Code Constructions and Bounds for Identification via Channels

Onur Günlü, *Member, IEEE,* Jörg Kliewer, *Senior Member, IEEE,*

Rafael F. Schaefer, *Senior Member, IEEE*, and

Vladimir Sidorenko, *Member, IEEE*

Abstract

Consider the identification (ID) via channels problem, where a receiver decides whether the transmitted identifier is its identifier, rather than decoding it. This model allows to transmit identifiers whose size scales doubly-exponentially in the blocklength, unlike common transmission codes with exponential scaling. Binary constant-weight codes (CWCs) suffice to achieve the ID capacity. Relating parameters

O. Günlü and R. F. Schaefer were supported by the German Federal Ministry of Education and Research (BMBF) within the national initiative for "Post Shannon Communication (NewCom)" under the Grant 16KIS1004. J. Kliewer was supported in part by U.S. National Science Foundation (NSF) under Grants 1815322 and 2107370. V. Sidorenko is on leave from the Institute for Information Transmission Problems, Russian Academy of Science. His work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No: 801434) and by the Institute for Communications Engineering at TU Munich. Parts of this work will be presented at the IEEE International Symposium on Information Theory 2021 in [1].

O. Günlü and R. F. Schaefer are with the Chair of Communications Engineering and Security, University of Siegen, 57076 Siegen, Germany (email:{onur.guenlue, rafael.schaefer}@uni-siegen.de).

J. Kliewer is with the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, 07102 Newark, NJ, USA (email: jkliewer@njit.edu)

V. Sidorenko is with the Institute for Communications Engineering, TU Munich, 80333 Munich, Germany (email: vladimir.sidorenko@tum.de)

of a binary CWC to the minimum distance of a code and using higher-order correlation moments, two upper bounds on binary CWC sizes are proposed. These bounds are also upper bounds on identifier sizes for ID codes constructed by using binary CWCs. We propose two constructions based on optical orthogonal codes (OOCs), which are used in optical multiple access schemes, have constant-weight codewords, and satisfy cyclic cross-correlation and auto-correlation constraints. These constructions are modified and concatenated with outer Reed-Solomon codes to propose new binary CWCs being optimal for ID. Improvements to the finite-parameter performance of both our and existing code constructions are shown by using outer codes with larger minimum distance vs. blocklength ratios. We illustrate ID regimes for which our ID code constructions perform significantly better than existing constructions. An extensive list of other modified OOCs that can be used as binary CWCs is provided.

Index Terms

Identification via channels, optical orthogonal codes, binary constant weight codes, hypothesis testing, constant composition codes.

I. Introduction

Shannon's paper on communication systems [2] establishes a fundamental asymptotic upper bound on the size of the message set, or *message size*, such that reliable transmission from a transmitter to a receiver through a memoryless noisy channel is possible by using an encoder-decoder pair. This point-to-point (P2P) channel model is one of the simplest communication models that are extended to numerous other channels; see, e.g., [3]. The common approach to provide reliable transmission is that the transmitter encodes a given message into a codeword with redundant information about the message such that the receiver can decode an observed noisy codeword reliably. Transmission codes with low decoding complexity and block error probability (BLER) are proposed for various communication and storage applications, including Reed-Solomon (RS) codes used for information storage and polar codes [4] for wireless communications with short blocklengths.

We consider a communication problem closely related to reliable communications via a P2P channel. Similar to the P2P channel model, a transmitter encodes an identifier, not known before encoding, into a codeword that is sent through a noisy channel such that a receiver observes a noisy codeword. Unlike the P2P channel model, the receiver is interested in the reliable result of the *binary hypothesis test* whether the transmitted identifier is the identifier of interest for him. Since the transmitted information of interest for each receiver is fixed, it is considered as an *identifier* for the corresponding receiver; therefore, this hypothesis testing problem is called the *identification via channels problem* [5].

The problem of authenticated communications through an adversarial multiple access channel is also closely related to the ID problem, since authentication of a message transmitted via unauthenticated communications through a nominal (i.e., no-attack state) channel is possible by using an ID code if authentication is possible, which does not affect the authenticated communication rate [6, Remark 4], [7]. This close relation is illustrated also in [8, Theorem 1] by showing the relation between the sizes of authentication and ID codes, which are equivalent in special cases. Thus, one can use various authentication problem results to obtain bounds on the parameters of the ID problem and vice versa.

One practical scenario for the identification (ID) problem is when there is a network of internet-of-things (IoT) devices, such as sensors, that are controlled by a mobile phone. Suppose we want to save energy to increase the battery life of these sensors. One way to achieve this is to insert a physical unclonable function (PUF) [9]–[11], which can be any digital circuit with unique outputs, into each sensor such that a uniformly distributed secret key is assigned to each device. Each secret key is an identifier for the corresponding sensor, which can be shared with the mobile phone when secure transmission is possible or by using public key cryptography. When the mobile phone intends to communicate with a particular sensor, this sensor's identifier and the content of the command to this sensor are encoded and broadcast through a noisy wireless channel. All sensors first apply a binary hypothesis test to decide whether they are the targeted

sensor. If this is not the case, they do not decode the command. Hence, all sensors with which the mobile phone does not communicate save energy by avoiding decoding. Similarly, see [12] for an application of the ID problem to digital watermarking.

The ID problem replaces the decoding operation in a P2P channel model with a binary hypothesis test; see, e.g., [13]–[15] for various discussions about this binary hypothesis test and its extensions. For a discrete memoryless channel (DMC) this problem is shown in [5] to allow the number of identifiers, or *identifier size*, to be doubly-exponential in the blocklength. This is actually achievable for any channel with a non-zero Shannon transmission capacity. This is also in contrast to the P2P channel problem for which the message size is exponential in the blocklength. Thus, asymptotically the number of identifiers that can be used in an ID problem is exponential in the message size of a P2P channel problem with the same blocklength. Furthermore, reliable ID is possible for a DMC with a maximum ID rate, defined below, being equal to its Shannon capacity [5], which is the maximum transmission rate at which reliable decoding is possible. The main difference between the encoders for the ID and the transmission problem, in the functional sense, is that randomization increases the performance of ID codes, whereas it suffices to use deterministic encoders for transmission. Randomized encoders allow to fully benefit from overlaps between sets of codewords assigned to different identifiers, which is the reason for the increase in the identifier size as compared to deterministic encoders [16]. Randomized encoders are required for several information-theoretic communication models, especially the ones with secrecy constraints, such as the wiretap channel (WTC) problem [17]. A local randomness source for a WTC transmitter is proposed in [18, Chapter 2] to be a digital PUF embodied in the WTC transmitter; see [11], [19] for other applications for which PUFs can be used as a randomness source. Such a uniformly-distributed PUF output can be used also for randomization by the ID transmitter. We assume that a codeword of a transmission code is selected by the ID transmitter for a given identifier uniformly at random over the pre-determined set of codewords assigned to the identifier. There exist uniformly-distributed randomized encoding algorithms with equally

sized codeword sets assigned to each identifier that achieve the *ID capacity* [5]. Therefore, we analyze binary constant-weight codes (CWCs), which are used to represent the equally sized codeword sets assigned to an identifier with symbol "1" and conversely codewords that cannot be chosen for a given identifier with symbol "0", respectively, as in [16], [20], [21].

An important family of binary CWCs is given by optical orthogonal codes (OOC), proposed in [22] as codes that have high synchronous auto-correlation, low synchronous and asynchronous cross-correlations, and low asynchronous auto-correlations. We remark that *synchronous* correlations consider sequences only with aligned symbols, whereas *asynchronous* correlations allow cyclic shifts in the sequences. OOCs are different from orthogonal (spreading) codes used in cellular asynchronous code division multiple access systems because OOCs consist of symbols "0" and "1", unlike orthogonal codes with symbols "1" and "-1". This property makes OOCs suitable for unipolar environments such as optical systems used for direction detection [22], where a symbol "1" represents a detected signal and symbol "0" no signal, respectively.

We modify OOCs to prove that modified OOCs are not only suitable but also optimal for the ID problem. First, binary CWCs that are shown in [20] to be optimal for the ID problem, have a parameter called *pairwise overlap* that corresponds to the maximum ratio of pairwise overlaps of "1" symbols between codewords. This parameter suggests that cross-correlation properties of such ID codes should not be affected in case symbols of different codewords do not overlap. This means that the cross-correlation value should increase when "1" symbols overlap, but it should not change when a "1" and the alternate symbol overlap. Therefore, OOCs are well suited for this purpose as they have binary (unmodulated) symbols of "0" and "1". Second, the set of binary CWCs is shown to contain OOCs and OOCs must satisfy extra cross-correlation constraints in addition to further auto-correlation constraints. Therefore, we modify OOC constructions to improve their ID performance and then illustrate the range of parameters for which optimal ID codes can be obtained by using modified OOCs. Unlike OOCs, our ID code constructions do not require correlation calculations at the receiver, which generally have a high

computational complexity. Furthermore, we propose a method to improve the finite-parameter performance of both our and existing ID code constructions by concatenating inner binary CWCs with suitable outer codes. We illustrate that our ID code constructions significantly outperform existing constructions at low ID rates, whereas at high ID rates existing constructions perform slightly better.

