \bar{Y}^n} = P_{XY}^{\otimes n}$ hold, where id is the identity mapping and P_{XY} is a distribution on a finite alphabet $\mathcal{X} \times \mathcal{Y}$, then by using Theorem 4 we have the following conclusion. Given $R < R_{\max}(R_c)$ then for any multi-user testing scheme with user-dependent compression mappings at the compression rate R_c such that the exponent of the miss detection probability is greater than $R_{\max}(R_c) - R$ then the false alarm probability goes to 1 exponentially.

Proof: Assume that the first statement holds. Let $(\phi_n, \phi_n^y, \psi_n)$ be an arbitrary multi-user testing scheme such that

$$\limsup_{n \rightarrow \infty} \max_{k \in [1:M]} \frac{1}{n} \log |\phi_{kn}| \leq R_c, \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\phi_{kn}^y| \leq R_y, \\ \text{and } \liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\beta_n} \geq E,$$

hold where $E > E_{\max}(R_c, R_y) - R$. Select $\hat{E} = E - 2\gamma$ for $\gamma > 0$ small enough such that $E + R - 4\gamma > E_{\max}(R_c, R_y)$. By applying Theorem 5 we obtain a single-user hypothesis testing scheme $(\bar{\phi}_n, \bar{\phi}_n^y, \bar{\psi}_n)$ such that $\bar{\phi}_n \in \phi_n$ and for all sufficiently large n the corresponding single-user false alarm and miss detection probabilities $(\bar{\alpha}_n, \bar{\beta}_n)$ are bounded by

$$\bar{\alpha}_n \leq \alpha_n + e^{-n\gamma} + 2 \exp(-n(E - 2\gamma)) \\ \bar{\beta}_n \leq \exp(-n(E + R - 4\gamma)). \quad (67)$$

Since $E + R - 4\gamma > E_{\max}(R_c, R_y)$ holds, the assumption implies that $\bar{\alpha}_n \rightarrow 1$ exponentially at a positive rate. Hence $\alpha_n \rightarrow 1$ exponentially at a positive rate.

Assume now that the second statement holds. Let $(\bar{\phi}_n, \bar{\phi}_n^y, \bar{\psi}_n)$ be an arbitrary single-user hypothesis testing scheme such that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\bar{\phi}_n| \leq R_c, \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\bar{\phi}_n^y| \leq R_y, \\ \text{and } \liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\bar{\beta}_n} \geq \bar{E},$$

where $\bar{E} > E_{\max}(R_c, R_y)$. Select an arbitrary $\gamma > 0$ small enough such that $\bar{E} - 2\gamma > E_{\max}(R_c, R_y)$. Then for all sufficiently large n we obtain by applying Theorem 5 a multi-user testing scheme such that the corresponding multi-user false alarm and miss detection probabilities (α_n, β_n) are upper bounded by

$$\alpha_n \leq \bar{\alpha}_n, \quad \beta_n \leq e^{-n(\bar{E} - 2\gamma - R)}. \quad (68)$$

Since $\bar{E} - R - 2\gamma > E_{\max}(R_c, R_y) - R$, the assumption implies that $\bar{\alpha}_n$ tends to 1 exponentially fast at a positive rate. Hence $\bar{\alpha}_n$ goes to 1 exponentially fast at a positive rate. ■

In the next theorem we provide some partial information about the behavior of errors of type I and II when the number of users exceeds the identifiable threshold, i.e., $R > R_{\max}(R_c)$.

Theorem 8. Suppose that $P_{\bar{Y}^n \bar{X}^n} = P_{YX}^{\otimes n}$ and $\phi_n^y = \text{id}$ hold where P_{XY} is a joint distribution on a finite alphabet $\mathcal{X} \times \mathcal{Y}$. Assume further that $R = R_{\max}(R_c) + 3\gamma$ where $\gamma > 0$ is arbitrary but given. Given any doubly indexed sequence (ϕ_{kn}) such that $\limsup_{n \rightarrow \infty} \max_{k \in [1:M]} \frac{1}{n} \log |\phi_{kn}| \leq R_c$ holds, then for any sequence of decision mappings (ψ_n) for the multi-user hypothesis testing problem we have

$$\lim_{n \rightarrow \infty} (\alpha_{n, \text{var}} + \beta_{n, \text{var}}) = 1. \quad (69)$$

Proof: Similarly as in the proof of Theorem 5 we define

$$\bar{\mathcal{A}}_{kn} = \{(y^n, x^n) \mid \iota_{\bar{Y}^n \phi_{kn}(\bar{X}^n)}(y^n; \phi_{kn}(x^n)) \\ > n(R_{\max}(R_c) + \gamma)\}, \quad \forall k \in [1:M].$$

By [31, Theorem VII.1] we obtain for all sufficiently large n that

$$\|P_{Y^n \phi_n(\mathbf{X}^n)} - P_{Y^n} \times P_{\phi_n(\mathbf{X}^n)}\|_{TV} \\ \leq \frac{1}{M} \sum_{k=1}^M P_{YX}^{\otimes n} \left(\{(y^n, x^n) \mid \iota_{\bar{Y}^n \phi_{kn}(\bar{X}^n)}(y^n; \phi_{kn}(x^n)) \\ > \log M - n\gamma\} \right) + e^{-n\gamma/2}/2 \\ \leq \frac{1}{M} \sum_{k=1}^M P_{YX}^{\otimes n}(\bar{\mathcal{A}}_{kn}) + e^{-n\gamma/2}/2. \quad (70)$$

For each n let $k^*(n)$ be such that

$$P_{YX}^{\otimes n}(\bar{\mathcal{A}}_{k^*n}) = \max_{k \in [1:M]} P_{YX}^{\otimes n}(\bar{\mathcal{A}}_{kn}). \quad (71)$$

Define $\bar{\phi}_n = \phi_{k^*n}$ and $\bar{\mathcal{A}}_n = \bar{\mathcal{A}}_{k^*n}$. We have

