entropy

MDPI

Article

Bounds for Coding Theory over Rings

Niklas Gassner [1], Marcus Greferath [2], Joachim Rosenthal [1] and Violetta Weger [3,*]

[1] Institute of Mathematics, University of Zurich, 8057 Zurich, Switzerland
[2] School of Mathematics and Statistics, University College of Dublin, D04 V1W8 Dublin, Ireland
[3] Department of Computer Engineering, Technical University of Munich, 80333 München, Germany
* Correspondence: violetta.weger@tum.de

Abstract: Coding theory where the alphabet is identified with the elements of a ring or a module has become an important research topic over the last 30 years. It has been well established that, with the generalization of the algebraic structure to rings, there is a need to also generalize the underlying metric beyond the usual Hamming weight used in traditional coding theory over finite fields. This paper introduces a generalization of the weight introduced by Shi, Wu and Krotov, called overweight. Additionally, this weight can be seen as a generalization of the Lee weight on the integers modulo 4 and as a generalization of Krotov's weight over the integers modulo 2^s for any positive integer s. For this weight, we provide a number of well-known bounds, including a Singleton bound, a Plotkin bound, a sphere-packing bound and a Gilbert–Varshamov bound. In addition to the overweight, we also study a well-known metric on finite rings, namely the homogeneous metric, which also extends the Lee metric over the integers modulo 4 and is thus heavily connected to the overweight. We provide a new bound that has been missing in the literature for homogeneous metric, namely the Johnson bound. To prove this bound, we use an upper estimate on the sum of the distances of all distinct codewords that depends only on the length, the average weight and the maximum weight of a codeword. An effective such bound is not known for the overweight.

Keywords: rings; coding theory; Johnson bound; Plotkin bound

Citation: Gassner, N.; Greferath, M.; Rosenthal, J.; Weger, V. Bounds for Coding Theory over Rings. *Entropy* **2022**, *24*, 1473. https://doi.org/10.3390/e24101473

Academic Editors: Onur Günlü, Rafael F. Schaefer, Holger Boche and H. Vincent Poor

Received: 15 September 2022
Accepted: 12 October 2022
Published: 16 October 2022

1. Introduction

Coding theoretic experience has shown that considering linear codes over finite fields often yields significant complexity advantages over the nonlinear counterparts, particularly when it comes to complex tasks such as encoding and decoding. On the other side, it was recognized early [1,2] that the class of binary block codes contains excellent code families, which were not linear (Preparata, Kerdock codes, Goethals and Goethals–Delsarte codes). For a long time, it could not be explained why these families exhibit formal duality properties in terms of their distance enumerators that occur only on those among linear codes and their duals.

A true breakthrough in the understanding of this behavior came in the early 1990s when, after preceding work by Nechaev [3], the paper by Hammons et al. [4] discovered that these families allow a representation in terms of $\mathbb{Z}_4$-linear codes.

A crucial condition for this ring-theoretic representation was that $\mathbb{Z}_4$ was equipped with an alternative metric, the Lee weight, rather than with the traditional Hamming weight, which only distinguishes whether an element is zero or non-zero. The Lee weight is finer, assigning 2 a higher weight than the other non-zero elements of this ring.

The fact that the traditional settings of linear coding theory (finite fields endowed with the Hamming metric) are actually too narrow, which suggests expanding the theory in at least two directions: on the algebraic part, the next more natural algebraic structure serving as alphabet for linear coding is that of finite rings (and modules). On the metrical part, the appropriateness of the Lee weight for $\mathbb{Z}_4$-linear coding suggests that the distance function for a generalized coding theory requires generalization as well.

Since these ground-breaking observations, an entire discipline arose within algebraic coding theory. A considerable community of scholars have been developing results in various directions, among them code duality, weight-enumeration, code equivalence, weight functions, homogeneous weights, existence bounds, code optimality and decoding schemes, to mention only a few.