We provide two finite-parameter bounds on the ID code size. For binary CWCs, we show that the pairwise overlap and constant weight parameters uniquely define the minimum distance of the code. This fact allows to provide two different bounds on the ID code size. Furthermore, we use the fundamental result from [5, Section II-A], which states that to asymptotically achieve the ID capacity it suffices to design a binary CWC optimally for a noiseless channel and to concatenate it with a (Shannon) capacity-achieving transmission code, where the concatenation operation is different than the code concatenation operation common in the coding theory literature. Therefore, one can combine our two bounds for the proposed binary CWCs with finite length bounds for error correction codes to obtain bounds for ID code parameters for noisy channels.

A. Summary of Novel Contributions

A summary of the novel contributions in this work is stated below, where both the contribution in Point 6 as well as the proof of the contribution in Point 2 extend beyond the material presented in the submitted conference version of this work [1].

1) The code minimum distance of a binary CWC is uniquely determined by its codeword weight and the maximum number of overlaps of non-zero codeword symbols. This result allows to apply the Unrestricted Johnson Bound (UJB) [23], [24, Theorem 2.3.6], which depends on the code minimum distance, to binary CWCs. We then provide an upper bound on the size of a binary CWC with given blocklength, codeword weight, and maximum number of overlaps of non-zero codeword symbols.
2) A new alternative upper bound on the size of a binary CWC is provided by defining a

metric called ℓ-th correlation moment and then finding an upper and a lower bound on the metric. Similar to the UJB, a recursion formula is applied to improve the obtained upper bound on the size of a binary CWC.

3) We show that the given upper bounds on the size of a binary CWC are upper bounds also for the size of an ID code that consists of a concatenation of a binary CWC and a transmission code.
4) We propose two novel ID code constructions by modifying existing OOC constructions and by doubly concatenating them with two outer RS codes. Furthermore, we prove that these constructions are optimal for ID if a set of asymptotic conditions are satisfied. We also show that the performance of the proposed ID code constructions for finite parameters is improved by replacing outer RS codes with doubly-extended RS codes.
5) The proposed ID constructions are compared with the best existing ID constructions along with a tight asymptotic bound on the weighted sum of the ID rate and an ID error exponent. For low ID rates, the proposed constructions perform significantly better than the existing constructions, whereas for high ID rates the existing constructions perform slightly better. Furthermore, the gaps between the (ID rate, type-II error exponent) tuples achieved by the considered constructions and the tight asymptotic bound illustrate the performance loss due to finite code parameters.
6) We provide an extensive list of other OOC constructions that can be modified to construct ID codes by applying entirely similar steps as the ones used to construct the proposed ID codes. These other OOC constructions are expected to perform in general worse than the proposed ID codes mainly due to their larger error exponents. However, we emphasize that these other constructions might be useful, e.g., to decrease the hardware complexity.

B. Organization

This paper is organized as follows. In Section II, we describe the ID via channels problem, binary CWCs, and OOCs. In Section III, we provide two upper bounds on the binary CWC sizes that are used for ID. In Section IV, two OOC constructions are modified and concatenated with outer error correction codes to propose new binary CWCs that are optimal for ID and whose finite-parameter performance can be improved by using different outer codes. In Section V, the proposed ID code constructions are compared with existing constructions and with a tight asymptotic bound. Section VI concludes the paper.

C. Notation

Upper case letters represent random variables and lower case letters their realizations. A superscript denotes a string of variables, e.g., $X^n = X_1, X_2, \ldots, X_j, \ldots, X_n$, and a subscript j denotes the position of a variable in a string. A random variable X has a probability distribution P_X. Calligraphic letters such as $\mathcal{X}$ denote sets and set sizes are written as $|\mathcal{X}|$. $[1:M]$ denotes the set $\{1, 2, \ldots, M\}$ for an integer $M \geq 1$. $\mathbb{Z}^*$ denotes the set $\{0, 1, 2 \ldots\}$ of non-negative integers and $\mathbb{Z}^+$ denotes the set $\{1, 2, \ldots\}$ of positive integers. GF(p) denotes the Galois field of order p. $x \to a$ indicates that the parameter x tends to the value a. $\mathbb{1}\{\cdot\}$ denotes the indicator function. A vector with elements $(\alpha_0, \alpha_1, \ldots)$ is denoted as $\vec{\alpha}$ and its support $\mathrm{supp}(\vec{\alpha})$ is another vector $\vec{\beta} = (\beta_0, \beta_1, \ldots)$ such that $\beta_s = \mathbb{1}\{\alpha_s > 0\}$ for all $s \geq 0$.

II. Problem Formulation

Consider $N_{\text{ID}} \geq 1$ identifiers $i \in [1 : N_{\text{ID}}]$. This set of identifiers represents N_{ID} different receivers that want to test whether they are the receiver with which the transmitter intends to communicate. To communicate with the i-th receiver, the transmitter sends a sequence $X^{n_{\text{ID}}}$ whose noisy version $Y^{n_{\text{ID}}}$, associated with a DMC $P_{Y|X}$, is observed by each receiver. The task of the i^*-th receiver is to apply a hypothesis test for its received noisy sequence to decide

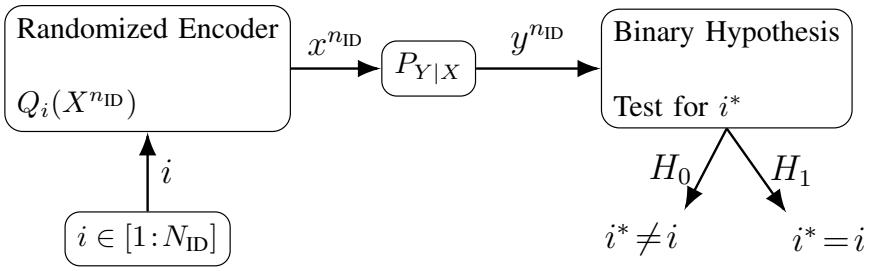

Fig. 1. Identification via channels problem. The i^*-th receiver applies the hypothesis test for its received sequence to determine whether it is the target of the intended communication.

whether the transmitted identifier is equal to the identifier $i^* \in [1 : N_{\text{ID}}]$ assigned to this receiver before transmission. The null hypothesis H_0 for each receiver is that the transmitted identifier is not the identifier assigned to it, and the alternative hypothesis H_1 is that the receiver is the one with which the transmitter aims to communicate. Fig. 1 illustrates the identifier encoding procedure at the transmitter that sends $X^{n_{\text{ID}}}$ through a channel $P_{Y|X}$, and the receiver observes $Y^{n_{\text{ID}}}$ for which the hypothesis test is applied.

There are two types of errors associated with the model shown in Fig. 1. *Type-I errors* occur when the receiver mistakenly decides that it is not the desired receiver. *Type-II errors* occur if the receiver mistakenly decides that it is the receiver for which the communication is intended. Consider a randomized encoding step that takes an identifier i as input and outputs a codeword $x^{n_{\text{ID}}} \in \mathcal{X}^{n_{\text{ID}}}$ according to a probability distribution $Q_i(X^{n_{\text{ID}}}) : i \to \mathcal{X}^{n_{\text{ID}}}$ for all $i \in [1 : N_{\text{ID}}]$. It is shown in [5] that in general a random encoder is necessary to achieve the ID capacity. Furthermore, type-I and type-II errors can be characterized by defining N_{ID} demapping regions $\mathcal{D}_i \subset \mathcal{Y}^{n_{\text{ID}}}$ for $i \in [1 : N_{\text{ID}}]$. The randomized encoding step allows to benefit from overlapping demapping regions, which is the main reason why the number of identifiers scales doubly-exponentially in the blocklength. This gain can be obtained as long as the two error probabilities can be made negligibly small [16]. Thus, we define the *identification via channels* problem as follows.

Definition 1. An $(n_{\text{ID}}, N_{\text{ID}}, \lambda_1, \lambda_2)$ ID code consists of N_{ID} encoding probability distributions $Q_i(X^{n_{\text{ID}}})$ and demapping regions $\mathcal{D}_i \subset \mathcal{Y}^{n_{\text{ID}}}$ such that, given a DMC $P_{Y|X}$, for all $i, i' \in [1 : N_{\text{ID}}]$ and $i \neq i'$ type-I and type-II error probabilities are upper bounded, respectively, as

$$1 - \sum_{y^{n_{\text{ID}}} \in \mathcal{D}_i} \sum_{x^{n_{\text{ID}}} \in \mathcal{X}^{n_{\text{ID}}}} Q_i(x^{n_{\text{ID}}}) P_{Y|X}^{n_{\text{ID}}}(y^{n_{\text{ID}}}|x^{n_{\text{ID}}}) \leq \lambda_1, \tag{1}$$

$$\sum_{y^{n_{\text{ID}}} \in \mathcal{D}_i} \sum_{x^{n_{\text{ID}}} \in \mathcal{X}^{n_{\text{ID}}}} Q_{i'}(x^{n_{\text{ID}}}) P_{Y|X}^{n_{\text{ID}}}(y^{n_{\text{ID}}}|x^{n_{\text{ID}}}) \leq \lambda_2. \tag{2}$$

$\Diamond$

Due to the doubly-exponential scaling of N_{ID} in the blocklength n_{ID}, the ID rate and ID capacity are defined as follows.