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\bar{\phi}_n| \leq R_c, \\ \liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{P_{Y^n} \times P_{\bar{\mathcal{A}}_n}^{\otimes n}} \geq R_{\max}(R_c) + \gamma. \quad (72)$$

By the second part of Theorem 4 we have $\lim_{n \rightarrow \infty} P_{YX}^{\otimes n}(\bar{\mathcal{A}}_n) = 0$ which implies that

$$\lim_{n \rightarrow \infty} \|P_{Y^n \phi_n(\mathbf{X}^n)} - P_{Y^n} \times P_{\phi_n(\mathbf{X}^n)}\|_{TV} = 0. \quad (73)$$

Hence the conclusion follows. ■

Remark 5. A related identification model was studied in [1, Section V], which is an extension of the one studied in [32]. Ahlswede and Csiszár considered a database $(x^n(i))_{i=1}^M \subset \mathcal{X}^{nM}$ iid generated from $P_X^{\otimes n}$ of M users. Under hypothesis H_0 , which happens with probability $\pi_0 > 0$, an observation y^n is related to a user i according to $P_{Y|X}^{\otimes n}$. While under hypothesis H_1 , which happens with probability $1 - \pi_0$, y^n is independent of all $(x^n(i))_{i=1}^M$. The authors approached the problem using the Bayesian framework and used the following

specific search procedure to calculate the expected cost. Given y^n , the processing searches for a list of indices j such that $x^n(j) \in \mathcal{G}(y^n) \subset \mathcal{X}^n$. The decision incurs a cost c for each matched index irrespective of whether H_0 or H_1 is true. Additionally, if H_0 is true a cost k is incurred if the true one is not in the list, i.e., $x^n(i) \notin \mathcal{G}(y^n)$. It was shown that the total expected cost can be expressed in terms of error probabilities of type I and II of testing P_{YX} against $P_Y \times P_X$ as $C_n = c(M - \pi_0)\beta_n + \pi_0(k - c)\alpha_n + \pi_0 c$. The objective of [1, Section V] is to obtain the minimum cost $C_n^*(R_c)$ of the total expected cost C_n when all data sequences $(x^n(i))_{i=1}^M$ are compressed using the same compression mapping at a rate R_c . When the false alarm cost is greater than the miss detection cost, i.e., $k > c$, an upper bound and a lower bound on $C_n^*(R_c)$ are given when $R < R_{\max}(R_c)$ and $R > R_{\max}(R_c)$, respectively. In contrast, we formulate the problem according to the Neyman-Pearson framework in which we do not assume any specific decision making procedure. Additionally, with this approach we have characterized the optimal error exponents when each user can have a distinct compression mapping.

C. Equivalence between single-user HT and Identification Systems

Let W be the hidden random variable that characterizes the uniformly chosen user in the identification setting in Fig. 1 which is independent of users' sequences $(X^n(i))_{i=1}^{M^{\text{id}}}$. The underlying joint probability distribution is given by

$$P_{Y^n \mathbf{X}^n W}(y^n, \mathbf{x}^n, w) = \frac{1}{M^{\text{id}}} P_{\bar{Y}^n \bar{X}^n}(y^n, x^n(w)) \times \prod_{k \neq w} P_{\bar{X}^n}(x^n(k)), \quad (74)$$

where $P_{\bar{Y}^n \bar{X}^n}$ is a general distribution on $\mathcal{X}_n \times \mathcal{Y}_n$, in which $\mathcal{X}_n$ and $\mathcal{Y}_n$ are Polish spaces. An identification scheme consists of a multiset of compression mappings for users' data $(\phi_{kn})_{k=1}^{M^{\text{id}}}$, a compression mapping for the observation sequence ϕ_n^y , and a stochastic decoding mapping ψ_n , where

$$\begin{aligned} \phi_{kn}: \mathcal{X}^n &\rightarrow \mathcal{M}_k, k \in [1 : M^{\text{id}}], \quad \phi_n^y: \mathcal{Y}^n \rightarrow \mathcal{M}_y, \\ \psi_n: \prod_{k=1}^{M^{\text{id}}} \mathcal{M}_k \times \mathcal{M}_y &\rightarrow \{1, \dots, M^{\text{id}}\} \cup \{e\}. \end{aligned} \quad (75)$$

Similarly we denote by $\phi_n(\mathbf{X}^n)$ the ensemble $(\phi_{1n}(X^n(1)), \dots, \phi_{M^{\text{id}}n}(X^n(M^{\text{id}})))$, $\phi_n(\mathbf{x}^n)$ a corresponding realization and ϕ_n the multiset of mappings $(\phi_{kn})_{k=1}^{M^{\text{id}}}$. Furthermore, we denote by $\hat{W} = \psi_n(\phi_n^y(Y^n), \phi_n(\mathbf{X}^n))$ the output of the decoding mapping. The identification setting is depicted in Fig. 8.

The probability of identifying an incorrect index is given by

$$\begin{aligned} \Pr\{\hat{W} \neq w | W = w\} &= \sum_{\phi_n^y(y^n), \phi_n(\mathbf{x}^n)} P_{\phi_n^y(Y^n) \phi_n(\mathbf{X}^n) | W=w}(\phi_n^y(y^n), \phi_n(\mathbf{x}^n)) \\ &\quad \times \sum_{w' \neq w} P_{\hat{W} | \phi_n^y(y^n), \phi_n(\mathbf{x}^n)}(w'), \quad w \in [1 : M^{\text{id}}]. \end{aligned} \quad (76)$$

For the iid discrete scenario in which $\phi_{kn} = \phi_n$ for all $k \in [1 : M^{\text{id}}]$ holds and ψ_n is a deterministic mapping, this setting

Fig. 8: Identification system with user-dependent compression mappings ϕ_n and a compression mapping for the observation sequence ϕ_n^y . $P_{Y^n | X^n}$ is a well-defined probability kernel since $\mathcal{Y}_n$ and $\mathcal{X}_n$ are Polish spaces.

was studied in [33], where inner bounds and outer bounds on the achievable rate region were derived. Additionally, when $\phi_n^y = \text{id}$ the restricted setting reduces to the setting in [3].

