Secure Computation from Random Error Correcting Codes

Hao Chen[1], Ronald Cramer[2], Shafi Goldwasser[3], Robbert de Haan[4], and Vinod Vaikuntanathan[5]

[1] Department of Computing and Information Technology, School of Information Science and Engineering, Fudan University, Shanghai, China. EMAIL: `chenhao@fudan.edu.cn`.

[2] CWI, Amsterdam & Mathematical Institute, Leiden University, The Netherlands. URL: `http://www.cwi.nl/~cramer`.

[3] MIT, Cambridge, Massachusetts, USA & Weizmann Institute of Science, Rehovot, Israel. URL: `http://theory.lcs.mit.edu/~shafi`.

[4] CWI, Amsterdam, The Netherlands. URL: `http://www.cwi.nl/~haan`.

[5] MIT, Cambridge, Massachusetts, USA. URL: `http://www.mit.edu/~vinodv`.

Abstract. Secure computation consists of protocols for secure arithmetic: secret values are added and multiplied securely by networked processors. The striking feature of secure computation is that security is maintained even in the presence of an adversary who corrupts a quorum of the processors and who exercises full, malicious control over them. One of the fundamental primitives at the heart of secure computation is secret-sharing. Typically, the required secret-sharing techniques build on Shamir's scheme, which can be viewed as a cryptographic twist on the Reed-Solomon error correcting code. In this work we further the connections between secure computation and error correcting codes. We demonstrate that threshold secure computation in the secure channels model can be based on arbitrary codes. For a network of size n, we then show a reduction in communication for secure computation amounting to a multiplicative polylogarithmic factor (in n) compared to classical methods for small, e.g., constant size fields, while tolerating $t < (\frac{1}{2} - \epsilon)n$ players to be corrupted, where $\epsilon > 0$ can be arbitrarily small. For large networks this implies considerable savings in communication. Our results hold in the broadcast/negligible error model of Rabin and Ben-Or, and complement results from CRYPTO 2006 for the zero-error model of Ben-Or, Goldwasser and Wigderson (BGW). Our general theory can be extended so as to encompass those results from CRYPTO 2006 as well. We also present a new method for constructing high information rate ramp schemes based on arbitrary codes, and in particular we give a new construction based on algebraic geometry codes.

[1] Hao Chen's research has been supported by NSFC grants 10225106 and 90607005.

[2] Ronald Cramer's research has been partially supported by NWO VICI.

[4] Robbert de Haan's research has been partially funded by the Dutch BSIK/BRICKS project PDC1.

1 Introduction

Secure computation consists of protocols for secure arithmetic: secret values are added and multiplied securely by networked processors. The striking feature of secure computation is that security is maintained even in the presence of an adversary who corrupts a quorum of the processors and who exercises full, malicious control over them. A crowning achievement of cryptography in the late '80s was the following result (stated informally):

Any function that can be computed, can be computed securely.

This statement (appropriately formalized) was shown in the computational setting by Goldreich, Micali and Wigderson [16] and in the information-theoretic setting by Ben-Or, Goldwasser and Wigderson [2] and Chaum, Crépeau and Damgaard [5]. Our focus in this paper will be on the information-theoretic setting.

One of the fundamental primitives at the heart of information-theoretic secure computation is secret-sharing. Typically, the required secret-sharing techniques build on Shamir's scheme, which can be viewed as a cryptographic twist on the Reed-Solomon error correcting code. In this work we further the study on the connections between secure computation and error correcting codes. We demonstrate that threshold secure computation in the secure channels model can be based on arbitrary codes, in two steps.

First we identify sufficient, specialized conditions on a secret sharing scheme in order that it can serve as an essentially seamless replacement of Shamir's scheme in the context of secure computation. Second, we show how arbitrary error correcting codes give rise to such dedicated secret sharing schemes, and we prove various bounds on the relevant achievable parameters. We also analyze high information rate ramp schemes based on arbitrary codes, and in particular we give a new construction based on algebraic geometry codes.

A t-threshold secret-sharing scheme among n players typically has the following complementary pair of guarantees: (1) *Privacy:* The shares of any set of at most t players reveal no information about the secret, and (2) *Reconstruction:* The shares of $t+1$ players, together, reveal the entire secret. Linear threshold secret sharing schemes are known to be equivalent to maximum-distance-separable (MDS) codes. By known lower bounds on MDS codes (or equivalently, on matroids), the smallest possible field K on which the shares can lie is of size at least $\max\{n-t, t+2\} \geq \frac{n+2}{2}$ [6]. We show that this obstacle can be circumvented by bounding corruption tolerance an arbitrary constant fraction of n away from its maximal value $\lfloor \frac{n-1}{2} \rfloor$.

In turn, we use this result to improve the existing results on information-theoretic secure computation. The existing approaches, which use variants of Shamir's threshold secret-sharing scheme, incur a communication overhead as the size of the working field is larger than n due to Shamir's scheme. This can

[6] In fact, the so-called Main Conjecture on MDS codes implies that $|K|$ is at least n minus a constant.

amount to a multiplicative factor of a (large) power of $\log n$ bits. Our results alleviate this and allow, for instance, constant size fields K as opposed to linear size, while corruption tolerance t is at most an (arbitrary) constant fraction of n away from optimal. Such a (small) loss is unavoidable over sub-linear size fields due to (the above-mentioned) impossibility results from combinatorics.

Concretely, by using Gilbert-Varshamov type of arguments, we show that for each ϵ there is a constant size field K and an infinite family of quasi-threshold (i.e., ramp) parameters (t_i, n_i) such that for each of them there is an ideal (or information rate 1/2) linear secret sharing scheme over K that has multiplication, t_i-privacy and $(n_i - t_i)$-reconstruction and $(\frac{1}{2} - \epsilon)n_i \leq t_i < \frac{1}{2}n_i$. Other interesting examples include schemes over $\mathbb{F}_2$ where corruption tolerance t is about $\frac{n}{10}$, or in fact, $t \approx \frac{n}{5}$ for $n \leq 100$.

Trading corruption tolerance for small fields was first used in [6] where a class of algebraic geometric secret sharing schemes was introduced that are ideal, linear, offer t-privacy and $(n-2t)$-reconstruction and satisfy the strong multiplication property rather than only the multiplication property. It was shown there how this enables low-communication threshold multi-party computation over small (e.g. constant) size fields in the zero-error/perfect security/active adversary model of Ben-Or, Goldwasser and Wigderson (BGW) [2]. This result owes to the special multi-linear algebraic structure induced by rational function evaluation (for the strong multiplication property, which also implies efficient error correction algorithms), the existence of families of algebraic curves with many rational points (to enable a small field), and reductions from secure computation to these dedicated secret sharing schemes. Of course, the techniques from [6] can be adapted to obtain the quasi-threshold schemes of the type we consider in this work (at least when $|K|$ is a square); their properties are different but similar enough to facilitate easy adaptation.

However, our first point is that quasi-threshold schemes of the type we consider here are much easier to design. In fact, they can be constructed from arbitrary (or even randomly chosen) error correcting codes. Our second point is that, although these quasi-threshold schemes cannot be used as the basis for BGW type of secure computation (as opposed to the schemes from [6]), they can serve as an essentially seamless replacement of Shamir's scheme in known secure computation protocols in the broadcast model of Rabin and Ben-Or (such as [26, 9, 12]) supplemented with preprocessing. In this model a broadcast primitive is given and small, non-zero errors are tolerated, but corruption tolerance is greater, i.e., up to $\frac{1}{2}n$ instead of $\frac{1}{3}n$ as in the BGW model. An important advantage of the use of our quasi-threshold schemes here is that they can lead to much more communication-efficient protocols. More concretely, when operating in Beaver's preprocessing model [1], we can obtain a reduction in communication amounting to a multiplicative polylogarithmic factor (in n), while tolerating a number of corrupted players that is arbitrarily close to the optimal value of $n/2$. Note that this may offer a considerable gain in case of very large networks. For similar results in the zero-error BGW model, see [6].