Definition 2. An ID rate R_{ID} is achievable if, given any $\lambda_1, \lambda_2, \epsilon > 0$, there exist some $n_{\text{ID}} \geq 1$, encoding probability distributions, and demapping regions satisfying Definition 1 and

$$\frac{1}{n_{\text{ID}}} \log(\log(N_{\text{ID}})) > R_{\text{ID}} - \epsilon. \tag{3}$$

The *ID capacity* C_{ID} is the supremum over all achievable ID rates R_{ID}. $\Diamond$

We next state the result that the ID capacity C_{ID} of a DMC $P_{Y|X}$ is equal to its Shannon capacity C_{Sh}.

Theorem 1 ([5])**.** *The ID capacity of a DMC $P_{Y|X}$ is*

$$C_{ID} = \max_{P_X} I(X;Y) = C_{Sh}. \tag{4}$$

We remark that if there is available common randomness shared between the transmitter and receiver, the ID capacity C_{ID} of a DMC increases by the entropy rate of the common randomness [12]. This provides an exponential increase in the number N_{ID} of identifiers with only a few bits of common randomness. Therefore, the performance of any ID code construction, including our constructions below, can be significantly improved if there is a source of common randomness available such as PUFs [25].

Achievability of the ID capacity C_{ID} is shown in [5, Section II-A] to be possible by using encoding probability distributions $Q_i(\cdot)$ that are uniformly distributed over equally sized support sets, which can be represented by binary CWCs [20]. Therefore, we next define the parameters of binary CWCs along with the conditions for them to be optimal ID codes.

Definition 3. An $(S_{\text{cw}}, N_{\text{cw}}, W_{\text{cw}}, K_{\text{cw}})$ binary CWC $\{x_1^{S_{\text{cw}}}, x_2^{S_{\text{cw}}}, \ldots, x_{N_{\text{cw}}}^{S_{\text{cw}}}\}$ consists of N_{cw} codewords of blocklength S_{cw} and Hamming weight W_{cw} with symbols $x_{j,s} \in \{0,1\}$ for $j = 1, 2, \ldots, N_{\text{cw}}$ and $s = 0, 1, \ldots, S_{\text{cw}} - 1$ such that the maximum number of overlaps of symbols $x_{j,s} = 1$ over all codeword pairs is K_{cw}, i.e., we have the cross-correlation

$$\gamma_{j,j'} \triangleq \sum_{s=0}^{S_{\text{cw}}-1} x_{j,s} x_{j',s} \leq K_{\text{cw}}, \quad \forall j, j' \in [1 : N_{\text{cw}}] \text{ s.t. } j \neq j'. \tag{5}$$

A set of binary CWCs is *optimal for ID* if we have [20]

$$\frac{\log(W_{\text{cw}})}{\log(S_{\text{cw}})} \to 1 \qquad \text{(weight factor)}, \tag{6}$$

$$\frac{\log(\log(N_{\text{cw}}))}{\log(S_{\text{cw}})} \to 1 \qquad \text{(second-order rate)}, \tag{7}$$

$$\frac{K_{\text{cw}}}{W_{\text{cw}}} \to 0 \qquad \text{(overlap fraction)}. \tag{8}$$

$\Diamond$

A closely related code family to binary ID CWCs is given by OOCs. We next define OOCs since the derivation of the bounds given below on the code size of binary ID CWCs follows similar steps as for OOCs. Note that our new ID code constructions also modify OOCs to improve their ID performance.

Definition 4. An $(S_{\text{ooc}}, N_{\text{ooc}}, W_{\text{ooc}}, \lambda_{\text{ooc,a}}, \lambda_{\text{ooc,c}})$ OOC $\{x_1^{S_{\text{ooc}}}, x_2^{S_{\text{ooc}}}, \ldots, x_{N_{\text{ooc}}}^{S_{\text{ooc}}}\}$ consists of N_{ooc} codewords of blocklength S_{ooc} and Hamming weight W_{ooc} with symbols $x_{j,s} \in \{0,1\}$ for $j = 1, 2, \ldots, N_{\text{ooc}}$ and $s = 0, 1, \ldots, S_{\text{ooc}} - 1$, such that for all $\tau_{\text{a}} \in [1 : (S_{\text{ooc}} - 1)]$, $\tau_{\text{c}} \in [0 : (S_{\text{ooc}} - 1)]$,

$j, j' \in [1 : N_{\text{ooc}}]$, and $j \neq j'$, we have

$$\sum_{s=0}^{S_{\text{ooc}}-1} x_{j,s} x_{j,(s+\tau_{\text{a}})} \leq \lambda_{\text{ooc},a} \qquad \text{(auto-correlation)}, \tag{9}$$

$$\sum_{s=0}^{S_{\text{ooc}}-1} x_{j,s} x_{j',(s+\tau_{\text{c}})} \leq \lambda_{\text{ooc},c} \qquad \text{(cross-correlation)} \tag{10}$$

where $(s + \tau_{\text{a}})$ and $(s + \tau_{\text{c}})$ additions are taken modulo S_{ooc}. $\Diamond$

We next give bounds on the size of ID codes that can be constructed by using binary CWCs with given parameters.

III. Upper Bounds On Binary CWC Sizes

We first consider the minimum distance of a binary CWC.

Lemma 1. *An $(S_{cw}, N_{cw}, W_{cw}, K_{cw})$ binary CWC has a minimum distance $d_{cw} = 2(W_{cw} - K_{cw})$.*

Proof: Since the $(S_{\text{cw}}, N_{\text{cw}}, W_{\text{cw}}, K_{\text{cw}})$ CWC is binary and since there are at most K_{cw} symbols of "1" overlapping between all codeword pairs, there are at least $(W_{\text{cw}} - K_{\text{cw}})$ symbols $x_{j,s} = 1$ of each codeword that are overlapping with $x_{j,s} = 0$ symbols of another codeword. Therefore, the number of symbols that are not the same is at least $2(W_{\text{cw}} - K_{\text{cw}})$ for each codeword pair of a $(S_{\text{cw}}, N_{\text{cw}}, W_{\text{cw}}, K_{\text{cw}})$ binary CWC. Furthermore, since there exist two binary CW codewords that have exactly K_{cw} overlapping symbols $x_{j,s} = 1$, the lemma follows. ■

Theorem 2. *Given a binary CWC with parameters S_{cw}, W_{cw}, and K_{cw}, we have*

$$N_{cw} \leq \left\lfloor \frac{S_{cw}}{W_{cw}} \left\lfloor \frac{(S_{cw}-1)}{(W_{cw}-1)} \right\rfloor \cdots \left\lfloor \frac{(S_{cw}-K_{cw})}{(W_{cw}-K_{cw})} \right\rfloor \cdots \right\rfloor \right\rfloor. \tag{11}$$

Proof: We first apply the Unrestricted Johnson Bound [23], [24, Theorem 2.3.6] to a CWC with parameters S_{cw}, W_{cw}, and d_{cw}, which can be proved by recursively puncturing codewords. Then, by using Lemma 1, the theorem follows. ■

The upper bound in Theorem 2 can in general be improved by treating codewords of a binary CWC as a set of sequences to bound their higher-order correlation moments. Such bounds are applied in [26] to OOCs, which compared to binary CWCs satisfy extra cyclic auto-correlation and cross-correlation constraints. Therefore, results in [26] cannot be directly used for binary CWCs. We next present in Theorem 3 another upper bound on the number N_{cw} of binary CW codewords given S_{cw}, W_{cw}, and K_{cw} by finding appropriate bounds on their higher-order correlation moments. See Appendix A for the complete proof of Theorem 3 and below for a proof sketch with discussions about why various results for OOCs cannot be applied to binary CWCs.