We point out in the following one important difference between the identification system setting and our previous multi-user hypothesis testing problem. In our multi-user hypothesis testing, the number of users M grows at a fixed rate R , whereas in the identification setting we want to maximize the growing rate R^{id} of the number of supportable users M^{id} in the database. We therefore use two different notations to differentiate the distinction.

The following lemma relates the probability of correct identifying with the probability that the compressed sequence of the chosen user and the observation sequence are jointly typical. We assume that Eq. (52) also holds for the chosen mappings (ϕ_n^y, ϕ_n) .

Lemma 2. For any given η we have the following inequality

$$\begin{aligned} \Pr\{\hat{W} = W\} &\leq \max_{k \in [1 : M^{\text{id}}]} \Pr\{\iota_{\phi_n^y(\bar{Y}^n) \phi_{kn}(\bar{X}^n)}(\phi_n^y(\bar{Y}^n); \phi_{kn}(\bar{X}^n)) \\ &\quad > \log M^{\text{id}} - \eta\} + e^{-\eta}, \end{aligned} \quad (77)$$

where $(\bar{Y}^n, \bar{X}^n) \sim P_{\bar{Y}^n \bar{X}^n}$.

Proof: Define for each $k \in [1 : M^{\text{id}}]$ the following jointly typical sets

$$\begin{aligned} \mathcal{A}_k &= \{(\phi_n^y(y^n), \phi_n(\mathbf{x}^n)) \mid \\ &\quad \iota_{\phi_n^y(\bar{Y}^n) \phi_{kn}(\bar{X}^n)}(\phi_n^y(y^n); \phi_{kn}(x^n(k))) > \log M^{\text{id}} - \eta\}. \end{aligned}$$

For notation simplicity in the following we define $Q_k = P_{\phi_n^y(Y^n) \phi_{kn}(X^n(k))} \times \prod_{j \neq k} P_{\phi_{jn}(X^n(j))}$ for all $k \in [1 : M^{\text{id}}]$ and $Q = P_{\phi_n^y(\bar{Y}^n)} \times P_{\phi_n(\mathbf{x}^n)}$. Then for each $k \in [1 : M^{\text{id}}]$ we have

$$\begin{aligned} \Pr\{\hat{W} = k | W = k\} &= \sum_{\phi_n^y(y^n), \phi_n(\mathbf{x}^n)} Q_k(\phi_n^y(y^n), \phi_n(\mathbf{x}^n)) \\ &\quad \times \Pr\{\hat{W} = k | \phi_n^y(y^n), \phi_n(\mathbf{x}^n)\} \end{aligned}$$

$$\begin{aligned}
&= \sum_{(\phi_n^y(y^n), \phi_n(\mathbf{x}^n)) \in \mathcal{A}_k^c} Q_k(\phi_n^y(y^n), \phi_n(\mathbf{x}^n)) \\
&\quad \times \Pr(\hat{W} = k | \phi_n^y(y^n), \phi_n(\mathbf{x}^n)) \\
&+ \sum_{(\phi_n^y(y^n), \phi_n(\mathbf{x}^n)) \in \mathcal{A}_k} Q_k(\phi_n^y(y^n), \phi_n(\mathbf{x}^n)) \\
&\quad \times \Pr(\hat{W} = k | \phi_n^y(y^n), \phi_n(\mathbf{x}^n)). \quad (78)
\end{aligned}$$

The second term, which is the probability of correct identification as well as the compressed observation sequence and the compressed user's data sequence being jointly typical, can be upper bounded by

$$\begin{aligned}
Q_k(\mathcal{A}_k) &= \Pr\{\iota_{\phi_n^y(\bar{Y}^n)\phi_{kn}(\bar{X}^n)}(\phi_n^y(\bar{Y}^n); \phi_{kn}(\bar{X}^n)) \\
&> \log M^{\text{id}} - \eta\}. \quad (79)
\end{aligned}$$

Similarly the first term inside the bracket, which is the probability of correct identification as well as the compressed observation sequence and the compressed user's data sequence being not jointly typical, is upper bounded by

$$\begin{aligned}
&M^{\text{id}} e^{-\eta} \sum_{(\phi_n^y(y^n), \phi_n(\mathbf{x}^n)) \in \mathcal{A}_k^c} Q(\phi_n^y(y^n), \phi_n(\mathbf{x}^n)) \\
&\quad \times \Pr(\hat{W} = k | \phi_n^y(y^n), \phi_n(\mathbf{x}^n)) \\
&\leq M^{\text{id}} e^{-\eta} \sum_{\phi_n^y(y^n), \phi_n(\mathbf{x}^n)} Q(\phi_n^y(y^n), \phi_n(\mathbf{x}^n)) \\
&\quad \times \Pr(\hat{W} = k | \phi_n^y(y^n), \phi_n(\mathbf{x}^n)). \quad (80)
\end{aligned}$$

Therefore we have

$$\begin{aligned}
&\Pr\{\hat{W} = W\} \\
&\leq \frac{1}{M^{\text{id}}} \sum_{k=1}^{M^{\text{id}}} \Pr\{\iota_{\phi_n^y(\bar{Y}^n)\phi_{kn}(\bar{X}^n)}(\phi_n^y(\bar{Y}^n); \phi_{kn}(\bar{X}^n)) \\
&\quad > \log M^{\text{id}} - \eta\} \\
&\quad + e^{-\eta} \sum_k \sum_{\phi_n^y(y^n), \phi_n(\mathbf{x}^n)} Q(\phi_n^y(y^n), \phi_n(\mathbf{x}^n)) \\
&\quad \times \Pr(\hat{W} = k | \phi_n^y(y^n), \phi_n(\mathbf{x}^n)) \\
&\leq \max_{k \in [1:M^{\text{id}}]} \Pr\{\iota_{\phi_n^y(\bar{Y}^n)\phi_{kn}(\bar{X}^n)}(\phi_n^y(\bar{Y}^n); \phi_{kn}(\bar{X}^n)) \\
&\quad > \log M^{\text{id}} - \eta\} + e^{-\eta}. \quad (81)
\end{aligned}$$

The conclusion of the lemma follows. $\blacksquare$

Using Lemma 2 we can establish the code transformation between the identification problem and the single-user hypothesis testing problem as follows.