We also consider high information rate ramp schemes based on arbitrary codes. These are schemes where the secret is a vector of field elements, but shares consist of a single field element (or at least a shorter vector than the secret). This of course is impossible in perfect secret sharing schemes, which necessarily have shares of size at least the size of the secret. In ramp schemes one has t-privacy and r-reconstruction, and one does care if there are sets of size in between these bounds whose joint shares reveal partial information about the secret. The earliest example of such a scheme we are aware of is the one by Blakley and Meadows [4] (see also [19, 24] in the references therein), which is a variation on Shamir's scheme. We give a full treatment of linear ramp schemes from arbitrary error correcting codes, and show various bounds. As an application we give a new scheme based on algebraic geometry that improves the high information rate scheme given at the end of [6].

1.1 Organization of the paper

This paper is organized as follows. In Section 3 we study linear quasi-threshold secret sharing schemes with multiplication and show how these can be constructed from codes. Additionally, we prove several bounds on the achievable parameters. We also argue there how these schemes can essentially seamlessly replace Shamir's scheme in secure computation in the Rabin/Ben-Or model with preprocessing and indicate what savings can be achieved due to our results.

In Section 4.1 and Section 4.2, we describe a general approach for constructing high information rate ramp schemes from linear codes. Finally, in Section 4.3, we present a new high information rate ramp scheme based on algebraic geometry that improves the one presented in [6] and demonstrate that we can obtain high information rate ramp schemes from randomly generated codes and can predict bounds on their parameters with high probability.

2 Preliminaries and Definitions

2.1 Basic Definitions from Coding Theory

We establish notational conventions that we will use throughout this paper. Let K be a finite field.

Definition 1 *The* Hamming weight $w_H(c)$ *of a vector* $c \in K^n$ *is the number of non-zero positions in* c. *For a subspace* $C \subset K^n$, *the* minimum distance $d_{min}(C)$ *is defined as* $\min\{w_H(c) \mid c \in C\backslash\{0\}\}$.

An $[n, k, d]$-*code* C over K is defined to be a k-dimensional subspace of K^n with $d_{\min}(C) = d$.

Definition 2 *The* dual code $C^\perp$ *for a code* C *consists of all vectors* $c^* \in K^n$ *such that* $\langle c^*, c\rangle = 0$ *for all* $c \in C$, *where* $\langle \cdot, \cdot \rangle$ *denotes the standard inner product. Whenever* d *is used to denote the minimum distance of* C, $d^\perp$ *is used to denote the minimum distance of* $C^\perp$.

2.2 Threshold and Ramp Secret Sharing Schemes

In what follows, the reader is assumed to be familiar with linear secret sharing schemes (For details, see [10, 11, 6]). However, we give a brief survey of the most relevant properties below.

A secret-sharing scheme with t-privacy and r-reconstruction over a field K is an algorithm that, on input a secret $s_0 \in K^{d_0}$, outputs a vector $(s_1, \ldots, s_n)$ of shares, where $s_i \in K^{d_i}$ for certain $d_i > 0$, such that for any $A \subset \{1, 2, \ldots, n\}$ the following properties hold:

1. If $|A| \geq r$, then the shares $(s_i)_{i \in A}$ jointly determine the value s_0.
2. If $|A| \leq t$, then the shares $(s_i)_{i \in A}$ jointly give no information about s_0.

Such a scheme is called a *t-threshold secret-sharing scheme* when $r = t + 1$. In general (that is, when this is not the case), the scheme is called a *ramp (quasi-threshold) scheme with t-privacy and r-reconstruction.*

The sets A for which the shares allow for reconstruction are referred to as the *accepted* sets, whereas the sets for which the shares give no information are called the *rejected* sets. The *information rate* of a secret sharing scheme is $d_0/\max\{d_1, \ldots, d_n\}$. A secret sharing scheme with information rate 1, which is maximal for threshold secret sharing schemes, is said to be *ideal.*

A secret sharing scheme is said to be *linear* if for any two secrets s and s' and respective share vectors $(s_1, s_2, \ldots, s_n)$ and $(s'_1, s'_2, \ldots, s'_n)$, the vectors $(s_1 + s'_1, s_2 + s'_2, \ldots, s_n + s'_n)$ and $(\lambda s_1, \lambda s_2, \ldots, \lambda s_n)$ are valid share vectors for the secrets $s + s'$ and λs respectively. It is said to have the *multiplication property* if given any two full share vectors $(s_1, s_2, \ldots, s_n)$ and $(s'_1, s'_2, \ldots, s'_n)$ for secrets s and s', there is a vector r such that $\langle r, (s_1 s'_1, s_2 s'_2, \ldots, s_n s'_n)\rangle = ss'$, where $\langle \cdot, \cdot \rangle$ denotes the standard inner product. It has *strong multiplication* with respect to a t-adversary structure if the multiplication property holds with respect to any combination of $n - t$ shares. The latter property allows for reconstruction of the secret after a pooling of all shares, even when the shares for up to t indices are replaced by random values.

3 Linear Ramp Schemes with Multiplication from Codes

3.1 Massey's Secret Sharing From Codes

Massey [22, 23] gave the following construction of a secret sharing scheme from an error correcting code. Let C be an $[n + 1, k, d]$-code over a finite field K. We use coordinates $(c_0, c_1, \ldots, c_n)$ for codewords. The dual code $C^\perp$ is then an $[n + 1, n + 1 - k, d^\perp]$-code. We tacitly assume in this section that C is non-degenerate, i.e., that the minimum distances of both C and $C^\perp$ are greater than 1.

Let $s \in K$ be a secret value. Select a codeword $c = (c_0, c_1, \ldots, c_n) \in C$ uniformly at random such that $c_0 = s$, and define the share-vector as $(c_1, \ldots, c_n)$. Let LSSS(C) denote this linear secret sharing scheme. The access structure $\Gamma(C)$, i.e., the collection of accepted sets, is as follows. For a vector x, define $\sup(x) =$

$\{i : x_i \neq 0\}$. Consider the set V_0 of all $c^* \in C^\perp$ such that $c_0^* = 1$. Then $\Gamma(C) = \{\sup(c^*) \setminus \{0\} : c^* \in V_0\}$.

We now extend this idea in several ways in order to obtain the claimed quasi-threshold schemes, and we prove bounds on their existence.

3.2 Extensions of Massey's Idea

We first report the following consequence (which appears to be part of folklore) about the ramp parameters of this scheme and include a proof.

THEOREM 1 *Let C be an $[n+1, k, d]$-code over a finite field K. Then LSSS(C) offers linearity, $(d^\perp - 2)$-privacy and $(n - d + 2)$-reconstruction.*

PROOF. Linearity is clear; the sum of two code-words is a share-vector for the sum of the secrets, and likewise for scalar multiplication. First, we argue that $\Gamma(C) = (\Gamma(C^\perp))^*$, i.e., the access structure of LSSS(C) is the dual of the access structure of LSSS$(C^\perp)$, and vice versa.[7] Indeed, $A \in \Gamma(C)$ if and only if there is $c^* \in C^\perp$ with $c_0^* = 1$ and $c_i = 0$ for all $i \in \{1, \ldots, n\} \setminus A$ $(:= \overline{A})$. The latter is a share vector with secret equal to 1 in LSSS$(C^\perp)$, with shares equal to 0 for $\overline{A}$. The existence of such a share vector is equivalent to $\overline{A} \notin \Gamma(C^\perp)$. Now, from the characterization of $\Gamma(C)$ it is immediate that LSSS(C) rejects all sets of size $d^\perp - 2$. Since LSSS$(C^\perp)$ rejects all sets of size $d - 2$ and since $\Gamma(C) = (\Gamma(C^\perp))^*$, it must be that LSSS$(C)$ accepts all sets of size $n - d + 2$. $\triangle$

The exact privacy threshold $t_{\max}$ is equal to $-2 + \min\{w_H(c^*) : c^* \in C^\perp : c_0^* = 1\}$, i.e., this is the largest cardinality such that the joint shares of any set of this cardinality give no information on the secret. The exact reconstruction threshold $r_{\min}$ is equal to $n + 2 - \min\{w_H(c) : c \in C : c_0 = 1\}$.