We first define functions and parameters that are used in Theorem 3. For a $d' \in [1 : K_{\text{cw}}]$, define $S'_{\text{cw}} = (S_{\text{cw}} - d')$, $W'_{\text{cw}} = (W_{\text{cw}} - d')$, and $K'_{\text{cw}} = (K_{\text{cw}} - d')$. For $\ell \in \mathbb{Z}^+$ and $u \in [1 : \ell]$, define

$$C_{\ell,u} = \sum_{k=0}^{u} (-1)^k \binom{u}{k} (u-k)^\ell. \tag{12}$$

Theorem 3. *Given a binary CWC with parameters S_{cw}, W_{cw}, and K_{cw}, we have the following upper bound on N_{cw} for any $\ell \in \mathbb{Z}^+$ and $d' \in [1 : K_{cw}]$ such that the innermost denominator is positive.*

$$N_{cw} \leq \left\lfloor \frac{S_{cw}}{W_{cw}} \left\lfloor \frac{(S_{cw}-1)}{(W_{cw}-1)} \left\lfloor \cdots \left\lfloor \frac{(S'_{cw}+1)}{(W'_{cw}+1)} \left\lfloor \frac{(W'_{cw})^\ell - (K'_{cw})^\ell}{\left(\frac{\sum_{u=1}^{\ell} C_{\ell,u} \binom{W'_{cw}}{u}^2}{\binom{S'_{cw}}{u}} \right) - (K'_{cw})^\ell} \right\rfloor \right\rfloor \cdots \right\rfloor \right\rfloor \right\rfloor. \tag{13}$$

Proof Sketch: Define the $\ell \geq 1$-th order correlation moment as

$$m_\ell = \frac{1}{N_{\text{cw}}(N_{\text{cw}}-1)} \left(\sum_{j=1}^{N_{\text{cw}}} \sum_{j'=1}^{N_{\text{cw}}} \gamma_{j,j'}^\ell - N_{\text{cw}} W_{\text{cw}}^\ell \right) \tag{14}$$

where $\gamma_{j,j'}$ is as defined in (5) such that $\gamma_{j,j'} = W_{\text{cw}}$ if $j = j'$. We provide a lower and an upper bound on the term $(N_{\text{cw}}-1)m_\ell$ by using the properties of binary CWCs so that a combination of

these bounds provides the bound in (13). We follow similar steps to the ones in [26, Appendix A] to obtain the lower bound for binary CWCs with two main differences. First, as compared to the correlation moment defined in [26, (A2)], our m_ℓ definition in (14) replaces $N_{\text{cw}}S_{\text{cw}}$ terms in the factors of the denominator given in [26, (A2)] by N_{cw} since binary CWCs do not impose any cyclic correlation constraints. Second, we remove the steps [26, (A16)] and [26, (A17)] that assume that the cyclic auto-correlation constraints in (9) are imposed, and we apply the Cauchy-Schwarz inequality for all cases as in [26, (A18)] to obtain the lower bound on $(N_{\text{cw}}-1)m_\ell$. The upper bound on $(N_{\text{cw}}-1)m_\ell$ used here is $(N_{\text{cw}}-1)K_{\text{cw}}^\ell$. Note that one cannot use similar steps as in [26, Appendix B] where upper bounds for OOCs are provided by using their cyclic correlation properties. Thus, by combining the obtained lower and upper bounds on $(N_{\text{cw}}-1)m_\ell$ and by applying a recursion formula for any $d' \in [1 : K_{\text{cw}}]$, which is applied also in the Unrestricted Johnson Bound and in [26, Theorem 4], the theorem follows. ■

Combining Lemma 1 and Theorem 3, the bound on N_{cw} in (13) can be written as a function of d_{cw}. This alternative formulation provides a lower bound on the minimum distance d_{cw} of binary CWCs with given parameters S_{cw}, N_{cw}, and W_{cw}, which can be useful to design ID binary CWCs.

We next prove that the upper bounds given in Theorems 2 and 3 are upper bounds also for the code size of ID binary CWCs.

Lemma 2. *If binary CWCs are used for ID, the upper bounds in (11) and (13) on N_{cw} are also upper bounds on the number N_{ID} of identifiers that can be reliably identified.*

Proof: $(S_{\text{cw}}, N_{\text{cw}}, W_{\text{cw}}, K_{\text{cw}})$ binary CWCs concatenated with a capacity C_{Sh} achieving transmission code are shown in [5, Section II-A] to be asymptotically optimal ID codes. To obtain an optimal $(n_{\text{ID}}, N_{\text{ID}}, \lambda_1, \lambda_2)$ ID code using this concatenation, the transmission code used for error correction should have a blocklength of n_{ID} and dimension of $\log(S_{\text{cw}})$; see [16, Section 4.1]. This scheme achieves $N_{\text{ID}} = N_{\text{cw}}$. This is because a given identifier $i \in [1 : N_{\text{ID}}]$ corresponds

to a CW codeword $x_i^{S_{\text{cw}}}$ such that the transmission codewords in the uniform encoding probability distributions $Q_i(x^{n_{\text{ID}}})$ are represented by symbols $x_{j,s}=1$ of the CW codeword $x_i^{S_{\text{cw}}}$, i.e., every $x_i^{S_{\text{cw}}}$ can choose W_{CW} transmission codewords. ■

Remark 1. The best upper bound obtained from Theorem 3, when applicable, provides generally a tighter result than the result of the bound given in Theorem 2, where only the first-order correlation properties are used. However, for various sets of parameters the denominator in (13) is not positive, so Theorem 3 cannot be applied, unlike Theorem 2.

IV. Modified OOC Constructions for ID

There are only a few constructive methods proposed for the ID via channels problem. In [16], [20], [21], [27] algebraic codes such as inner pulse position modulation (PPM) codes, which are binary CWCs with $W_{\text{cw}}=1$ and $K_{\text{cw}}=0$, concatenated with two outer codes are constructed to obtain binary CWCs optimal for ID. Similarly, in [28] ϵ-almost strongly universal hash functions are concatenated with an outer code. These constructions concatenate a set of inner binary CWCs with one or more outer codes such that the constraints in (6)-(8) are satisfied for the set of binary CWCs. The following lemma characterizes the parameters of binary CWCs obtained by such a concatenation.

Lemma 3 ([16])**.** *Consider the concatenation of an inner* $(S_{icw}, N_{icw}, W_{icw}, K_{icw})$ *binary CWC with an outer error correction code with blocklength* n_o*, code dimension* k_o*, minimum distance* d_o*, i.e., an* (n_o, k_o, d_o) *code. The resulting concatenated code is an* $(S_{icw}n_o,\ N_{icw}^{k_o},\ W_{icw}n_o,\ W_{icw}(n_o - d_o) + K_{icw}n_o)$ *binary CWC.*

Lemma 3 suggests that to achieve a small overlap fraction (8) the outer error correction code should have a large minimum distance vs. blocklength ratio $\frac{d_{\text{o}}}{n_{\text{o}}}$, whose maximum $\frac{(n_{\text{o}}-k_{\text{o}}+1)}{n_{\text{o}}}$ is obtained by maximum distance separable (MDS) codes. In [16], $[q_{\text{o}}-1, k_{\text{o}}]$ RS codes over GF(q_{o}), which are $(q_{\text{o}}-1, k_{\text{o}}, q_{\text{o}}-k_{\text{o}})$ error correction codes with $k_{\text{o}}<q_{\text{o}}-1$ and a prime power

$q_{\rm o}$, are used as outer codes. In [20], [21], [28], $[q_{\rm o}, k_{\rm o}]$ extended RS codes with parameters $(q_{\rm o}, k_{\rm o}, q_{\rm o} - k_{\rm o} + 1)$ are used as outer codes, which provide a larger minimum distance vs. blocklength ratio than RS codes because we have that

$$\frac{q_{\rm o} - k_{\rm o} + 1}{q_{\rm o}} > \frac{q_{\rm o} - k_{\rm o}}{q_{\rm o} - 1}. \tag{15}$$

This extension decreases the overlap fraction value of the concatenated CWC. To further improve the overlap fraction for the same field size $q_{\rm o}$, we propose to use $[q_{\rm o}+1, k_{\rm o}]$ *doubly-extended RS codes* that are also MDS with parameters $(q_{\rm o}+1, k_{\rm o}, q_{\rm o}-k_{\rm o}+2)$ as outer codes.

We next propose modified OOC constructions adapted to the ID via channels problem as new inner binary CWCs such that their concatenations with outer (doubly-extended) RS codes are optimal. A requirement to use Lemma 3 for outer (doubly-extended) RS codes is to set $q_{\rm o} = N_{\rm icw}$ such that each symbol of the outer code can be represented as a different codeword of the inner code [20]. Therefore, we propose modified OOC constructions with code sizes $N_{\rm icw}$ that are prime powers.

Construction 1: Prime sequences are proposed in [29], [30] as a $(p^2, p, p, p-1, 2)$ OOC, where p is a prime. A prime sequence is generated by multiplying in modulo-p all field elements of GF(p) with one of the field elements, where we map each field element to an integer in the range $[0 : p-1]$. For instance, prime sequences for $p = 5$ are $\{(00000), (01234), (02413), (03142), (04321)\}$. Each symbol is then mapped to an index in a binary sequence of length p such that at the corresponding index there is the symbol "1" and the other indices contain symbol "0". This symbol-to-binary-sequence mapping is called *one-hot encoding*. For instance, the prime sequence (01234) is mapped to the binary sequence $(10000\ 01000\ 00100\ 00010\ 00001)$. The number of pairwise overlaps of symbols $x_{j,s} = 1$ over the binary representations of prime sequences is $K_{\rm icw} = 1$ due to the first symbol being symbol "0", common in all prime sequences. We remove this "0" (i.e., for $p = 5$, we have sequences $\{(0000), (1234), (2413), (3142), (4321)\}$) to obtain binary representations of modified prime sequences that constitute a $(p^2-p, p, p-1, 0)$ binary

CWC, where p is prime.