The paper at hand aims at providing a further contribution to this discipline, by introducing the overweight on a finite ring. This weight is a generalization of the Lee weight over $\mathbb{Z}_4$, as well as of the weight introduced in [5] by Krotov over $\mathbb{Z}_{2^s}$ for any positive integer s, which was further generalized to $\mathbb{Z}_{p^k}$ in [6].

We study the relations of this new weight to other well-known weights over rings and state several properties of the overweight, such as its extremal property. We also develop a number of standard existence bounds, such as a Singleton bound, a sphere-packing bound, a Plotkin bound and a version of the (assertive) Gilbert–Varshamov bound.

In the final part of this article, we derive a general Johnson bound for the homogeneous weight on a finite Frobenius ring. This result is important, as it is closely connected to list decoding capabilities.

2. Preliminaries

Throughout this paper, we will consider R to be a finite ring with identity, denoted by 1. If R is a finite ring, we denote by $R^\times$ its group of invertible elements, also known as units.

Let us recall some preliminaries in coding theory, where we focus on ring-linear coding theory.

For a prime power q, let us denote by $\mathbb{F}_q$ the finite field with q elements and, for a positive integer m, we denote by $\mathbb{Z}_m$ the ring of integers modulo m.

In traditional coding theory, we consider a linear code to be a subspace of a vector space over a finite field.

Definition 1. *Let q be a prime power, and let $k \leq n$ be non-negative integers. A linear subspace C of $\mathbb{F}_q^n$ of dimension k is called a linear $[n, k]$ code.*

We define a weight in a general way.

Definition 2. *Let R be a finite ring. A real-valued function w on R is called a* weight *if it is a non-negative function that maps 0 to 0.*

It is natural to identify w with its additive extension to R^n, and so we will always write $w(x) = \sum_{i=1}^n w(x_i)$ for all $x \in R^n$. Every weight w on R induces a *distance* $d : R \times R \longrightarrow \mathbb{R}$ by $d(x, y) = w(x - y)$. Again, we will identify d with its natural additive extension to $R^n \times R^n$.

If the weight additionally is positive definite, symmetric and satisfies the triangular inequality, that is,

1. $w(0) = 0$ and $w(x) > 0$ for all $x \neq 0$,
2. $w(x) = w(-x)$ for all $x \in R$,
3. $w(x + y) \leq w(x) + w(y)$ for all $x, y \in R$,

then the induced distance inherits these properties, i.e.,

1. $d(x, y) \geq 0$ for all $x, y \in R$ and $d(x, y) = 0$ if and only if $x = y$.
2. $d(x, y) = d(y, x)$ for all $x, y \in R$,
3. $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in R$.

The most prominent and best studied weight in traditional coding theory is the Hamming weight.

Definition 3. *Let $n \in \mathbb{N}$. The Hamming weight of a vector $x \in R^n$ is defined as the size of its support*

$$w_H(x) = |\{i \in \{1, \ldots, n\} \mid x_i \neq 0\}|,$$

and the Hamming distance between x and $y \in R^n$ is given by

$$d_H(x, y) = |\{i \in \{1, \ldots, n\} \mid x_i \neq y_i\}| = w_H(x - y).$$

The minimum Hamming distance of a code is then defined as the minimum distance between two different codewords

$$d_H(C) = \min\{d_H(x, y) \mid x, y \in C,\ x \neq y\}.$$

Note that the concept of minimum distance can be applied for any underlying distance d.

In the paper at hand, we focus on a more general setting where the ambient space is a module over a finite ring.

Definition 4. *Let $n \in \mathbb{N}$ and let R be a finite ring. A submodule C of ${}_R R^n$ of size $M = |C|$ is called a left R-linear (n, M) code.*

The most studied ambient space for ring-linear coding theory is the integers modulo 4, denoted by $\mathbb{Z}_4$, endowed with the Lee metric.