For $A \subset \{1, \ldots, n\}$, let $\phi_A(C)$ denote the code restricted to the coordinates from the set $i \in A \cup \{0\}$, i.e., consisting of all codewords of C stripped of the coordinates not in $A \cup \{0\}$.

DEFINITION 3 *A* self-dual code *C is one for which $C = C^\perp$. A code is* weakly self-dual *if it there is a diagonal matrix $W \in K^{n+1,n+1}$ such that $w_{00} = 1$ and $Wc \in C^\perp$ for all $c \in C$. A code C is t-*locally weakly self-dual *if for all sets $B \subset \{1, \ldots, n\}$ with $|B| = n - t$ the code $\phi_B(C)$ is weakly self-dual.*

The definition of self-dual is standard in the coding literature, while our definition for weakly self-dual codes is a slight relaxation of the notion of quasi self-orthogonal[8] codes. The t-local variation appears to be novel. Simple examples are the following: the $[n+1, t+1, n-t+1]$-Reed Solomon code is weakly self-dual if $t < \frac{n}{2}$ and t-locally weakly self-dual if $t < \frac{n}{3}$. The following theorem demonstrates the relevance of these notions in secure computation.

---

[7] The dual Γ^* is defined as $A \in \Gamma^*$ if and only if $\{1, \ldots, n\} \setminus A \notin \Gamma$. It holds that $(\Gamma^*)^* = \Gamma$.

[8] For quasi self-orthogonal codes, the matrix W is required to be regular.

Theorem 2 *If C is a self-dual code of length $n+1$ with minimum distance d, then LSSS(C) offers linearity, t-privacy and $(n-t)$-reconstruction with $t = d-2$, and it has the multiplication property. If C is weakly self-dual, then C has the multiplication property and $t = d^{\perp} - 2$ if the matrix W is regular and otherwise $t = min\{d-2, d^{\perp}-2\}$. If C is t-locally weakly self-dual then LSSS(C) has the strong multiplication property with respect to the t-adversary structure.*

Proof. Since $d = d^{\perp}$ for self-dual codes, the privacy and reconstruction claims follow from Theorem 1. From $\langle c, c' \rangle = 0$ for all $c, c' \in C$ we get $c_0 c'_0 = -c_1 c'_1 - \cdots - c_n c'_n$. This implies the multiplication property (see [10, 11, 6] for the definition). For weakly self-dual codes, if W is regular then the minimum distance of WC is the same as that of C. Since $WC \subset C^{\perp}$, we must have $d^{\perp} \leq d$, and we apply Theorem 1. As to multiplication, we now have $\langle Wc, c' \rangle = 0$, so $c_0 c'_0 = -w_1 c_1 c'_1 - \cdots - w_n c_n c'_n$. The claim about the strong multiplication property is now obvious from the definition. △

We can generalize this as follows, using a twist on an idea from [10]. Let C be a code of length $n+1$ and minimum distance d. Consider the linear secret sharing scheme LSSS†(C) defined as follows. Take the secret s, and generate random shares $(c_1, \ldots, c_n)$ according to LSSS(C), and generate independently random shares $(c^*_1, \ldots, c^*_n)$ according to LSSS($C^{\perp}$). The share vector is then defined as $((c_1, c^*_1), \ldots, (c_n, c^*_n))$.

Theorem 3 *Let C be a code of length $n+1$ and minimum distance d. Define $t(C) = min\{d-2, d^{\perp}-2\}$. Then: LSSS†($C$) offers $t(C)$-privacy and $(n - t(C))$-reconstruction and it has the multiplication property. In particular, $t(C) < n/2$.*

The claim that $t(C) < n/2$ can for instance be verified by applying the Singleton-bound to C as well as to $C^{\perp}$. Note however that this scheme has information rate 1/2.

Strong multiplication is much more elusive and is not achieved by the construction above. In fact, the only way known to ensure strong multiplication (with respect to the t-adversary structure) for LSSS(C) is when C is an algebraic geometry code defined by the Riemann-Roch space of a divisor of degree $2g + t$ on a genus g algebraic curve over a finite field, where $3t < n - 4g$ [6]. If $2t < n - 4g$ it is weakly self-dual. For the special case where $g = 0$, these correspond to the well-known Reed-Solomon codes with the appropriate parameters.

3.3 Existence and Bounds

Our main objective in this section is to prove several lower bounds on the maximal value T taken over all values $t = \min\{d-2, d^{\perp}-2\}$ as C ranges over all K-linear codes of length $n+1$. In the following, an $[n+1, k]$-*code* C is simply a k-dimensional subspace of $\mathbb{F}_q^{n+1}$ and q is some fixed prime power. Where the parameters n and k are clear, $[n+1, k]$-code is simply abbreviated to code.

General lower bounds on T In Theorem 5 we give a general lower bound on the maximal t. In Corollary 2 we treat the general case when $K = \mathbb{F}_2$. In Theorem 6 we show that one can asymptotically get arbitrarily close to $\frac{1}{2}n$, over some constant size field. We also treat in that same corollary the parameterized case where C is randomly selected and a security parameter regulates the error probability that t is below a certain bound.

DEFINITION 4 *Let $n \in \mathbb{Z}_{>0}$ be fixed. Then $T(n+1,q) := \max_C t(C)$, where C ranges over all subcodes of $\mathbb{F}_q^{n+1}$. Similarly, $T'(n+1,q) := \max_C t(C)$, where C ranges over all weakly self-dual subcodes of $\mathbb{F}_q^{n+1}$.*

DEFINITION 5 *Let $\mathcal{C}_k$ have the uniform distribution over the set of $[n+1,k]$-subcodes of $\mathbb{F}_q^{n+1}$. Then we define*

$$T(n+1,q,m,k) := \max\{d-2 : P(\min\{d_{min}(\mathcal{C}_k), d_{min}(\mathcal{C}_k^{\perp})\} < d) < 2^{-m}\}$$

and $T(n+1,q,m) := \max_k T(n+1,q,m,k)$.

It is easy to see that $T(n+1,q) \geq T(n+1,q,0)$. The following lemma is trivial.

LEMMA 1 *Suppose $k \leq n$. For each pair (x,y) with $x \in \mathbb{F}_q^k \setminus \{0\}$ and $y \in \mathbb{F}_q^n \setminus \{0\}$ there exists an $n \times k$ matrix M of rank k such that $Mx = y$.*

The following theorem bounds the probability that a randomly chosen code has a minimum distance less than some fixed value d. It is used for most of the bounds that follow later.

THEOREM 4 *Let $\mathcal{C}$ have the uniform distribution over the set of $[n,k]$-subcodes of $\mathbb{F}_q^n$. Furthermore assume that $d = \alpha n \in \mathbb{Z}$, where $0 < \alpha < \frac{1}{2}$. Then*

$$P(\exists y \in \mathcal{C} : w_H(y) < d) < q^{k+n(H_q(\alpha)-1)},$$

where $H_q(\lambda) = \lambda \log_q(q-1) - \lambda \log_q \lambda - (1-\lambda)\log_q(1-\lambda)$.