Theorem 9. Fix an arbitrary $\eta > 0$. Given an identification scheme $(\phi_n, \phi_n^y, \psi_n)$, we can construct a single-user HT scheme $(\bar{\phi}_n, \bar{\phi}_n^y, \bar{\psi}_n)$ such that for each n , $\bar{\phi}_n \in \phi_n$ holds and the corresponding error probabilities of type I and II are given by

$$\begin{aligned}
\bar{\alpha}_n &\leq \Pr\{\hat{W} \neq W\} + e^{-\eta}, \\
\bar{\beta}_n &\leq \frac{e^\eta}{M^{\text{id}}}. \quad (82)
\end{aligned}$$

Conversely, given a testing scheme $(\bar{\phi}_n, \bar{\phi}_n^y, \bar{\psi}_n)$ for the single-user HT problem, we can construct an identification scheme with M^{id} users $(\phi_n, \phi_n^y, \psi_n)$ such that

$$\Pr\{\hat{W} \neq w | W = w\} \leq \bar{\alpha}_n + M^{\text{id}} \bar{\beta}_n, \quad \forall w \in [1 : M^{\text{id}}]. \quad (83)$$

Proof: ID $\Rightarrow$ Single-user HT: From a given identification scheme with user-dependent compression mappings $(\phi_n, \phi_n^y, \psi_n)$ we directly use the mapping ϕ_n^y to compress the sequence y^n in the single-user HT, i.e. we set $\bar{\phi}_n^y = \phi_n^y$. For a given n we define events $\bar{\mathcal{A}}_{kn} = \{\iota_{\phi_n^y(\bar{Y}^n)\phi_{kn}(\bar{X}^n)}(\phi_n^y(\bar{Y}^n); \phi_{kn}(\bar{X}^n)) > \log M^{\text{id}} - \eta\}$ for all $k \in [1 : M^{\text{id}}]$. Similarly as in the proof of Theorem 5 we select the mapping $\bar{\phi}_n$ to compress the sequence x^n in the single-user HT to be $\bar{\phi}_n = \phi_{k^*n}$ where k^* is selected such that $\Pr\{\bar{\mathcal{A}}_{k^*n}\} = \max_{k \in [1:M^{\text{id}}]} \Pr\{\bar{\mathcal{A}}_{kn}\}$. We define an acceptance region for the single-user HT setup as

$$\begin{aligned}
\bar{\mathcal{A}}_n &= \{(\bar{\phi}_n(x^n), \bar{\phi}_n^y(y^n)) \mid \iota_{\bar{\phi}_n^y(\bar{Y}^n)\bar{\phi}_n(\bar{X}^n)}(\bar{\phi}_n^y(y^n); \bar{\phi}_n(x^n)) \\
&> \log M^{\text{id}} - \eta\}. \quad (84)
\end{aligned}$$

The probability of type I of error is then given by

$$\bar{\alpha}_n = P_{\bar{\phi}_n(\bar{X}^n)\bar{\phi}_n^y(\bar{Y}^n)}(\bar{\mathcal{A}}_n) \leq \Pr\{\hat{W} \neq W\} + e^{-\eta}, \quad (85)$$

where the inequality follows from Lemma 2. By the change of measure we also obtain

$$\bar{\beta}_n = P_{\bar{\phi}_n(\bar{X}^n)} \times P_{\bar{\phi}_n^y(\bar{Y}^n)}(\bar{\mathcal{A}}_n) \leq \frac{e^\eta}{M^{\text{id}}}. \quad (86)$$

ID $\Leftarrow$ Single-user HT: Given a testing scheme $(\bar{\phi}_n, \bar{\phi}_n^y, \bar{\psi}_n)$ of the single-user HT and a number of users M^{id} . We use the mapping $\bar{\phi}_n$ to compress each user's information and store it into a database and the mapping $\bar{\phi}_n^y$ to compress the observation sequence y^n in the identification setting, i.e., we set $\phi_{kn} = \bar{\phi}_n$, $\forall k \in [1 : M^{\text{id}}]$, and $\phi_n^y = \bar{\phi}_n^y$. We define the decoding rule as follows. We search for a unique $\hat{w}$ such that

$$\bar{\psi}_n(\bar{\phi}_n^y(y^n), \bar{\phi}_n(x^n(\hat{w}))) = 0. \quad (87)$$

If there exists none or there is more than one of such index, we output e . Define the following error events

$$\begin{aligned}
\mathcal{E}_1 &= \{\bar{\psi}_n(\bar{\phi}_n^y(Y^n), \bar{\phi}_n(X^n(W))) = 1\} \\
\mathcal{E}_2 &= \{\exists \hat{w} \neq W \mid \bar{\psi}_n(\bar{\phi}_n^y(Y^n), \bar{\phi}_n(X^n(\hat{w}))) = 0\}. \quad (88)
\end{aligned}$$

Then

$$\begin{aligned}
&\Pr\{\hat{W} \neq w | W = w\} \\
&\leq \Pr\{\mathcal{E}_1 | W = w\} + \Pr\{\mathcal{E}_2 | W = w\}. \quad (89)
\end{aligned}$$

It can be seen that the first term is equal to $\bar{\alpha}_n$ while the second term is upper bounded by $M^{\text{id}} \bar{\beta}_n$. $\blacksquare$

As a consequence of Theorem 9 we establish an intuitive however not obvious result on the under-capacity operation of the identification system. Assume that at the planning step, an identification scheme is built for M^{id} users. However, when the system is put in operation only M users are available where $M < M^{\text{id}}$, i.e., $M^{\text{id}} - M$ memory slots for storing compressed indices are empty. The identification mapping ψ_n needs to be redesigned, as it depends on the compressed user data sequences. The question is whether we also need to redesign the compression mappings. The following result states that we

need not to do so. This behavior is quite different from the channel coding problem. Therein the decoding mapping input does not contain any messages. Hence in the inactive message case it needs not to be redesigned.