If modified prime sequences are doubly concatenated with an outer $[p-1, k_{\mathrm{o}}]$ RS code over GF(p) and again with another outer $[p^{k_{\mathrm{o}}}-1, k_{\mathrm{oo}}]$ RS code over GF($p^{k_{\mathrm{o}}}$) (the latter is also called the *second outer RS code*), we obtain a binary CWC with parameters

$$S_{\mathrm{cw}} = p(p-1)^2(p^{k_{\mathrm{o}}}-1), \tag{16}$$

$$N_{\mathrm{cw}} = p^{k_{\mathrm{o}}k_{\mathrm{oo}}}, \tag{17}$$

$$W_{\mathrm{cw}} = (p-1)^2(p^{k_{\mathrm{o}}}-1), \tag{18}$$

$$K_{\mathrm{cw}} = (p-1)^2(k_{00}-1) + (p-1)(k_{\mathrm{o}}-1)(p^{k_{\mathrm{o}}}-1) \tag{19}$$

which follows from Lemma 3. It is straightforward to show that the binary CWCs constructed from modified prime sequences are optimal for ID if we have the following four conditions

$$\log(k_{\mathrm{oo}}) \to \infty, \tag{20}$$

$$\frac{\log(k_{\mathrm{oo}})}{k_{\mathrm{o}}} \to 1, \tag{21}$$

$$\frac{k_{\mathrm{o}}}{p} \to 0, \tag{22}$$

$$\frac{k_{\mathrm{oo}}}{p^{k_{\mathrm{o}}}} \to 0. \tag{23}$$

The last two conditions, i.e., (22) and (23), require the (first-order) code rates of outer codes to be asymptotically zero although the construction is optimal for ID, i.e., the second-order rate is optimal. Furthermore, the second outer RS code we use is more general than the second outer RS code used in [16], [20], [21], [28], where the code dimension is enforced to be $k_{\mathrm{oo}}=p^t$ for some $t \in [1 : k_{\mathrm{o}}-1]$. Thus, our optimality conditions for ID given in (20)-(23) are more general than the conditions in [20, Proposition 3].

If the outer RS codes are replaced with corresponding doubly-extended RS codes, then we obtain a binary CWC with parameters in (16)-(19) after replacing the $(p-1)^2$ terms with (p^2-1) and $(p^{k_{\mathrm{o}}}-1)$ terms with $(p^{k_{\mathrm{o}}}+1)$, respectively. The asymptotic optimality conditions for ID are

the same for constructions with two outer RS codes and doubly-extended RS codes, i.e., (20)-(23). However, using doubly-extended RS codes decreases the overlap fraction as compared to RS codes. Therefore, the type-II error probability λ_2 of the ID code, which can be obtained by concatenating the binary CWC with a capacity-achieving transmission code, also decreases by using outer doubly-extended RS codes. This is because λ_2 is shown in [27, Proposition 1] to be equal to the sum of overlap fraction of the binary CWC and the block error probability of the capacity-achieving transmission code. This result suggests that binary CWC constructions that have outer codes with large minimum distance vs. blocklength ratio $\frac{d_{\rm o}}{n_{\rm o}}$ should be used to decrease λ_2 of the ID code. Furthermore, doubly-extended RS codes can be obtained by adding two parity check symbols to RS codes, which has only small extra encoding complexity.

Construction 2: The following sequences are proposed in [26] as $(p^{2m}-1, p^m-2, p^m+1, 2, 2)$ OOCs, where p is a prime and $m \in \mathbb{Z}^+$. Let α be a primitive element of GF(p^{2m}) and consider p^m-2 sets with elements x satisfying

$$(x-1)^{p^m+1} = \alpha^{i(p^m+1)} \tag{24}$$

for $i \in [1 : p^m-2]$, where we then map each nonzero x to an integer equal to the exponent with respect to α, i.e., we calculate the integer $\log_\alpha(x)$, in modulo-$(p^{2m}-1)$. We obtain p^m-2 sets each containing p^m+1 integers in the range $[1 : p^{2m}-1]$ that correspond to the indices at which a binary CW codeword of blocklength $p^{2m}-1$ has the symbol "1". An example for $p=2$ and $m=3$ is given in [26, Table IV]. Since the field elements satisfying (24) are different for different i, this construction provides sequences that are $(p^{2m}-1, p^m-2, p^m+1, 0)$ binary CWCs, where p is prime and $m \in \mathbb{Z}^+$.

We now can concatenate these binary CWCs with outer codes such as RS codes to obtain optimal parameters for ID. However, unlike in Construction 1, $N_{\rm icw}=q_{\rm o}=p^m-2$ is not a prime power for all (p,m) pairs. For instance, $(p,m)=(2,\forall m\geq 3),(3,7),(3,8),(5,3),(11,2),(23,3)$ do not result in prime power values $N_{\rm icw}$, whereas various pairs such as $(p,m)=(2,2),(3,m\in$

$[2:6]), (3,9), (7,2), (13,2), (19,2)$ do. Therefore, if (doubly-extended) RS codes are used as outer codes, it is necessary to check the prime power condition since there may not exist a general condition to obtain prime powers of the form p^m-2 from a prime p and $m \in \mathbb{Z}^+$. One can alternatively decrease the size N_{icw} of this binary CWC to the maximum prime power p' such that $p' \leq p^m-2$.

If binary sequences obtained from the solution of (24) are doubly concatenated with an outer $[p^m\text{–}3, k_{\text{o}}]$ RS code over GF(p^m–2) and again with another outer $[(p^m-2)^{k_{\text{o}}}\text{–}1, k_{\text{oo}}]$ RS code over GF($(p^m-2)^{k_{\text{o}}}$), we obtain binary CWCs that are optimal for ID if the same four conditions given above for the optimality of Construction 1, i.e., (20)-(23), are satisfied here as well. Furthermore, the type-II error probability λ_2 of the ID codes constructed from these binary CWCs can be decreased by using outer codes with larger minimum distance vs. blocklength ratios $\frac{d_{\text{o}}}{n_{\text{o}}}$ than RS codes, as discussed for Construction 1. These constructions can be further modified to allow feedback [31], [32] and provide secrecy [33], [34].

V. ID Code Comparisons

ID codes that consist of $(S_{\text{cw}}, N_{\text{cw}}, W_{\text{cw}}, K_{\text{cw}})$ binary CWCs and a capacity C_{Sh} achieving transmission code are asymptotically optimal ID codes [5, Section II-A] with $N_{\text{ID}} = N_{\text{cw}}$, as discussed in the proof of Lemma 2. Thus, we consider noiseless channels $P_{Y|X}(y|x) = \mathbb{1}\{x = y\}$. For these channels, the capacity-achieving transmission code has a code rate of $C_{\text{Sh}} = 1$ symbol/channel-use, so we have $n_{\text{ID}} = \log(S_{\text{cw}})$. Furthermore, the type-I error probability is zero, i.e., $\lambda_1 = 0$, and the type-II error probability is upper bounded by the overlap fraction of the binary CWC, i.e., $\lambda_2 \leq \frac{K_{\text{cw}}}{W_{\text{cw}}}$. Define the type-I and type-II error exponents as $\text{E}_1 = -\frac{\log(\lambda_1)}{n_{\text{ID}}}$ and $\text{E}_2 = -\frac{\log(\lambda_2)}{n_{\text{ID}}}$, respectively.

Theorem 4 ([5], [20])**.** *If there exists an $(n_{ID}, N_{ID}, \lambda_1, \lambda_2)$ ID code that achieves the triple (R_{ID}, E_1, E_2) with $E_1 > 0$ for a DMC $P_{Y|X}$ with channel capacity C_{Sh} , then $R_{ID} + 2E_2 \leq C_{Sh}$. This bound is tight for noiseless channels.*

We compare Constructions 1 and 2 with the best existing ID constructions to illustrate the achieved $(R_{\text{ID}}, \text{E}_2)$ tuples for a noiseless channel. As benchmark schemes we consider the CWC construction in [20], where a PPM code is concatenated with two outer extended RS codes, and in [28], where ϵ-almost strongly universal hash functions are concatenated with an outer extended RS code, respectively. The choice of the finite field used for Constructions 1 and 2 affects the encoding complexity. We therefore choose the parameters

$$p_{\text{Constr.1}} = p_{\text{Constr.2}}^{m} - 2 \tag{25}$$

to have the same finite fields for both constructions, where $p_{\text{Constr.1}}$ is the parameter p for Construction 1 and $p_{\text{Constr.2}}$ is the parameter p for Construction 2, respectively. We assign $p_{\text{Constr.2}} = 5$ and $m = 2$ for Construction 2, and $p_{\text{Constr.1}} = 23$ as the parameter p for both Construction 1 and the constructions in [20], [28]. Fig. 2 depicts the $(R_{\text{ID}}, \text{E}_2)$ tuples achieved by these four constructions in addition to the tight upper bound given in Theorem 4; see [35] for its extensions to ID of multiple identifiers. We remark that all four constructions are optimal for ID for noiseless channels, i.e., they achieve the upper bound given in Theorem 4 asymptotically.