PROOF. Let $\mathcal{H}$ have the uniform distribution over the set of $n \times k$ matrices of rank k over $\mathbb{F}_q$. Every such matrix corresponds to an ordered basis for a subcode V of $\mathbb{F}_q^n$. Since there is a one-to-one correspondence between the ordered bases for V and the linear isomorphisms between V and $\mathbb{F}_q^k$, each such subcode has the same number of ordered bases. Therefore, the variable $\mathcal{H}$ induces a uniformly random selection of an $[n,k]$-subcode of $\mathbb{F}_q^n$.

Fix some non-zero $x \in \mathbb{F}_q^k$. The variable $\mathcal{H}x$ then corresponds to a uniformly random selection from $\mathbb{F}_q^n$, which can be seen as follows: First, by Lemma 1 for any non-zero $y \in \mathbb{F}_q^n$ there exists an $n \times k$ matrix M of rank k such that $Mx = y$. Now fix some $y \in \mathbb{F}_q^n$ and assume that $Mx = y$ for some $n \times k$-matrix M of rank k. Then $\#\{M' : M'x = y\} = \#\{M' : (M-M')x = 0\} = \#\{M' : M'x = 0\}$, so for every $y \in \mathbb{F}_q^n$ there are the same number of matrices of rank k such that $Mx = y$.

Now let x range over the elements of $\mathbb{F}_q^k$. It follows that

$$P(\exists y \in \mathcal{C} : w_H(y) < d) = P(\exists x \in F_q^k : w_H(\mathcal{H}x) < d) \leq \sum_{x \in (\mathbb{F}_q^k)^*} P(w_H(\mathcal{H}x) < d)$$
$$= \frac{q^k - 1}{q^n - 1} \cdot \sum_{i=1}^{d-1} \binom{n}{i} (q-1)^i < \frac{q^k}{q^n} \cdot (q-1)^d \sum_{i=1}^{d-1} \binom{n}{i}$$
$$< \frac{q^k}{q^n} \cdot q^{\alpha n \log_q(q-1)} \cdot 2^{nH_2(\alpha)} = q^{k+n(H_q(\alpha)-1)}.$$

△

Since there is a one-to-one correspondence between subcodes C of $\mathbb{F}_q^n$ and their dual codes $C^\perp$, the random variable $\mathcal{C}^\perp$ corresponds to a uniformly random selection from the set of $[n, n-k]$-subcodes of $\mathbb{F}_q^n$. Therefore, we immediately obtain the following corollary.

Corollary 1 *Let $\mathcal{C}$ have the uniform distribution on the set of $[n,k]$-subcodes of $\mathbb{F}_q^n$. Furthermore assume that $d^* = \alpha n \in \mathbb{Z}$, where $0 < \alpha < \frac{1}{2}$. Then*

$$P(\exists y \in \mathcal{C}^\perp : w_H(y) < d^*) < q^{nH_q(\alpha)-k}.$$

Using the fact that $-\lambda \ln \lambda - (1-\lambda)\ln(1-\lambda) < 3.3\lambda$ for $1/10 \leq \lambda \leq 1/2$, we obtain that

$$H_q(\lambda) < \lambda \log_q(q-1) - \frac{3.3}{\ln q}\lambda \tag{1}$$

for $1/10 \leq \lambda \leq 1/2$. This gives rise to the following theorem.

Theorem 5 $T(n+1, q, m) \geq \lfloor \beta(n+1, q, m) \rfloor - 2$ *with*

$$\beta(n+1, q, m) = \frac{(n+1)\ln q - 2(m+1)\ln 2}{2\ln(q-1) + 6.6},$$

provided that $\lfloor \beta(n+1, q, m) \rfloor \geq n/10$.

Proof. Set $k = (n+1)/2$ and let $\mathcal{C}$ be as in Theorem 4. By Theorem 4 and Corollary 1,

$$P(\min\{d_{\min}(\mathcal{C}), d_{\min}(\mathcal{C}^\perp)\} < d) \leq P(d_{\min}(\mathcal{C}) < d) + P(d_{\min}(\mathcal{C}^\perp) < d)$$
$$< 2 \cdot q^{(n+1)H_q(\alpha) - (n+1)/2}.$$

We want $P(\min\{d_{\min}(\mathcal{C}), d_{\min}(\mathcal{C}^\perp)\} < d) < 2^{-m}$. Filling in (1) and rewriting, we see that this is the case if

$$d \leq \frac{(n+1)\ln q - 2(m+1)\ln 2}{2\ln(q-1) + 6.6}.$$

△

COROLLARY 2 *If $n \geq 21$, then $T(n+1,2) \geq \lfloor 0.1n \rfloor - 2$.*

THEOREM 6 *Fix any arbitrarily small $\epsilon > 0$ and any $m \in \mathbb{Z}_{>0}$. Then there exists a fixed finite field $\mathbb{F}_q$ over which for infinitely many n there exist $[n+1,k]$-codes $C \subset \mathbb{F}_q^{n+1}$ with $(1/2 - \epsilon)n \leq t(C) \leq n/2$ where such a code can be selected with probability at least $1 - 2^{-m}$ using a random selection among the $[n,k]$-subcodes of $\mathbb{F}_q^{n+1}$.*

PROOF. Let d be the minimum distance of C and $d^{\perp}$ the minimum distance of $C^{\perp}$. By Theorem 3, $t(C) < n/2$. Therefore, it suffices to show that $(d-2)$ and $(d^{\perp} - 2)$ can simultaneously get arbitrarily close to $n/2$ (relative to n) with probability at least $1 - 2^{-m}$.

By Theorem 5,

$$T(n+1,q,m) \geq \beta(n+1,q,m) - 2 = \frac{(n+1)\ln q - 2(m+1)\ln 2}{2\ln(q-1) + 6.6} - 2$$

and we have that

$$\begin{aligned}\lim_{q\to\infty} \frac{(n+1)\ln q - 2(m+1)\ln 2}{2\ln(q-1)+6.6} - 2 &= \lim_{q\to\infty} \frac{(n+1)\ln q}{2\ln(q-1)+6.6} - 2 \\ &\geq \lim_{q\to\infty} \frac{(n+1)\ln q}{2\ln q + 6.6} - 2.\end{aligned}$$

Since $\lim_{x\to\infty} \frac{x}{x+3.3} = \lim_{y\to\infty} \frac{y-3.3}{y} = \lim_{y\to\infty}(1 - \frac{3.3}{y}) = 1$, the final term converges to $(n+1)/2 - 2$ as $q \to \infty$. We can therefore for any $\delta > 0$ select a q large enough such that $T(n,q,m) \geq n/2 - 3/2 - \delta$. For large enough n, $(3/2+\delta)/n < \epsilon$ and the claim follows. △

So far we have assumed a random selection from the set of $[n,k]$-subcodes of $\mathbb{F}_q^n$. The lemma below demonstrates, together with the proof of Theorem 4, that we can in fact perform this random selection by selecting $n \times k$ matrices at random, where we obtain a matrix of rank k with probability at least $1/4$.

LEMMA 2 *The probability that a randomly selected $n \times k$-matrix over $\mathbb{F}_q$ has full rank is larger than $1 - 1/q - 1/q^2$.*

Bounds from (Weakly) Self-Dual Codes In Corollary 3 we prove a general lower bound on T for binary self-dual codes, and Theorem 8 shows that for $n < 100$ the situation is much better than the bound indicates. We are especially interested in self-dual codes, because secret sharing schemes based on self-dual codes do not suffer from the 1/2 information rate loss that occurs in the general case. Finally, in Theorem 9 we prove a much better lower bound for weakly self-dual codes based on algebraic geometry, and not random codes. Note that the results based on algebraic geometry are only known to hold if the size of the field is a square.