Corollary 2. Let $(\phi_n, \phi_n^y, \psi_n)$ be an identification scheme such that for all $k \in [1 : M^{\text{id}}]$ we have $\phi_{kn} = \phi_n$ for some mapping ϕ_n . Furthermore, let $M < M^{\text{id}}$, $\eta > 0$ be given numbers and W' be a uniform random variable on $[1 : M]$. Then we can construct a mapping ψ'_n such that the identification scheme $((\phi'_{kn})_{k=1}^M, \phi_n^y, \psi'_n)$ with $\phi'_{kn} = \phi_n$ induces the following error probability

$$\begin{aligned} & \Pr\{\hat{W}' \neq w' | W' = w'\} \\ & \leq \Pr\{\hat{W} \neq W\} + e^{-\eta} + \frac{M}{M^{\text{id}}} e^\eta, \quad \forall w' \in [1 : M]. \end{aligned} \quad (90)$$

Proof: The result can be seen by applying Theorem 9 twice as follows:

ID scheme $(\phi_n, \phi_n^y, \psi_n) \mapsto$ Single-user HT scheme
 $(\phi_n, \phi_n^y, \bar{\psi}_n) \mapsto$ An ID scheme $((\phi'_{kn})_{k=1}^M, \phi_n^y, \psi'_n)$.

The following result is an immediate consequence of Theorem 9 and Theorem 5. It states that an identification database can be reused for the membership testing purpose when the identification system operates below capacity.

Corollary 3. Let (ϕ_n, ϕ_n^y) be compression mappings. Suppose that we use ϕ_n to enroll M user data sequences $(x^n(i))_{i=1}^M$ into a database. Let $M^{\text{id}} > M$ be an arbitrarily given number. Furthermore, let $(x^n(i))_{i=M+1}^{M^{\text{id}}}$ be virtual sequences generated from the same distribution $P_{\bar{X}^n}$ which are mapped into indices using the same mapping ϕ_n . Also let $\eta, \hat{E}$ be given positive numbers.

- 1) Let ψ_n^{id} be an identification mapping for the extended system involving M^{id} sequences with the corresponding average error probability $\Pr\{\hat{W} \neq W\}$. Then we can construct a decision mapping for the membership testing problem ψ_n^{t} with the given database of M compressed sequences such that the multi-user error probabilities are bounded by

$$\alpha_n \leq \Pr\{\hat{W} \neq W\} + e^{-\eta}, \quad \beta_n \leq M \frac{e^\eta}{M^{\text{id}}}. \quad (91)$$

- 2) Let ψ_n^{t} be a decision mapping for the multi-user hypothesis testing problem using the database of M compressed sequences with the corresponding error probabilities α_n, β_n . Then we can construct an identification mapping ψ_n^{id} on the extended system involving M^{id} users such that the error probability is given by

$$\begin{aligned} & \Pr\{\hat{W} \neq w | W = w\} \\ & \leq \alpha_n + e^{n\hat{E}} \beta_n + 2 \exp(-n\hat{E}) + 2M^{\text{id}} e^{-n\hat{E}} / M, \\ & \quad \forall w \in [1 : M^{\text{id}}]. \end{aligned} \quad (92)$$

In the following we present further consequences of the transformation between the single-user HT problem and the identification problem. For that purpose we need some additional definitions, which we state in the following.

Definition 4. For an arbitrary but fixed $\epsilon \in [0, 1)$, define $\mathcal{R}_{\text{ID}, \epsilon}$ to be the closure of all tuples $(R_c, R_y, R^{\text{id}})$ such that there exists a sequence of identification schemes $(\phi_n, \phi_n^y, \psi_n)$ which satisfies

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \max_{k \in [1 : M^{\text{id}}]} \frac{1}{n} \log |\phi_{kn}| \leq R_c, \\ & \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\phi_n^y| \leq R_y, \end{aligned} \quad (93a)$$

$$\limsup_{n \rightarrow \infty} \Pr\{\hat{W} \neq W\} \leq \epsilon, \quad \liminf_{n \rightarrow \infty} \frac{1}{n} \log M^{\text{id}} \geq R^{\text{id}}. \quad (93b)$$

Define the ϵ -identification capacity for a given compression rate pair (R_c, R_y) as

$$R_{\epsilon, \text{var}}^*(R_c, R_y) = \sup\{R^{\text{id}} \mid (R_c, R_y, R^{\text{id}}) \in \mathcal{R}_{\text{ID}, \epsilon}\}.$$

For an ϵ -achievable identification tuple $(R_c, R_y, R^{\text{id}})$ we say that a second-order rate $\hat{R}$ is achievable if the second condition in (93b) is replaced by

$$\liminf_{n \rightarrow \infty} \frac{1}{\sqrt{n}} (\log M^{\text{id}} - nR^{\text{id}}) \geq \hat{R}. \quad (94)$$

Similarly, we define the maximum second-order identification rate $\hat{R}_\epsilon^*(R_c, R_y, R^{\text{id}})$ for a given rate pair $(R_c, R_y, R^{\text{id}})$ as the supremum of achievable second-order rates $\hat{R}$.