Fig. 2 illustrates that Constructions 1 and 2 achieve rate tuples that are close, and Construction 1 achieves slightly larger R_{ID} and E_2 values than Construction 2. Tuples achieved by Constructions 1 and 2 follow a similar pattern, whereas code constructions in [20] and [28] follow a pattern that is different from the patterns of Constructions 1 and 2. Furthermore, at low ID rates R_{ID} Constructions 1 and 2 achieve significantly larger type-II error exponents E_2 than being achieved by existing constructions, but at high ID rates the constructions in [20] and [28] can achieve slightly larger type-II error exponents. Thus, the choice of the ID code construction should depend on the required ID rate and the allowed encoding complexity.

A. ID Codes Constructed from Other OOCs

In addition to Constructions 1 and 2 given above, there are numerous other OOC constructions in the literature proposed for unipolar environments. One can modify these OOC constructions

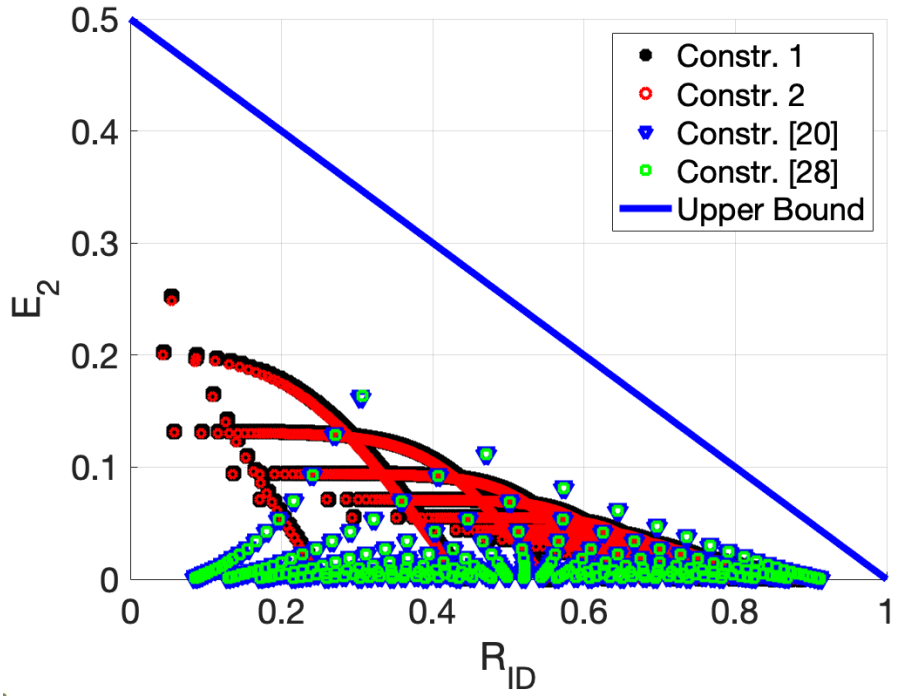

Fig. 2. Achieved (ID rate, type-II error exponent) tuples and the tight upper bound for a noiseless channel with $p_{\text{Constr.1}} = p^{m}_{\text{Constr.2}} - 2 = 23$.

to obtain further binary CWCs that can be used as an inner code in an ID code construction, as outlined in Section IV. In Table I, we provide a list of parameters of binary CWCs obtained from other modified OOCs. We also remark that the inner binary CWC size N_{icw} should be chosen as a prime power if it is concatenated with outer (doubly-extended) RS codes to obtain an ID code; see the discussions in Section IV for further details. Furthermore, Constructions 1 and 2 provide in general a smaller overlap fraction (8) than the other OOCs listed in Table I. This is due to the fact that the parameter K_{icw} of the inner binary CWC, which is concatenated with outer codes to obtain an ID code, is non-zero for these constructions compared to Constructions 1 and 2. However, different combinatorial constraints on the code parameters for each construction and the difference in the hardware complexity should also be considered to choose the modified OOC to be used for an ID via channels application.

TABLE I

A LIST OF OOCS AND THEIR PARAMETERS AS MODIFIED BINARY CWCS.

OOC Source	$\mathbf{S}_{\text{cw}}$	$\mathbf{N}_{\text{cw}}$	$\mathbf{W}_{\text{cw}}$	$\mathbf{K}_{\text{cw}}$	**Constraints**
Projective Geometry [22, Section IV-D]	$\frac{p^{d+1}-1}{p-1}$	$\left\lfloor \frac{S_{\text{cw}}-1}{(p+1)p} \right\rfloor$	$p+1$	1	p prime, $d \in \mathbb{Z}^+$
Combinatorial [22, Section IV-E]	m	$\left\lfloor \frac{m-1}{6} \right\rfloor$	3	1	$m \neq 2 \pmod 6$
Quadratic Congruence (QC) [36]	p^2	$p-1$	p	2	p prime
Extended QC [37]	$p(2p-1)$	$p-1$	p	2	p prime
Balanced Incomplete Block Designs (BIBDs) [38]	$W_{\text{cw}}(W_{\text{cw}}-1)r+1$	r	$2m$ or $2m+1$	1	$m, r \in \mathbb{Z}^+$, S_{cw} prime
Unequal Error Protection (UEP) 1 [39, Section V-A]	$W_{\text{cw}}^2 r/2+1$ or $(W_{\text{cw}}^2-1)r/2+1$	r	$2m$ or $2m+1$	1	$m, r \in \mathbb{Z}^+$, S_{cw} prime
UEP 2 [39, Section V-B]	$W_{\text{cw}}^2 r/2+1$ or $(W_{\text{cw}}^2-1)r/2+1$	r	$4m$ or $4m+1$	1	$m, r \in \mathbb{Z}^+$, S_{cw} prime

VI. CONCLUSION

We proposed two upper bounds on the size of a binary CWC by obtaining the minimum distance of the code and by finding bounds on its higher-order correlation moments. These bounds were shown to be upper bounds also on the corresponding ID code size. Two new asymptotically optimal ID code constructions were proposed by modifying OOCs, whose relation to the ID problem was discussed for the first time in the literature. Two RS codes were doubly concatenated with an inner binary CWC to obtain optimal ID codes, and we illustrated that the finite-parameter of all ID code constructions using such outer codes can be improved by

increasing the minimum distance vs. blocklength ratios of the outer codes. Furthermore, modified OOC constructions designed for ID were illustrated to perform significantly better than existing methods at low ID rates. We also provided an extensive list of other modified OOC constructions that can be used as binary CWCs. In future work, we will combine our two bounds for binary CWCs with finite length bounds on transmission codes used for error correction to provide bounds on the overall performance of ID codes.

ACKNOWLEDGMENT

O. Günlü thanks Rick Fritschek for his useful suggestions on Theorem 3.

APPENDIX A
PROOF OF THEOREM 3

Consider the $\ell \geq 1$-th order correlation moment defined in (14) as

$$m_\ell = \frac{1}{N_{\text{cw}}(N_{\text{cw}}-1)} \left(\sum_{j=1}^{N_{\text{cw}}} \sum_{j'=1}^{N_{\text{cw}}} \left(\sum_{s=0}^{S_{\text{cw}}-1} x_{j,s} x_{j',s} \right)^\ell - N_{\text{cw}} W_{\text{cw}}^\ell \right) \tag{26}$$

where we use the definition of $\gamma_{j,j'}$ given in (5). We equivalently have

$$\begin{aligned}
&m_\ell (N_{\text{cw}}-1) \\
&= \frac{1}{N_{\text{cw}}} \left(\sum_{j=1}^{N_{\text{cw}}} \sum_{j'=1}^{N_{\text{cw}}} \left[\left(\sum_{s_1=0}^{S_{\text{cw}}-1} x_{j,s_1} x_{j',s_1} \right) \left(\sum_{s_2=0}^{S_{\text{cw}}-1} x_{j,s_2} x_{j',s_2} \right) \cdots \left(\sum_{s_\ell=0}^{S_{\text{cw}}-1} x_{j,s_\ell} x_{j',s_\ell} \right) \right] - N_{\text{cw}} W_{\text{cw}}^\ell \right) \\
&= \frac{1}{N_{\text{cw}}} \left(\sum_{s_1=0}^{S_{\text{cw}}-1} \sum_{s_2=0}^{S_{\text{cw}}-1} \cdots \sum_{s_\ell=0}^{S_{\text{cw}}-1} \left[\left(\sum_{j=1}^{N_{\text{cw}}} \left(x_{j,s_1} x_{j,s_2} \ldots x_{j,s_\ell} \right) \right)^2 \right] - N_{\text{cw}} W_{\text{cw}}^\ell \right) \\
&= \frac{1}{N_{\text{cw}}} \left(\sum_{s_1=0}^{S_{\text{cw}}-1} \sum_{s_2=0}^{S_{\text{cw}}-1} \cdots \sum_{s_\ell=0}^{S_{\text{cw}}-1} \left[\left(\sum_{j=1}^{N_{\text{cw}}} \left(\prod_{s'=s_1,s_2,\ldots,s_\ell} x_{j,s'} \right) \right)^2 \right] - N_{\text{cw}} W_{\text{cw}}^\ell \right).
\end{aligned} \tag{27}$$