Theorem 7 *Let n be any positive integer and let d_{GV} be the largest integer such that*

$$\sum_{\substack{0<i<d \\ 2|i}} \binom{n}{i} < 2^{n/2-1} + 1.$$

Then there exists a self-dual binary code of length n and minimum distance at least d_{GV}.

Proof. See [21, 29, 27]. $\triangle$

Corollary 3 *Fix $\epsilon > 0$. For large enough n, $T'(n,2) \geq \lfloor(\delta - \epsilon)n\rfloor - 2$, where $\delta \approx 0.11002786$ is any truncated approximation of the unique solution less than $1/2$ of $H_2(\delta) = 1/2$.*

Proof. ([21, 29, 27]) Let $d = \alpha(n+1)$. Since for $\alpha < 1/2$, $\sum_{0<i<d} \binom{n+1}{i} \leq 2^{(n+1)H(\alpha)}$, the conditions of Theorem 7 are met if

$$(n+1)H(\alpha) \leq \frac{n+1}{2} - 1 \Leftrightarrow H(\alpha) \leq \frac{1}{2} - \frac{1}{n+1}.$$

The solution for α then comes arbitrarily close to δ as n increases.

$\triangle$

Theorem 8 *There exist self-dual binary codes C of length $n+1 < 100$ for which $d_{min}(C) > n/5$. In particular, there exist self-dual binary codes C with the following parameters:*

$n+1$	$d_{min}(C)$
12	4
22	6
24	8
46	10
48	12

Proof. See [14]. $\triangle$

Theorem 9 *When we take the maximum over algebraic geometry codes, then*

$$T(n+1, q^2) > \left(\frac{1}{2} - \frac{1}{q-1}\right) n.$$

Proof. This follows from a suitable choice of parameters for algebraic geometry codes and their duals and the existence of Garcia-Stichtenoth curves, using techniques similar to those in [6]. $\triangle$

For a corresponding result that ranges over t-locally weakly self-dual codes, see [6].

3.4 Application to VSS and Secure Computation

Using the results from Sections 3.2 and 3.3, we are now ready to discuss the fact that our specialized secret sharing schemes can essentially seamlessly replace Shamir's scheme in the broadcast model of Rabin/Ben-Or, yielding significant reductions in communication when working over a small field. More concretely, when operating in Beaver's preprocessing model [1] with a network of size n, this results in a reduction in communication amounting to a multiplicative polylogarithmic factor (in n) in the on-line phase, while tolerating $(\frac{1}{2} - \epsilon)n$ corrupted players, where $\epsilon > 0$ is arbitrarily small. Note that this may offer a considerable gain in case of very large networks. For similar results in the zero-error BGW model, see [6].

As an illustration, Theorem 6 together with Theorem 3 implies that for any $\epsilon > 0$, there exists a (fixed) finite field K and an infinite family of specialized secret sharing schemes tolerating a $(\frac{1}{2} - \epsilon)n$-fraction of corrupted players. We now focus on the communication-efficient protocol of Cramer, Damgaard and Fehr [12] and outline the main changes necessary to enable the use of these specialized secret-sharing schemes.The CDF protocol is stated in the broadcast model of Rabin and Ben-Or [26] supplemented with a preprocessing phase as introduced by Beaver [1]. The claimed reduction in communication will be achieved in the on-line phase of the adapted CDF protocol.

The model of Rabin and Ben-Or assumes the presence of a broadcast channel and induces a non-zero (negligible) error probability. In Beaver's model, an independent preprocessing phase is implemented, which can take place even before the selection of the type of computation, that is used to compute VSSes of random values and secret-shared "multiplication tables" of random values. The attractive feature of this model is that, during the subsequent *on-line phase* when the actual computation is performed, players only need to open a constant number of VSSes for every secure multiplication (which saves a lot of communication). Moreover, no secure channels are required at all during this on-line phase, as all communication is by broadcast. [9]

Briefly, the main changes are as follows. First, in VSS we modify the usual bivariate Shamir-sharing by using a technique from [10] for extending a linear secret sharing scheme so as to enable the pair-wise checking protocols for VSS. This is by having the fixed secret sharing matrix operating on random symmetric matrices, rather than on random vectors. This can trivially be adapted to our scenario here. Exactly as in the CDF protocol, the resulting two-level secret-sharings are then augmented with unconditionally secure Information Checking (IC) signatures. This completes the basis for VSS with a two-level sharing, where all shares and sub-shares are signed. Multiplication of VSS'ed values can be performed based on the linearity of the scheme and the multiplication property, while addition essentially comes for free due to linearity of the VSS itself.

The preprocessing in the CDF protocol is a secure multi-party computation that prepares VSSes of random multiplication tables, as well as VSSes of random

[9] In some implementations broadcast isn't even necessary in the on-line phase, but in our case it is.

inputs of players. The point however, is that by a specialized secure multi-party computation the CDF preprocessing strips off one layer of shares, resulting in VSSes with just a single layer of signed shares. This makes an on-line phase possible that is much more communication-efficient. We assume now that the security parameters are set so that these signatures in these one-level sharings are correct except with negligible probability. This can be done by repeating the information checking step sufficiently many times; the total amount of communication in this preprocessing phase would be the same as in CDF though, since our field is small.

In the on-line phase each player first VSSes his real inputs, by broadcasting the difference of this input with the random VSSed input that he has been given in the pre-processing. The corresponding VSS is accordingly updated (non-interactively). Secure computation in the on-line phase can subsequently take place. Note that, as opposed to CDF, we are working here over a constant size field. This means that, though the signatures themselves are correct with high probability as a result of the CDF preprocessing as instructed above, they are "so small" (as a matter of fact, equal to field elements) that successful forgeries can be constructed with high probability. Thus, when opening such a (stripped) VSS, a corrupted player could in principle make an individual honest player accept a false share with high probability, by guessing the "small signature value" for this individual player. An additional concern would be the following. For their use in secure addition and secure multiplication, these signatures enjoy a certain linearity property [9]. This requires, for each ordered pair of players, a secret key part held by one of those players. This part remains fixed throughout the protocol. Now, this fixed key part can be extracted from an honest player in a single successful forgery, which, as we have seen above, has a high probability of success. So, at first sight, there seems to be a risk that security might degrade fatally over time, if there was any in the first place.

What saves the day completely is the ϵ-gap with $n/2$ in the number of corrupted players, in combination with a simple elimination strategy regarding corrupted players. Consider a corrupted player, and focus on his very first attempt at cheating in the on-line phase. It is easy to see that if he doesn't modify his correct share, he can predict the behavior all of all honest players; rejection if the corresponding correct signature was modified and acceptance otherwise. This is due to the fact that the signature is deterministic given all secret information held by the receiver and the purported share. So, he cannot gain advantage unless he modifies the correct share. In our adaptation of the CFD protocol, we instruct that he broadcasts that purported share. Thus, if the correct share is modified, he must also modify the corresponding correct signatures for many honest players individually. More precisely, we instruct that a purported share is accepted only if a majority of the players individually accept it. This is done by local verification of individual signatures followed by majority voting using broadcast. [10] This means that he must guess the signatures for roughly ϵn hon-

[10] There is a slightly more sophisticated strategy involving error correction that gives still better error probabilities.

est players, so as to get a majority (assuming that the adversary appropriately coordinates this with the actions of the other $t-1$ corrupted players). Now, if the field size $|K|$ is, say, about $2/\epsilon$, then this probability is exponentially small in n. Note that we can always replace our original fixed finite field K with a large enough fixed extension field so that this condition holds, without changing the other parameters and properties of the underlying specialized secret sharing scheme. Thus, if a corrupted player makes his first attempt, he will be caught in the voting phase with very high probability, and he is subsequently eliminated from the network. This also means that the entropy of the fixed secret keys of all honest players remains essentially intact, so the error probability analysis is essentially the same throughout the on-line phase if n is indeed very large. The network then moves to the next computation with the remaining players, applying the same strategy as above. All in all, this reduces the communication by a multiplicative factor $(\log n)^2$, due to the fact that in the stripped VSS each of the n shares now carries a signature for each individual receiving player that is a $\log n$ factor smaller.