Definition 5. Let $\mathcal{R}_{\text{HT}, \epsilon}$ be the closure of all tuples (R_c, R_y, E) such that there exists a sequence of single-user HT schemes $(\bar{\phi}_n, \bar{\phi}_n^y, \bar{\psi}_n)$ such that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\bar{\phi}_n| \leq R_c, \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\bar{\phi}_n^y| \leq R_y, \quad (95a)$$

$$\limsup_{n \rightarrow \infty} \bar{\alpha}_n \leq \epsilon, \quad \liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\bar{\beta}_n} \geq E. \quad (95b)$$

Similarly define the maximum ϵ -achievable error exponent for a given compression rate pair (R_c, R_y) as

$$E_\epsilon^*(R_c, R_y) = \sup\{E \mid (R_c, R_y, E) \in \mathcal{R}_{\text{HT}, \epsilon}\}.$$

For an ϵ -achievable hypothesis testing tuple (R_c, R_y, E) we say that a second-order rate $\hat{E}$ is achievable if the second condition in (95b) is replaced by

$$\liminf_{n \rightarrow \infty} \frac{1}{\sqrt{n}} (\log \frac{1}{\bar{\beta}_n} - nE) \geq \hat{E}. \quad (96)$$

We define the maximum second-order HT error exponent $\hat{E}_\epsilon^*(R_c, R_y, E)$ as the supremum of all achievable $\hat{E}$.

When the compression mapping for the observation sequence y^n is the identity mapping, $\phi_n^y = \text{id}$, we denote the ϵ -identification capacity by $R_{\epsilon, \text{var}}^*(R_c)$. We show in the following theorem a connection between the maximum ϵ -achievable identification rate in the ID setting and the maximum ϵ -achievable error exponent in the single-user HT setting. In the iid discrete case, a connection between the achievable region of the single-user HT and the one of the identification system with fixed compression mapping was given recently in [34] via the entropy characterization.

Theorem 10. For all $\epsilon \in [0, 1)$, and for all $(R_c, R_y) \in \mathbb{R}_+^2$, the following equality holds

$$E_\epsilon^*(R_c, R_y) = R_{\epsilon, \text{var}}^*(R_c, R_y).$$

In particular, when $P_{\bar{X}^n \bar{Y}^n} = P_{XY}^{\otimes n}$, and $\phi_n^y = \bar{\phi}_n^y = \text{id}$ hold, where P_{XY} is a joint distribution on a finite alphabet $\mathcal{X} \times \mathcal{Y}$, then we have $R_{\epsilon, \text{var}}^*(R_c) = R_{\max}(R_c)$.

Strong converse proofs for the identical compression mapping setting in [3] have been given in [35] and [36]. Theorem 10 shows that allowing user-dependent compression mappings does not change the optimal performance of the system in the strong converse sense.

Proof: Assume that both quantities are finite. Given $\gamma > 0$ there exists an identification scheme $(\phi_n, \phi_n^y, \psi_n)$ such that all conditions in (93) are satisfied for $(R_c + \gamma, R_y + \gamma, R_\epsilon^*(R_c, R_y) - \gamma)$. This implies that for all sufficiently large n we have $M^{\text{id}} \geq e^{n(R_\epsilon^*(R_c, R_y) - 2\gamma)}$. Then by the first part of Theorem 9 with $\eta = n\gamma$ the probabilities of error of the corresponding single-user HT scheme are bounded by

$$\limsup_{n \rightarrow \infty} \bar{\alpha}_n \leq \epsilon, \quad \liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\bar{\beta}_n} \geq R_\epsilon^*(R_c, R_y) - 3\gamma. \quad (97)$$

This implies that $E_\epsilon^*(R_c, R_y) \geq R_\epsilon^*(R_c, R_y)$, by taking $\gamma \rightarrow 0$.

Conversely, there exists a single-user testing scheme $(\bar{\phi}_n, \bar{\phi}_n^y, \bar{\psi}_n)$ such that all conditions in (95) are satisfied for $(R_c + \gamma, R_y + \gamma, E_\epsilon^*(R_c, R_y) - \gamma)$. This implies that for all sufficiently large n we have $\bar{\beta}_n \leq e^{-n(E_\epsilon^*(R_c, R_y) - 2\gamma)}$. By choosing $M^{\text{id}} = e^{n(E_\epsilon^*(R_c, R_y) - 3\gamma)}$ we obtain for $\eta = n\gamma$

$$\limsup_{n \rightarrow \infty} \Pr\{\hat{W} \neq W\} \leq \epsilon, \quad (98)$$

which implies that $(R_c + \gamma, R_y + \gamma, E_\epsilon^*(R_c, R_y) - 3\gamma) \in \mathcal{R}_{\text{ID}, \epsilon}$. Hence, we have $E_\epsilon^*(R_c, R_y) \leq R_\epsilon^*(R_c, R_y)$.

Next, if $E_\epsilon^*(R_c, R_y) = \infty$ for some pair $(R_c, R_y) \in \mathbb{R}_+^2$ and $\epsilon > 0$, then we can modify the proof as follows: Let $\{E_m\}_{m=1}^\infty$ be a sequence such that $E_m < \infty, \forall m$, and $E_m \rightarrow \infty$ as $m \rightarrow \infty$. Then we replace $E_\epsilon^*(R_c, R_y)$ with E_m in the last paragraph to get $(R_c, R_y, E_m) \in \mathcal{R}_{\text{ID}, \epsilon}$. This holds for any m , hence $R_\epsilon^*(R_c, R_y) = \infty$ as well. The case $R_\epsilon^*(R_c, R_y) = \infty$ can be handled similarly. ■

It can be seen that for a given rate pair (R_c, R_y) , $E_\epsilon^*(R_c, R_y)$ is finite, cf. Footnote 2. Hence both $E_\epsilon^*(R_c, R_y)$ and $R_{\epsilon, \text{var}}^*(R_c, R_y)$ are finite and equal to each other. Furthermore we have the following result whose proof can be inferred similarly as the ones of Theorems 4 and 7.