Definition 5. *For $x \in \mathbb{Z}_m$, its Lee weight is defined as*

$$w_L(x) = \min\{x, \mid m - x \mid\}.$$

One of the most prominent generalizations of the Lee weight over $\mathbb{Z}_4$ is the homogeneous weight.

Definition 6. *Let R be a Frobenius ring. A weight $w : R \longrightarrow \mathbb{R}$ is called (left) homogeneous of average value $\gamma > 0$, if $w(0) = 0$ and the following conditions hold:*

(i) For all x, y with $Rx = Ry$, we have that $w(x) = w(y)$.

(ii) For every non-zero ideal $I \leq {}_R R$, it holds that

$$\frac{1}{|I|} \sum_{x \in I} w(x) \ = \ \gamma.$$

We will denote the homogeneous weight with wt.

The homogeneous weight was first introduced by Constantinescu and Heise in [7] in the context of coding over integer residue rings. It was later generalized by Greferath and Schmidt [8] to arbitrary finite rings, where the ideal I in Definition 6 was assumed to be a principal ideal. In its original form, however, the homogeneous weight only exists on finite Frobenius rings. It can be shown that a left homogeneous weight is at the same time right homogeneous, and for this reason, we will omit the reference to any side for the sequel. In [9], Honold and Nechaev finally generalized the notion of homogeneous weight to some finite modules, called weighted modules, over a (not necessarily commutative) ring R with identity.

Since we will establish a Plotkin bound for a new weight, let us recall here the Plotkin bound over finite fields equipped with the Hamming metric.

Theorem 1 (Plotkin bound). *Let C be an (n, M) block code over $\mathbb{F}_q$ with minimum Hamming distance d. If $d > \frac{q-1}{q}n$, then*

$$M \leq \frac{d}{d - \frac{q-1}{q}n}.$$

For the homogeneous weight, the following Plotkin bound was established in [10].

Theorem 2 (Plotkin bound for homogeneous weights, [10]). *Let wt be a homogeneous weight of average value γ on R, and let C be an (n, M) block code over R with minimum homogeneous distance d. If $\gamma n < d$, then*

$$M \leq \frac{d}{d - \gamma n}.$$

3. Overweight

As the Hamming weight defined over the binary can be generalized to larger ambient spaces in different ways resulting in different metrics, such as the Hamming weight over $\mathbb{F}_q$ or the Lee weight over $\mathbb{Z}_{p^s}$; in addition, the Lee weight over $\mathbb{Z}_4$ can be generalized in different ways. For example, the weight defined in [5] over $\mathbb{Z}_{2m}$ for any positive integer m is a possible generalization, but the most prominent generalization is the homogeneous weight (see for example [10]). In this section, we introduce a new generalization, called the *overweight*. This weight shows some interesting properties and relations to the homogeneous weight and can additionally be seen as a generalization of the weight defined in [5] over $\mathbb{Z}_{2^s}$ for any positive integer s and the weight defined in [6] over $\mathbb{Z}_{p^s}$.

Definition 7. *Let R be a finite ring. The* overweight *on R is defined as*

$$W : R \longrightarrow \mathbb{R}, \quad x \mapsto \begin{cases} 0 & \text{if } x = 0, \\ 1 & \text{if } x \in R^\times, \\ 2 & \text{otherwise.} \end{cases}$$

We also denote by W its additive expansion to R^n, given by $W(x) = \sum_{i=1}^{n} W(x_i)$.

Let us call the distance which is induced by the overweight the *overweight distance*, and denote it by D, i.e., $D(x, y) = W(x - y)$.

The motivation of introducing this new weight is twofold: on one hand, it is theoretically interesting to explore a new generalization of the Lee weight over $\mathbb{Z}_4$ and its connections to other known weights over rings. On the other hand, the overweight would also be perfectly suitable for a channel, where unit errors are more likely.