A Concrete Example. The case $K = \mathbb{F}_2$ is especially interesting, since the algebraic geometry results have no known strong bearing on this case. Our results show that in the secure channels model (passive case), secure multiplication over $\mathbb{F}_2$ can be done with just n^2 *bits* communication, with corruption tolerance of a constant fraction of n. This saves a multiplicative factor of $O(\log n)^2$ bits compared to the standard approach based on Shamir's scheme. For n below 100, about 20 percent of the network may be corrupted, while the underlying scheme is ideal due to the use of a self-dual code. For instance, with $n = 48 - 1 = 47$, an adversary corrupting $t = 12 - 2 = 10$ players can be tolerated. In the active adversary case (with preprocessing, as in [12]), the savings also amount to a multiplicative factor of $O(\log n)^2$ bits. For large networks these savings in communication can be rather substantial.

4 Ramp Schemes with High Information Rate

In a secret sharing scheme each subset of the player set is either rejected, which means that the shares held by the players in the given set jointly do not give any information about the underlying secret-shared value, or it is accepted, which means that those shares jointly determine that secret uniquely. In other words, there is no way in between. As a consequence (by an argument very similar to the one used to show that the key is at least the size of the plain-text in the perfectly secure one-time pad encryption scheme), the size of a share is at least the size of the secret.

In what is sometimes called a non-perfect secret sharing scheme, there is a third category of subsets, consisting of subsets whose joint shares gives some partial (but not full) information about the secret. In such schemes it is possible to have high information rate, i.e., the size of a share may be much smaller than the size of the secret.

Ramp schemes are a special case, and a variation on Shamir's threshold secret sharing scheme constitutes a well-known example [13]. This goes as follows. Let K be a finite field with $|K| > n+\ell$, let $x_1, \ldots, x_\ell, y_1, \ldots, y_n \in K$ be distinct and let the y_i's be non-zero. Let τ, ℓ be positive integers with $1 \leq \ell \leq \tau$. Consider a secret vector $\alpha \in K^\ell$ of length ℓ. Sample a polynomial $f(X) \in K[X]$ uniformly at random such that its degree is at most τ and such that $f(x_1) = \alpha_1, \ldots, f(x_\ell) = \alpha_\ell$, and define the shares as $s_1 = f(y_1), \ldots, s_n = f(y_n)$. This is a scheme on n players, and using Lagrange interpolation one proves that all player sets of size at least $\tau + 1$ are accepted, while all player sets of size at most $\tau + 1 - \ell$ are rejected. Note that the scheme has information rate ℓ, i.e., each player gets one element of K as a share while in fact the secret is a K-vector of length ℓ. In other words, this is an $(n, \tau + 1, \tau + 1 - \ell, \ell)$-ramp scheme over K. It is also linear in that each share is a K-linear combination of the coordinates of the secret vector and (random) field elements.

An alternative [7] is to encode the secret vector in the first ℓ lower order coefficients of the polynomial f instead. This yields a ramp scheme with the same parameters, except that the requirement on the size of the field K can be relaxed, namely, $|K| > n$ suffices here. Later we analyze this scheme in terms of our general results from Section 4.1 and in Section 4.3 we generalize this result in terms of algebraic geometry codes.

Interestingly, these two schemes give rise to complementary applications in secure computation. The first one to parallel secure multi-party computation with good amortized communication complexity [13], and the second to secure atomic multiplication with low communication [7].

We generalize Massey's scheme from Section 3.1 to high information rate ramp schemes in Section 4.1. In Section 4.2, we give a completely general construction that does not consume codelength (which corresponds to the number of players in the scheme) for an increased information rate. As an application we use this theory to analyse the alternative high information rate ramp scheme based on Shamir presented above. Also, our general method gives rise to a new high information rate ramp scheme based on algebraic geometry code which we introduce in Section 4.3.

4.1 A High Information Rate Ramp Scheme

Let C be an $[n + \ell, k, d]$-code over a finite field K. We now extend Massey's scheme from Section 3.1 in the direction of high information rate as follows. Let ℓ be a non-negative integer such that $\ell < d^\perp$.

Let $s \in K^\ell$. Select a codeword $c = (c'_0, \ldots, c'_{\ell-1}, c_1, \ldots, c_n) \in C$ at random such that $s = (c'_0, \ldots, c'_{\ell-1})$. Such c always exists. Define the coefficients of $(c_1, \ldots, c_n)$ to be the shares. We claim that this is a linear ramp scheme with information rate ℓ that has $(d^\perp - \ell - 1)$-privacy and $(n+l-d+1)$-reconstruction. This can be verified from the following facts.

Reconstruction follows from the fact that if there would exist two codewords in C that agreed on $n + l - d + 1$ share locations, their difference would give a codeword in C with Hamming weight less than d. As for privacy, note that in a generator matrix for C, any collection of $m < d^{\perp}$ rows (the code is generated by the columns) are linearly independent. So the corresponding columns span K^m. Therefore, for each $j \in \{0, \ldots, \ell-1\}$ and for each $A \subset \{1, \ldots, n\}$ with $|A| \leq d^{\perp} - \ell - 1$ there exists a codeword c such that $c'_j = 1$ and $c'_i = 0$ for all $i \in \{0, \ldots, \ell-1\} \setminus \{j\}$ and $c_u = 0$ for all $u \in A$. This implies privacy as claimed.

4.2 A More Fruitful Approach

A disadvantage of the scheme above is that it consumes code-length in exchange for secret-length. Below we describe an entirely general approach that doesn't have this disadvantage, and by means of which one can prove the existence of improved ramp schemes (see Section 4.3).

Let $\hat{C}$ and C be linear codes of length n over K, i.e., they are subspaces of the vector space K^n. Assume that C has dimension greater than 0 and that it is a proper subspace of $\hat{C}$. Choose an arbitrary linear code S such that

$$\hat{C} = S + C \text{ and } S \cap C = \{0\},$$

i.e., a direct sum. This is always possible of course, for instance by completing a basis of C to one of $\hat{C}$. Write

$$\ell = \dim_K \hat{C} - \dim_K C \ (= \dim_K S)$$

and fix an arbitrary isomorphism $\psi : K^{\ell} \longrightarrow S$.

We now define the following linear ramp scheme. Let $s \in K^{\ell}$ be the secret vector. Sample uniformly at random $c \in C$ and define the share vector $\hat{c}$ as $\hat{c} = \psi(s) + c$. [11]

Note that this is a generalization of a scheme used by Ozarow and Wyner [25], who considered the case $\hat{C} = K^n$. In fact, all possible linear ramp schemes are captured by this general scheme we consider here.

For $A \subset \{1, \ldots, n\}$, let ϕ_A denote the function $\phi_A : K^n \longrightarrow K^{|A|}$ where $(x_1, \ldots, x_n) \mapsto (x_i)_{i \in A}$, i.e., restriction to the coordinates labeled with A. Given A, consider the restriction of ϕ_A to $\hat{C}$. The set A is said to offer privacy if the collection of shares $\{\hat{c}_i\}_{i \in A}$ give no information on the secret vector, and reconstruction if those shares always determine the secret vector uniquely.