Define vectors $\vec{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_{S_{\text{cw}}-1})$ such that

$$\alpha_s \in \mathbb{Z}^*, \qquad \forall s \in [0 : S_{\text{cw}}-1] \qquad \text{and} \qquad \sum_{s=0}^{S_{\text{cw}}-1} \alpha_s = \ell \tag{28}$$

and multinomial coefficients as

$$\binom{\ell}{\alpha_0, \alpha_1, \ldots, \alpha_{S_{\text{cw}}-1}} = \frac{\ell!}{\alpha_0!\, \alpha_1!\, \ldots \alpha_{S_{\text{cw}}-1}!}. \tag{29}$$

Using these definitions in (27), we obtain

$$m_\ell(N_{\text{cw}}-1) = \frac{1}{N_{\text{cw}}}\left(\sum_{\vec{\alpha}} \left[\binom{\ell}{\alpha_0, \alpha_1, \ldots, \alpha_{S_{\text{cw}}-1}} \left(\sum_{j=1}^{N_{\text{cw}}} \Big(\prod_{s=0}^{S_{\text{cw}}-1} x_{j,s}^{\alpha_s} \Big) \right)^2 \right] - N_{\text{cw}} W_{\text{cw}}^\ell \right) \tag{30}$$

where the outermost summation is over all vectors $\vec{\alpha}$ that satisfy the properties given in (28) and where we define $0^0 = 1$. Denote next the support of a vector $\vec{\alpha}$ as $\vec{\beta} = \text{supp}(\vec{\alpha})$ and accordingly define index sets

$$\Omega_u = \left\{ \vec{\beta} \in \{0,1\}^{S_{\text{cw}}} \;\Big|\; \sum_{s=0}^{S_{\text{cw}}-1} \beta_s = u \right\}, \tag{31}$$

$$\Gamma_{\vec{\beta}} = \left\{ \vec{\alpha} \in \mathbb{Z}^{*S_{\text{cw}}} \;\Big|\; \sum_{s=0}^{S_{\text{cw}}-1} \alpha_s = \ell, \;\; \vec{\beta} = \text{supp}(\vec{\alpha}) \right\}, \tag{32}$$

$$\Gamma_u = \left\{ \vec{\alpha} \in \Gamma_{\vec{\beta}} \;\Big|\; \vec{\beta} \in \{0,1\}^{S_{\text{cw}}} \text{ fixed}, \;\; \sum_{s=0}^{S_{\text{cw}}-1} \beta_s = u \right\}. \tag{33}$$

Thus, we can represent (30) equivalently as

$$\begin{aligned}
&m_\ell(N_{\text{cw}}-1) \\
&= \frac{1}{N_{\text{cw}}}\left(\sum_{u=1}^{\ell} \sum_{\vec{\beta}\in\Omega_u} \sum_{\vec{\alpha}\in\Gamma_{\vec{\beta}}} \left[\binom{\ell}{\alpha_0, \alpha_1, \ldots, \alpha_{S_{\text{cw}}-1}} \left(\sum_{j=1}^{N_{\text{cw}}} \Big(\prod_{s=0}^{S_{\text{cw}}-1} x_{j,s}^{\alpha_s} \Big) \right)^2 \right] - N_{\text{cw}} W_{\text{cw}}^\ell \right) \\
&\overset{(a)}{=} \frac{1}{N_{\text{cw}}}\left(\sum_{u=1}^{\ell} \sum_{\vec{\beta}\in\Omega_u} \sum_{\vec{\alpha}\in\Gamma_{\vec{\beta}}} \left[\binom{\ell}{\alpha_0, \alpha_1, \ldots, \alpha_{S_{\text{cw}}-1}} \left(\sum_{j=1}^{N_{\text{cw}}} \Big(\prod_{s=0}^{S_{\text{cw}}-1} x_{j,s}^{\beta_s} \Big) \right)^2 \right] - N_{\text{cw}} W_{\text{cw}}^\ell \right) \\
&= \frac{1}{N_{\text{cw}}}\left(\sum_{u=1}^{\ell} \sum_{\vec{\beta}\in\Omega_u} \left[\left(\sum_{\vec{\alpha}\in\Gamma_{\vec{\beta}}} \binom{\ell}{\alpha_0, \alpha_1, \ldots, \alpha_{S_{\text{cw}}-1}} \right) \left(\sum_{j=1}^{N_{\text{cw}}} \Big(\prod_{s=0}^{S_{\text{cw}}-1} x_{j,s}^{\beta_s} \Big) \right)^2 \right] - N_{\text{cw}} W_{\text{cw}}^\ell \right) \\
&\overset{(b)}{=} \frac{1}{N_{\text{cw}}}\left(\sum_{u=1}^{\ell} \left[\left(\sum_{\vec{\alpha}\in\Gamma_u} \binom{\ell}{\alpha_0, \alpha_1, \ldots, \alpha_{S_{\text{cw}}-1}} \right) \left(\sum_{\vec{\beta}\in\Omega_u} \left(\sum_{j=1}^{N_{\text{cw}}} \Big(\prod_{s=0}^{S_{\text{cw}}-1} x_{j,s}^{\beta_s} \Big) \right)^2 \right) \right] - N_{\text{cw}} W_{\text{cw}}^\ell \right)
\end{aligned}$$

$$\overset{(c)}{=} \frac{1}{N_{\text{cw}}}\left(\sum_{u=1}^{\ell}\left[C_{\ell,u}\left(\sum_{\vec{\beta}\in\Omega_u}\left(\sum_{j=1}^{N_{\text{cw}}}\left(\prod_{s=0}^{S_{\text{cw}}-1}x_{j,s}^{\beta_s}\right)\right)^2\right)\right]-N_{\text{cw}}W_{\text{cw}}^{\ell}\right)$$

$$\overset{(d)}{\geq} \frac{1}{N_{\text{cw}}}\left(\sum_{u=1}^{\ell}\left[C_{\ell,u}\,\frac{\left(\sum_{\vec{\beta}\in\Omega_u}\left(\sum_{j=1}^{N_{\text{cw}}}\left(\prod_{s=0}^{S_{\text{cw}}-1}x_{j,s}^{\beta_s}\right)\right)\right)^2}{\sum_{\vec{\beta}\in\Omega_u}1^2}\right]-N_{\text{cw}}W_{\text{cw}}^{\ell}\right)$$

$$= \frac{1}{N_{\text{cw}}}\left(\sum_{u=1}^{\ell}\left[C_{\ell,u}\,\frac{N_{\text{cw}}^2\binom{W_{\text{cw}}}{u}^2}{\binom{S_{\text{cw}}}{u}}\right]-N_{\text{cw}}W_{\text{cw}}^{\ell}\right)$$

$$= N_{\text{cw}}\sum_{u=1}^{\ell}\left[C_{\ell,u}\,\frac{\binom{W_{\text{cw}}}{u}^2}{\binom{S_{\text{cw}}}{u}}\right]-W_{\text{cw}}^{\ell} \tag{34}$$

where (a) follows because $x_{j,s}\in\{0,1\}$ for all $j=1,2,\ldots,N_{\text{cw}}$ and $s=0,1,\ldots,S_{\text{cw}}-1$, (b) follows due to the symmetry in the function that defines multinomial coefficients, (c) follows from the definition of $C_{\ell,u}$ given in (12) and the following result from [40, pp. 29] [26, (A15)]

$$\sum_{\vec{\alpha}\in\Gamma_u}\binom{\ell}{\alpha_0,\alpha_1,\ldots,\alpha_{S_{\text{cw}-1}}} = S_u^{\ell}\,u! = C_{\ell,u} \tag{35}$$

where S_u^{ℓ} denotes a Stirling number of the second kind, i.e., the number of different ways to divide a set of size ℓ into u non-empty non-overlapping subsets such that their union is the whole set with size ℓ, and (d) follows from the Cauchy-Schwarz inequality. Thus, (34) is a lower bound on the ℓ-th order correlation moment m_ℓ. Furthermore, we can obtain a simple upper bound on m_ℓ as follows.