Note that the overweight is designed to satisfy the following criteria: it is positive definite, symmetric, satisfies the triangular inequality and distinguishes between units and non-zero non-units. Furthermore, it is extremal in the sense that, on a big family of rings, any increase of the weight of non-zero non-units would violate the triangular inequality, thus the name *overweight*. We will now study this extremal property in more details.

We can consider weights with values in $\{0, 1, \alpha\}$, for some $\alpha > 0$, without fixing the subsets of R where these values are attained. Thus, we are considering the generic weight function

$$f(x) = \begin{cases} 0 & \text{if } x = 0, \\ 1 & \text{if } x \in A_1, \\ \alpha & \text{if } x \in A_2, \end{cases}$$

where $A_1 \subset R \setminus \{0\}$ and $A_2 = R \setminus (A_1 \cup \{0\})$. Such a weight is always positive definite. In addition, the weight is symmetric if and only if A_1 and A_2 contain all additive inverses of

their elements. Let us now consider the triangular inequality: if there exist $x, y \in A_1$ such that $x + y \in A_2$, then we must have

$$\alpha = f(x+y) \leq f(x) + f(y) = 2.$$

Thus, in order for f to be an extremal weight, one chooses $\alpha = 2$.

The overweight is a special case of such a weight function f with the choice $A_1 = R^\times$. The existence of elements $x, y \in R^\times$ such that $x + y \in R \setminus (\{0\} \cup R^\times)$ is satisfied for many rings—for example, for rings with a non-trivial Jacobson radical.

Relations to Other Weights

Clearly, the homogeneous weight and the overweight coincide with the Lee weight on $\mathbb{Z}_4$, with the Hamming metric on finite fields $\mathbb{F}_q$, and finally with the weight [6] on $\mathbb{Z}_{p^s}$.

Proposition 1. *The overweight over finite chain rings gives an upper bound on the normalized homogeneous weight.*

Proof. Over a finite chain ring with socle S and residue field size q, we have that the normalized homogeneous weight is defined as

$$wt(x) = \begin{cases} 0 & \text{if } x = 0, \\ \frac{q}{q-1} & \text{if } x \in S \setminus \{0\}, \\ 1 & \text{else.} \end{cases}$$

If $x \in S \setminus \{0\}$, then also $x \in R \setminus R^\times$, and

$$wt(x) = \frac{q}{q-1} \leq 2 = W(x).$$

If $x \in R^\times$, then $wt(x) = 1 = W(x)$ and finally, if $x \in R \setminus (S \cup R^\times)$, we have that

$$wt(x) = 1 \leq 2 = W(x),$$

which implies the result. □

In [11], Bachoc defines the following weight on $\mathbb{F}_p$-algebras A, with units $A^\times$ as follows:

$$w_B(x) = \begin{cases} 0 & \text{if } x = 0, \\ 1 & \text{if } x \in A^\times, \\ p & \text{else.} \end{cases}$$

This is in the same spirit as the overweight. The weight of Bachoc is, however, only assuming positive definiteness. We note that, whenever we have a $\mathbb{F}_2$-algebra, the two weights coincide. The overweight can thus also be seen as a generalization of Bachoc's weight to a general finite ring.

Let us illustrate this connection with some examples: we consider the ring $M_2(\mathbb{F}_p)$ of 2×2 matrices over $\mathbb{F}_p$ and the ring $\mathbb{F}_p[x]/(x^2)$. In both cases, the Bachoc weight only coincides with the homogeneous and the overweight in the case $p = 2$.

Finally, in [5], Krotov defines the following weight over $\mathbb{Z}_{2m}$, for any positive integer m:

$$w_K(x) = \begin{cases} 0 & \text{if } x = 0, \\ 2 & \text{if } 2 \mid x, x \neq 0, \\ 1 & \text{else.} \end{cases}$$

Clearly, this is a further generalization of the Lee weight over $\mathbb{Z}_4$ and thus coincides there with the homogeneous and the overweight. However, even more is true: the weight of Krotov and the overweight coincide over $\mathbb{Z}_{2^s}$, for any positive integer s. Thus, the overweight may be considered as a generalization of Krotov's weight over $\mathbb{Z}_{2^s}$ for any positive integer s.