Theorem 10 *Let $\ell = \dim \hat{C} - \dim C$. The set A offers privacy if and only if $\dim \phi_A(\hat{C}) - \dim \phi_A(C) = 0$. The set A offers reconstruction if and only if $\dim \phi_A(\hat{C}) - \dim \phi_A(C) = \ell$. More generally, the uncertainty about the secret vector s, given the shares of A, is equal to r elements of K, where r is such that $\ell - r = \dim \phi_A(\hat{C}) - \dim \phi_A(C)$.*

[11] Equivalently, one can say that we fixed an arbitrary isomorphism from K^{ℓ} to $\hat{C}/C$, and that the share vector is selected by mapping s to the residue-class of $\psi(s)$ modulo C, and that $\hat{c}$ is chosen uniformly at random from that residue-class.

Proof. Privacy (for the set A) is equivalent to saying that for each possible secret vector $s \in K^{\ell}$, there is a share vector $\hat{c}$ that "encodes" s and that satisfies $\phi_A(\hat{c}) = 0$. This is the same as saying that for each $z \in S$, there exists $c \in C$ such that $0 = \phi_A(z + c) = \phi_A(z) + \phi_A(c)$. Thus, $\phi_A(\hat{C}) \subset \phi_A(C)$. Since the other inclusion holds regardless of A, the privacy claim follows. As for unique reconstruction (for the set A), this is equivalent to saying that there are no two distinct $z, z' \in S$ so that $\phi_A(z + c) = \phi_A(z' + c')$ for some $c, c' \in C$. This is equivalent to saying that $\dim \phi_A(S) = \ell$ and $\phi_A(S) \cap \phi_A(C) = \{0\}$. Since $\dim \phi_A(\hat{C}) - \dim \phi_A(C) = \dim \phi_A(S) - \dim \phi_A(S) \cap \phi_A(C)$, the reconstruction claim follows. The cases in between these two extremes should now be obvious.

△

We give the following estimate with respect to privacy and reconstruction (which, as one can prove by giving counter-examples, is not always sharp).

Corollary 4 *The set A offers privacy if $|A| < d_{min}(C^{\perp})$. The set A offers reconstruction if $|A| > n - d_{min}(\hat{C})$.*

Proof. As for privacy, if $|A| < d_{\min}(C^{\perp})$, then $\phi_A(C)$ clearly has rank $|A|$, since otherwise we could construct a codeword in $C^{\perp}$ whose weight is smaller than $d_{\min}(C^{\perp})$. Since $\phi_A(C) \subset \phi_A(\hat{C}) \subset K^{|A|}$, we must have $\phi_A(C) = \phi_A(\hat{C})$, and privacy follows from the theorem. As for reconstruction, if $|A| > n - d_{\min}(C)$, then $\phi_A(\hat{c}) = 0$ if and only if $\hat{c} = 0$, since otherwise C would contain a codeword whose weight is smaller than $d_{\min}(C)$. Thus, ϕ_A is injective when restricted to $\hat{C}$, and $\hat{c}$ follows uniquely from $\phi_A(\hat{c})$. Since $S \cap C = \{0\}$, $\psi(s)$ and c follow uniquely from $\hat{c}$. The secret vector s now follows uniquely from $\psi(s)$ since ψ is bijective.

△

Note that from the Singleton-bound, we have $\dim_K \hat{C} \leq n - d_{\min}(\hat{C}) + 1$ and $d_{\min}(C^{\perp}) - 1 \leq n - \dim_K C^{\perp} = \dim_K C$. Thus, $r - t \geq \dim_K \hat{C} - \dim_K C$ in any linear ramp scheme.

Before presenting constructive results, we argue as an example that the Shamir ramp scheme discussed earlier can be easily analyzed with this theory. Suppose $n > |K|$, and let $x_1, \ldots, x_n$ be distinct non-zero elements of K. Consider the Vandermonde matrix M with n rows and t columns whose i-th row is $(1, x_i, \ldots, x^t)$. Let $\hat{C}$ be the code generated by all the columns. This is an $(n, t+1, n-t)$-MDS code. So its dual is an $(n, n-t-1, t+2)$-code. Let C be the code generated by the last $t + 1 - \ell$ columns. Clearly $C \subset \hat{C}$. By appropriately scaling the rows of C it is immediate that C is equivalent to an $(n, t+1-\ell, n-t+\ell)$-code. This is an MDS code, so its dual is an $(n, n-t-1+\ell, t+2-\ell)$-code. So by our theorem the resulting ramp scheme rejects all sets of size $t + 1 - \ell$, and accepts all sets of size $t + 1$. Note that the gap between the two bounds here is ℓ, so that is optimal.

4.3 High Information Rate Ramp Schemes: Existence and Bounds

In this section we demonstrate two methods for constructing high information rate ramp schemes. First, we present a new high information rate ramp scheme that improves the one presented in [6], where $\hat{C}$ will be an algebraic geometry code and C will be a carefully selected algebraic geometry subcode of $\hat{C}$. Then, we demonstrate that high information rate ramp schemes can be obtained from random codes and bound the error probabilities on their predicted parameters.

Algebraic Geometry Codes Select an absolutely irreducible smooth projective curve over a finite field K, write g for its genus and let $\{Q, P_1, P_2, \ldots, P_n\}$ denote distinct points on the curve. Consider the rational divisor $\hat{D} = (2g+t) \cdot Q$, and let $\mathcal{L}(\hat{D})$ denote the corresponding Rieman-Roch space of rational functions. Write $\hat{C}$ for the Goppa-code consisting of the codewords $(f(P_1), \ldots, f(P_n))$, where f ranges over $\mathcal{L}(\hat{D})$. Also define the rational divisor $D = (2g + t - \ell) \cdot Q$, and let $\mathcal{L}(D)$ denote the corresponding Rieman-Roch space of rational functions. Write C for the Goppa-code consisting of the codewords $(f(P_1), \ldots, f(P_n))$, where f ranges over $\mathcal{L}(D)$.

By the Riemann-Roch Theorem the dimension of $\hat{C}$ is $g + t + 1$, whereas the dimension of C is $g + t + 1 - \ell$. Since $\hat{D} \geq D$, we have $\mathcal{L}(D) \subset \mathcal{L}(\hat{D})$, and hence $C \subset \hat{C}$. It is fact that the minimum distance of $C^{\perp}$ is at least $\deg D - 2g + 2 = t - \ell + 2$. Furthermore, it has been proven in [6] that we have reconstruction for $\deg \hat{D} + 1 = 2g + t + 1$ shares. Thus, by our theorem, we have a linear ramp scheme over K with $t - \ell + 1$ privacy, $2g + t + 1$ reconstruction and information rate ℓ. Note that the improvement consists in the fact that the scheme above does not use up any points on the curve in order to encode the secret vector. Also note that by taking the projective line (i.e., $g = 0$) we recover the earlier Shamir ramp scheme example. Using Garcia-Stichtenoth towers [15] our ramp scheme can be defined over constant size fields. See [6] for more details.

Random Codes Finally, the results in Section 3.3 demonstrate that we can also obtain high information rate ramp schemes from randomly selected codes $\hat{C}$ and C, provided that $C \subset \hat{C}$. Theorem 10 demonstrates that for such codes C and $\hat{C}$, the corresponding ramp scheme provides privacy for any subset consisting of at most $d_{\text{min}}\mathcal{C}^{\perp}) - 1$ players and reconstruction for any subset consisting of at least $n - d_{\text{min}}(\hat{\mathcal{C}}) + 1$ players.