$$m_\ell(N_{\text{cw}}-1) = \frac{1}{N_{\text{cw}}}\left(\sum_{j=1}^{N_{\text{cw}}}\sum_{j'=1}^{N_{\text{cw}}}\gamma_{j,j'}^{\ell}-N_{\text{cw}}W_{\text{cw}}^{\ell}\right)$$

$$= \frac{1}{N_{\text{cw}}}\left(\sum_{j=1}^{N_{\text{cw}}}\sum_{\substack{j'=1\\ j'\neq j}}^{N_{\text{cw}}}\gamma_{j,j'}^{\ell}+\sum_{j=1}^{N_{\text{cw}}}\gamma_{j,j}^{\ell}-N_{\text{cw}}W_{\text{cw}}^{\ell}\right)$$

$$\overset{(a)}{\leq} \frac{1}{N_{\text{cw}}} \left(\sum_{j=1}^{N_{\text{cw}}} \sum_{\substack{j'=1 \\ j' \neq j}}^{N_{\text{cw}}} K_{\text{cw}}^{\ell} + N_{\text{cw}} W_{\text{cw}}^{\ell} - N_{\text{cw}} W_{\text{cw}}^{\ell} \right) = K_{\text{cw}}^{\ell} (N_{\text{cw}} - 1) \tag{36}$$

where (a) follows by (5). Thus, combining (34) and (36) it is straightforward to obtain the following upper bound

$$N_{\text{cw}} \leq \frac{{W_{\text{cw}}}^{\ell} - {K_{\text{cw}}}^{\ell}}{\left(\frac{\sum_{u=1}^{\ell} C_{\ell,u} \binom{W_{\text{cw}}}{u}^2}{\binom{S_{\text{cw}}}{u}} \right) - {K_{\text{cw}}}^{\ell}}. \tag{37}$$

Applying a recursion formula to (37), which is entirely similar to the recursion formulas applied in [26, Theorem 4] and in the Unrestricted Johnson Bound that is used in Theorem 2 above, we obtain (13).

REFERENCES

[1] O. Günlü, J. Kliewer, R. F. Schaefer, and V. Sidorenko, "Doubly-exponential identification via channels: Code constructions and bounds," in *IEEE Int. Symp. Inf. Theory*, Melbourne, Victoria, Australia, July 2021, accepted.

[2] C. E. Shannon, "A mathematical theory of communication," *Bell Sys. Tech. J.*, vol. 27, no. 3, pp. 379–423, July 1948.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ: John Wiley & Sons, 2012.

[4] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, July 2009.

[5] R. Ahlswede and G. Dueck, "Identification via channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 15–29, Jan. 1989.

[6] N. Sangwan, M. Bakshi, B. K. Dey, and V. M. Prabhakaran, "Multiple access channels with byzantine users," in *IEEE Inf. Theory Workshop*, Visby, Sweden, Aug. 2019, pp. 1–5.

[7] A. Beemer, E. Graves, J. Kliewer, O. Kosut, and P. Yu, "Authentication and partial message correction over adversarial multiple-access channels," in *IEEE Conf. Commun. Network Security*, Avignon, France, July 2020, pp. 1–6.

[8] L. A. Bassalygo and M. V. Burnashev, "Authentication, identification, and pairwise separated measures," *Problems Inf. Transmission*, vol. 32, no. 1, pp. 41–47, Mar. 1996.

[9] O. Günlü, T. Kernetzky, O. İşcan, V. Sidorenko, G. Kramer, and R. F. Schaefer, "Secure and reliable key agreement with physical unclonable functions," *Entropy*, vol. 20, no. 5, May 2018.

[10] L. Kusters, O. Günlü, and F. M. Willems, "Zero secrecy leakage for multiple enrollments of physical unclonable functions," in *Symp. Inf. Theory Sign. Process. Benelux*, Twente, The Netherlands, May-June 2018, pp. 119–127.

[11] O. Günlü and R. F. Schaefer, "An optimality summary: Secret key agreement with physical unclonable functions," *Entropy*, vol. 23, no. 1, Jan. 2021.

[12] Y. Steinberg and N. Merhav, "Identification in the presence of side information with application to watermarking," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1410–1422, May 2001.

[13] S. Watanabe, "Minimax converse for identification via channels," Jan. 2021, [Online]. Available: arxiv.org/abs/2011.14741.

[14] Y. Steinberg, "New converses in the theory of identification via channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 984–998, May 1998.

[15] A. Winter, *Identification Via Quantum Channels in the Presence of Prior Correlation and Feedback*. Springer Verlag: Berlin Heidelberg, Germany, 2006, pp. 486–504.

[16] K. Eswaran, "Identification via channels and constant-weight codes," 2005, [Online]. Available: people.eecs.berkeley.edu/∼ananth/229BSpr05/Reports/KrishEswaran.pdf.

[17] A. D. Wyner, "The wire-tap channel," *Bell Labs Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[18] O. Günlü, "Key agreement with physical unclonable functions and biometric identifiers," Ph.D. dissertation, TU Munich, Germany, Nov. 2018, published by Dr. Hut Verlag in Feb. 2019.

[19] M. Bloch, O. Günlü, A. Yener, F. Oggier, H. V. Poor, L. Sankar, and R. F. Schaefer, "An overview of information-theoretic security and privacy: Metrics, limits and applications," *IEEE J. Selected Areas Inf. Theory*, vol. 2, no. 1, pp. 5–22, Mar. 2021.

[20] S. Verdú and V. K. Wei, "Explicit construction of optimal constant-weight codes for identification via channels," *IEEE Trans. Inf. Theory*, vol. 39, no. 1, pp. 30–36, Jan. 1993.

[21] S. Derebeyoğlu, C. Deppe, and R. Ferrara, "Performance analysis of identification codes," *Entropy*, vol. 22, no. 10, Oct. 2020.

[22] F. R. K. Chung, J. A. Salehi, and V. K. Wei, "Optical orthogonal codes: Design, analysis, and applications," *IEEE Trans. Inf. Theory*, vol. 35, no. 3, pp. 595–604, May 1989.

[23] S. Johnson, "Upper bounds for constant weight error correcting codes," *Elsevier Discrete Math.*, vol. 3, no. 1–3, pp. 109–124, Jan. 1972.

[24] W. C. Huffman and V. Pless, *Fundamentals of Error-correcting Codes*. Cambridge, NY: Cambridge University Press, 2010.

[25] O. Günlü, "Multi-entity and multi-enrollment key agreement with correlated noise," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 1190–1202, 2021.

[26] H. Chung and P. V. Kumar, "Optical orthogonal codes - New bounds and an optimal construction," *IEEE Trans. Inf. Theory*, vol. 36, no. 4, pp. 866–873, July 1990.

[27] P. Moulin and R. Koetter, "A framework for the design of good watermark identification codes," in *Security, Steganography, Watermarking Multimedia Contents VIII*, vol. 6072, Jan. 2006, pp. 565 – 574.

[28] K. Kurosawa and T. Yoshida, "Strongly universal hashing and identification codes via channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2091–2095, Sep. 1999.

[29] A. A. Shaar and P. A. Davies, "Prime sequences: Quasi-optimal sequences for or channel code division multiplexing," *Electron. Lett.*, vol. 19, no. 21, pp. 888–890, Oct. 1983.

[30] P. Prucnal, M. Santoro, and T. Fan, "Spread spectrum fiber-optic local area network using optical processing," *IEEE J. Lightw. Technol.*, vol. 4, no. 5, pp. 547–554, May 1986.

[31] R. Ahlswede and G. Dueck, "Identification in the presence of feedback - A discovery of new capacity formulas," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 30–36, Jan. 1989.

[32] H. Boche, R. F. Schaefer, and H. V. Poor, "Identification capacity of channels with feedback: Discontinuity behavior, super-activation, and Turing computability," *IEEE Trans. Inf. Theory*, vol. 66, no. 10, pp. 6184–6199, Oct. 2020.

[33] R. Ahlswede and Z. Zhang, "New directions in the theory of identification via channels," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 1040–1050, July 1995.

[34] H. Boche and C. Deppe, "Secure identification for wiretap channels; robustness, super-additivity and continuity," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 7, pp. 1641–1655, July 2018.

[35] M. V. Burnashev and H. Yamamoto, "On optimal error exponents in noiseless channel identification," in *IEEE Int. Symp. Inf. Theory*, Aachen, Germany, June 2017, pp. 2737–2740.

[36] S. V. Marić, Z. I. Kostić, and E. L. Titlebaum, "A new family of optical code sequences for use in spread-spectrum fiber-optic local area networks," *IEEE Trans. Commun.*, vol. 41, no. 8, pp. 1217–1221, Aug. 1993.

[37] S. V. Marić, "New family of algebraically designed optical orthogonal codes for use in CDMA fibre-optic networks," *IET Electron. Lett.*, vol. 29, no. 6, pp. 538–539, Mar. 1993.

[38] R. M. Wilson, "Cyclotomy and difference families in elementary Abelian groups," *Elsevier J. Number Theory*, vol. 4, no. 1, pp. 17–47, Feb. 1972.

[39] G.-C. Yang and T. E. Fuja, "Optical orthogonal codes with unequal auto- and cross-correlation constraints," *IEEE Trans. Inf. Theory*, vol. 41, no. 1, pp. 96–106, Jan. 1995.

[40] G. Pólya, R. Taryan, and D. R. Woods, *Notes on Introductory Combinatorics*. Stanford, CA: Stanford Univ., Apr. 1979.