Corollary 4. For an arbitrarily given positive number $E_{\max}(R_c, R_y)$ the two following statements are equivalent.

- 1) For any single-user HT scheme $(\bar{\phi}_n, \bar{\phi}_n^y, \bar{\psi}_n)$ with $\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\bar{\phi}_n| \leq R_c$, $\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\bar{\phi}_n^y| \leq R_y$ and $\liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{1}{\bar{\beta}_n} \geq E$, if $E > E_{\max}(R_c, R_y)$, then $\bar{\alpha}_n \rightarrow 1$ exponentially fast at a positive convergence rate.
- 2) For any identification scheme with user-dependent compression mappings $(\phi_n, \phi_n^y, \psi_n)$ with $\liminf_{n \rightarrow \infty} \frac{1}{n} \log M^{\text{id}} \geq R$, $\limsup_{n \rightarrow \infty} \frac{1}{n} \log |\phi_n^y| \leq R_y$ and $\limsup_{n \rightarrow \infty} \max_{k \in [1: M^{\text{id}}]} \frac{1}{n} \log |\phi_{kn}| \leq R_c$, if $R > E_{\max}(R_c, R_y)$, then $\Pr\{\hat{W} \neq W\} \rightarrow 1$ exponentially fast at a positive convergence rate.

When $\phi_n^{\text{id}} = \bar{\phi}_n^{\text{id}} = \text{id}$, and $P_{\bar{X}^n \bar{Y}^n} = P_{XY}^{\otimes n}$, where P_{XY} is a joint distribution on a finite alphabet $\mathcal{X} \times \mathcal{Y}$, the exponentially

strong converse for the identification problem with identical compression mapping with $E_{\max}(R_c, \log |\mathcal{Y}|) = R_{\max}(R_c)$ is given in [35]. Corollary 4 shows that the exponentially strong converse holds for a more general setting with user-dependent compression mappings. In the next theorem we establish a connection between the maximum second order quantities in these two settings.

Theorem 11. For all $\epsilon \in [0, 1)$ and for all ϵ -achievable pair (R_c, R_y, T) , we have $\hat{R}_\epsilon^*(R_c, R_y, T) = \hat{E}_\epsilon^*(R_c, R_y, T)$.

Proof: To show this relation we apply Theorem 9 with $\eta = \sqrt{n}\gamma$ with an arbitrary $\gamma > 0$. Given a second-order identification rate $\hat{R}$ by the first part of Theorem 9 we obtain $\bar{\beta}_n \leq e^{-n(T + (\hat{R} - 2\gamma)/\sqrt{n})}$ since we have $M^{\text{id}} \geq e^{n(T + (\hat{R} - \gamma)/\sqrt{n})}$ for all sufficiently large n . This means that $\hat{R} - 2\gamma$ is a second-order achievable exponent for the hypothesis testing problem.

Given a second-order error exponent $\hat{E}$ by the second part of Theorem 9 we can choose $M^{\text{id}} = e^{n(T + (\hat{E} - 2\gamma)/\sqrt{n})}$ to achieve the desired performance since we have $\bar{\beta}_n \leq e^{-n(T + (\hat{E} - \gamma)/\sqrt{n})}$ for all sufficiently large n . This means that $\hat{E} - 2\gamma$ is a second-order achievable exponent for the identification problem.

Since $\hat{E}$, $\hat{R}$ and γ are arbitrary the theorem follows. ■

Remark 6. For the iid, discrete, and uncompressed scenario such that $V = \text{Var}[\iota(X; Y)] > 0$, Strassen's result [37] implies that $\hat{E}_\epsilon(\log |\mathcal{X}|, \log |\mathcal{Y}|, I(X; Y)) = \sqrt{V}\Phi^{-1}(\epsilon)$ where $\Phi(\cdot)$ is the cumulative distribution function of the standard Gaussian. Recently, Zhou et.al [35, Theorem 8] obtained the second-order identification rate $\hat{R}_\epsilon(\log |\mathcal{X}|, \log |\mathcal{Y}|, I(X; Y)) = \sqrt{V}\Phi^{-1}(\epsilon)$. Our result shows that the equality of these two results are a part of a more general relation.

Remark 7. Operational equivalences between source-coding problem and simple hypothesis testing have been established in the literature such as in the works [18, Section 4.7], [19, Section VIII]. Although these works share similar ideas with our work of using code transformations to relate similar settings to each other, there are few crucial differences as follows. In the source coding problem in [18], [19] only a single distribution is involved. Therefore, to relate the source-coding to the simple hypothesis testing problem an additional quantity is introduced as an alternative hypothesis, namely the counting measure. The arguments are however limited to discrete scenarios. In our case, the distributions for the single-user hypothesis testing problem are naturally induced from the identification problem. The applicable scope of our relation goes beyond the discrete settings. Additionally, the favorable properties of the settings in [18] and [19] allow quantities such as r -achievable dual coding rate to be deduced quickly from the r -achievable correct probability exponent. As the settings in our works are more complex than those, it is not obviously to obtain similar feats in our work.

IV. ACKNOWLEDGMENT

The authors would like to thank the associate editor Shun Watanabe and the anonymous reviewers for their helpful comments which help to improve the manuscript significantly.