One method of obtaining the appropriate distribution for C and $\hat{C}$, as demonstrated in the proof of Theorem 4, is to randomly select a matrix M from the set of $n \times \hat{k}$-matrices of rank $\hat{k}$ and let $\hat{C}$ be the code spanned by the columns. It is easy to see that if we now look at the last k columns of M, these columns in turn span a random $[n, k]$-subcode C of K^n that is furthermore contained in $\hat{C}$. Clearly, the corresponding scheme allows for a secret vector of length $\ell = \hat{k} - k$.

Suppose that we want the scheme to provide privacy for up to t players and reconstruction for at least $n - \hat{t}$ players. Using a similar argument as in Theorem 4 and using the fact that $-\lambda \ln \lambda - (1 - \lambda) \ln(1 - \lambda) < 1.2\sqrt{\lambda}$ for $0 \leq \lambda \leq 1/2$, the following theorem is now straightforward to obtain. It provides, for many different parameters and with arbitrarily high probability, a lower bound on t and $\hat{t}$ when we select the codes C and $\hat{C}$ at random.

THEOREM 11 *Select an $[n,k]$-code C and an $[n,\hat{k}]$-code $\hat{C}$ over $\mathbb{F}_q$ at random under the restriction that $C \subset \hat{C}$. Then*

$$P(d_{min}(\mathcal{C}^{\perp}) < t) < q^{-(k - t\log_q(q-1) - \frac{1.2\sqrt{tn}}{\ln q})}$$

and

$$P(d_{min}(\hat{\mathcal{C}}) < \hat{t}) < q^{-(n-\hat{k}-\hat{t}\log_q(q-1) - \frac{1.2\sqrt{\hat{t}n}}{\ln q})}.$$

References

1. D. Beaver. Efficient multiparty protocols using circuit randomization. In *Proceedings of CRYPTO '91*, volume 576, pages 420–432. Springer Verlag LNCS, 1992.
2. M. Ben-Or, S. Goldwasser, and A. Wigderson. Completeness theorems for non-cryptographic fault-tolerant distributed computation. In *Proceedings of STOC 1988*, pages 1–10. ACM Press, 1988.
3. G. R. Blakley. Safeguarding cryptographic keys. In *Proceedings of National Computer Conference '79*, volume 48 of *AFIPS Proceedings*, pages 313–317, 1979.
4. G. R. Blakley and C. Meadows. Security of ramp schemes. In *Proceedings CRYPTO '85*, volume 196, pages 242–269. Springer Verlag LNCS, 1985.
5. D. Chaum, C. Crépeau, and I. Damgaard. Multi-party unconditionally secure protocols. In *Proceedings of STOC 1988*, pages 11–19. ACM Press, 1988.
6. H. Chen and R. Cramer. Algebraic Geometric Secret Sharing Schemes and Secure Multi-Party Computation over Small Fields. In *Proceedings of 26th Annual IACR CRYPTO*, volume 4117, pages 516–531, Santa Barbara, Ca., USA, August 2006. Springer Verlag LNCS.
7. R. Cramer, I. Damgaard, and R. de Haan. Atomic Secure Multi-Party Multiplication with Low Communication. In *Proceedings of EUROCRYPT 2007*, May 2007.
8. R. Cramer, I. Damgaard, and S. Dziembowski. On the complexity of verifiable secret sharing and multi-party computation. In *Proceedings of STOC 2000*, pages 325–334. ACM Press, 2000.
9. R. Cramer, I. Damgaard, S. Dziembowski, M. Hirt, and T. Rabin. Efficient Multi-Party Computations with Dishonest Minority. In *Proceedings of 18th Annual IACR EUROCRYPT*, volume 1592, pages 311–326, Prague, Czech Republic, May 1999. Springer Verlag LNCS.
10. R. Cramer, I. Damgaard, and U. Maurer. General secure multi-party computation from any linear secret sharing scheme. In *Proceedings of EUROCRYPT 2000*, volume 1807 of *LNCS*, pages 316–334. Springer Verlag, 2000.
11. R. Cramer, V. Daza, I. Gracia, J. Jimenez Urroz, G. Leander, J. Martí-Farré, and C. Padró. On codes, matroids and secure multi-party computation from linear secret sharing schemes. In *Proceedings of CRYPTO 2005*, volume 3621 of *LNCS*, pages 327–343. Springer-Verlag, 2005.
12. R. Cramer, I. Damgård, and S. Fehr. On the Cost of Reconstructing a Secret– Or: VSS with Optimal Reconstruction. In *Proceedings of 21th Annual IACR CRYPTO*, volume 2139, pages 503–523, Santa Barbara, Ca., USA, August 2001. Springer Verlag LNCS.

13. M. Franklin and M. Yung. Communication complexity of secure computation. In *Proceedings of STOC 1992*, pages 699–710. ACM Press, 1992.
14. P. Gaborit and A. Otmani. Experimental constructions of self-dual codes. Manuscript. Available from `http://www.unilim.fr/pages_perso/philippe.gaborit/SD/`, 2002.
15. A. García and H. Stichtenoth. On the asymptotic behavior of some towers of function fields over finite fields. *J. Number Theory*, 61:248–273, 1996.
16. O. Goldreich, S. Micali, and A. Wigderson. How to Play Any Mental Game. In *Proceedings of STOC 1987*, pages 218–229. ACM Press, 1987.
17. V. D. Goppa. Codes on algebraic curves. *Soviet Math. Dokl*, 24:170–172, 1981.
18. M. Karchmer and A. Wigderson. On span programs. In *Proceedings of the Eight Annual Structure in Complexity Theory Conference*, pages 102–111. IEEE, 1993.
19. K. Kurosawa, K. Okada, K. Sakano, W. Ogata, and S. Tsujii. Nonperfect Secret Sharing Schemes and Matroids. In *Proceedings EUROCRYPT 1993*, pages 126–141. Springer Verlag, 1993.
20. S. Lang. *Algebra*. Addison-Wesley Publishing Company, 1997.
21. F. J. MacWilliams, N. J. A. Sloane, and J. G. Thompson. Good self-dual codes exist. *Discrete Math.*, 3:153–162, 1972.
22. J. L. Massey. Minimal codewords and secret sharing. In *Proceedings of the 6-th Joint Swedish-Russian Workshop on Information Theory*, pages 269–279, Molle, Sweden, August 1993.
23. J. L. Massey. Some applications of coding theory in cryptography. *Codes and Ciphers: Cryptography and Coding IV*, pages 33–47, 1995.
24. W. Ogata and K. Kurosawa. Some Basic Properties of General Nonperfect Secret Sharing Schemes. *J. UCS*, 4(8):690–704, 1998.
25. L. H. Ozarow and A. D. Wyner. "Wire-tap-channel II". *AT&T Bell Labs Tech. J.*, 63:2135–2157, 1984.
26. T. Rabin and M. Ben-Or. Verifiable secret sharing and multiparty protocols with honest majority. In *Proceedings of ACM STOC 1989*, pages 73–85, 1989.
27. E. M. Rains and N. J. A. Sloane. Self-Dual Codes. A long survey article written for the Handbook of Coding Theory. Available from `http://www.research.att.com/~njas/`, 1998.
28. A. Shamir. How to share a secret. *Communications of the ACM*, 22(11):612–613, 1979.
29. J. G. Thompson. Weighted averages associated to some codes. *Scripta Math.*, 29:449–452, 1973.
30. J. H. van Lint. *Introduction to Coding Theory*. Graduate Texts in Mathematics. Springer Verlag, 1999.
31. V. K. Wei. Generalized Hamming Weights for Linear Codes. *IEEE Transactions on Information Theory*, 37(5):1412–1418, 1991.