REFERENCES

- [1] R. Ahlswede and I. Csiszár, "Hypothesis testing with communication constraints," *IEEE Trans. Inf. Theory*, vol. 32, no. 4, pp. 533–542, 1986.
- [2] F. M. J. Willems, T. Kalker, and J.-P. Linnartz, "On the capacity of a biometrical identification system," in *International Symposium on Information Theory*, 2003.
- [3] E. Tuncel, "Capacity/storage tradeoff in high-dimensional identification systems," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2097–2106, 2009.
- [4] E. Tuncel and D. Gündüz, "Identification and lossy reconstruction in noisy databases," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 822–831, 2014.
- [5] A. Wyner, "On source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 21, no. 3, pp. 294–300, 1975.
- [6] R. Ahlswede and J. Körner, "Source coding with side information and a converse for degraded broadcast channels," *IEEE Trans. Inf. Theory*, vol. 21, no. 6, pp. 629–637, 1975.
- [7] S. Voloshynovskiy, O. Koval, F. Beekhof, F. Farhadzadeh, and T. Holotyak, "Information-theoretical analysis of private content identification," in *IEEE Information Theory Workshop (ITW)*. IEEE, 2010, pp. 1–5.
- [8] N. A. Schmid, "Large deviations performance analysis for biometrics recognition," in *40th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2002.
- [9] P. Moulin, "Statistical modeling and analysis of content identification," in *Information Theory and Applications Workshop (ITA)*. IEEE, 2010, pp. 1–5.
- [10] V. Yachongka and H. Yagi, "Reliability function and strong converse of biomedical identification systems," in *International Symposium on Information Theory and Its Applications (ISITA)*. IEEE, 2016, pp. 547–551.
- [11] G. Dasarathy and S. C. Draper, "On reliability of content identification from databases based on noisy queries," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*. IEEE, 2011, pp. 1066–1070.
- [12] N. Merhav, "Reliability of universal decoding based on vector-quantized codewords," *IEEE Trans. Inf. Theory*, vol. 63, no. 5, pp. 2696–2709, 2017.
- [13] C. Tian and J. Chen, "Successive refinement for hypothesis testing and lossless one-helper problem," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4666–4681, 2008.
- [14] S. Sreekumar, D. Gündüz, and A. Cohen, "Distributed hypothesis testing under privacy constraints," in *2018 IEEE Information Theory Workshop (ITW)*. IEEE, 2018, pp. 1–5.
- [15] P. Escamilla, M. Wigger, and A. Zaidi, "Distributed hypothesis testing with concurrent detections," in *2018 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2018, pp. 166–170.
- [16] S. Salehkalaibar, M. Wigger, and L. Wang, "Hypothesis testing over the two-hop relay network," *IEEE Trans. Inf. Theory*, 2019.
- [17] D. Cao, L. Zhou, and V. Y. Tan, "Strong converse for hypothesis testing against independence over a two-hop network," *arXiv preprint arXiv:1808.05366*, 2018.
- [18] T. S. Han, *Information-Spectrum Methods in Information Theory*. Springer-Verlag Berlin Heidelberg, 2003.
- [19] H. Nagaoka and M. Hayashi, "An information-spectrum approach to classical and quantum hypothesis testing for simple hypotheses," *IEEE Transactions on Information Theory*, vol. 53, no. 2, pp. 534–549, 2007.
- [20] T. Han, "Hypothesis testing with multiterminal data compression," *IEEE Trans. Inf. Theory*, vol. 33, no. 6, pp. 759–772, 1987.
- [21] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
- [22] A. El Gamal and Y.-H. Kim, *Network information theory*. Cambridge university press, 2011.
- [23] Y. Polyanskiy and Y. Wu, "Lecture notes on information theory," *MIT (6.441), UIUC (ECE 563)*, 2017.
- [24] R. Ahlswede, P. Gács, and J. Körner, "Bounds on conditional probabilities with applications in multi-user communication," *Probability Theory and Related Fields*, vol. 34, no. 2, pp. 157–177, 1976.
- [25] S. Watanabe, "A converse bound on wyner-ahlswede-körner network via gray-wyner network," in *2017 IEEE Information Theory Workshop (ITW)*. IEEE, 2017, pp. 81–85.
- [26] J. Liu, R. van Handel, and S. Verdú, "Second-order converses via reverse hypercontractivity," *arXiv preprint arXiv:1812.10129*, 2018.
- [27] Y. Oohama, "Exponent function for one helper source coding problem at rates outside the rate region," in *2015 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2015, pp. 1575–1579. [Online]. Available: <https://arxiv.org/abs/1504.05891>
- [28] —, "Exponential strong converse for one helper source coding problem," *Entropy*, vol. 21, no. 6, p. 567, 2019.
- [29] H. Tyagi and S. Watanabe, "Strong converse using change of measure arguments," *IEEE Trans. Inf. Theory*, vol. 66, no. 2, pp. 689–703, 2019.
- [30] J. Liu, P. Cuff, and S. Verdú, " E_γ -resolvability," *IEEE Trans. Inf. Theory*, vol. 63, no. 5, pp. 2629–2658, 2017.
- [31] P. Cuff, "Distributed channel synthesis," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7071–7096, 2013.
- [32] H. Chernoff *et al.*, "The identification of an element of a large population in the presence of noise," *The Annals of Statistics*, vol. 8, no. 6, pp. 1179–1197, 1980.
- [33] M. B. Westover and J. A. O'Sullivan, "Achievable rates for pattern recognition," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 299–320, 2008.
- [34] G. Pichler, P. Piantanida, and G. Matz, "Distributed information-theoretic clustering," *arXiv preprint arXiv:1602.04605*, 2016.
- [35] L. Zhou, V. Y. Tan, L. Yu, and M. Motani, "Exponential strong converse for content identification with lossy recovery," *IEEE Trans. Inf. Theory*, vol. 64, no. 8, pp. 5879–5897, 2018.
- [36] M. T. Vu, T. J. Oechtering, M. Skoglund, and H. Boche, "Uncertainty in identification systems," *IEEE Trans. Inf. Theory*, vol. 67, no. 3, pp. 1400–1414, 2021.
- [37] V. Strassen, "Asymptotische Abschätzungen in Shannons Informationstheorie." *Trans. Third Prague Conference on Information Theory*, 1962.