Let us give some examples to illustrate the differences between the above-mentioned weights.

Example 1. *In the following table, w_H denotes the Hamming weight, wt the normalized homogeneous weight, w_L denotes the Lee weight, w_K denotes Krotov's weight, w_B denotes Bachoc's weight and finally W denotes the overweight. Let us consider two easy but pathological cases, namely $\mathbb{Z}_6$ for Table 1 and $\mathbb{Z}_2 \times \mathbb{Z}_2$ for Table 2.*

Table 1. Comparison of weights in $\mathbb{Z}_6$.

	w_H	wt	w_L	w_K	W
0	0	0	0	0	0
1	1	1/2	1	1	1
2	1	3/2	2	2	2
3	1	2	3	1	2
4	1	3/2	2	2	2
5	1	1/2	1	1	1

Table 2. Comparison of weights in $\mathbb{Z}_2 \times \mathbb{Z}_2$.

	w_H	wt	w_B	W
(0,0)	0	0	0	0
(0,1)	1	2	2	2
(1,0)	1	2	2	2
(1,1)	2	0	1	1

Finally, another interesting connection to the Hamming weight arises by considering the following linear injective isometry.

Lemma 1. *The map*

$$\begin{aligned} \psi : (\mathbb{F}_2[x]/(x^2), W) &\to (\mathbb{F}_2^2, w_H) \\ a + bx &\mapsto (a + b, b) \end{aligned}$$

is a linear isometry.

Recall that, over $\mathbb{F}_2[x]/(x^2)$, the overweight coincides with the weight of Bachoc and the homogeneous weight.

4. Bounds for the Overweight

In this section, we develop several bounds for the overweight, such as a Singleton bound, a sphere-packing bound, a Gilbert–Varshamov bound and a Plotkin bound.

For this, let us first define the minimum overweight distance of a code.

Definition 8. *Let $C \subseteq R^n$ be a code. The minimum overweight distance of C is then denoted by $D(C)$ and defined as*

$$D(C) = \min\{D(x, y) \mid x, y \in C, x \neq y\}.$$

4.1. A Singleton Bound

The Singleton bound usually follows a puncturing argument, which is possible for the overweight, but gives the same result as applying the following observation:

Remark 1. *For all $x \in R$, we have that*

$$0 \leq w_H(x) \leq W(x) \leq 2w_H(x) \leq 2n,$$

where w_H denotes the Hamming weight.

Hence, using the Singleton bound for the Hamming metric directly gives a Singleton bound for the overweight.

Proposition 2. *Let $C \subseteq R^n$ be a code of size M and minimum overweight distance d. Then,*

$$d \leq 2(n - \lceil \log_{|R|}(M) \rceil + 1).$$

Example 2. *A trivial example for a code achieving the Singleton bound in Proposition 2 is given by $C = \langle (p, \ldots, p) \rangle \subset \mathbb{Z}_{p^s}^n$, having $\log_{p^s}(|\ C\ |) = \frac{s-1}{s}$ and minimum overweight distance $d = 2n$.*

However, if we define the rank of a linear code C, denoted by $rk(C)$, to be the minimal number of generators of C, then the following bound is known for principal ideal rings [12,13]

$$d_H(C) \leq n - rk(C) + 1.$$

Codes achieving this bound are called Maximum Distance with respect to Rank (MDR) codes, in order to differentiate from MDS codes. This is a sharper bound than the usual Singleton bound, since for non-free codes we have $rk(C) > \log_{|R|}(M)$.

In the case of linear codes, the rank thus also leads to a sharper Singleton-like bound for the overweight.

Proposition 3. *Let R be a principal ideal ring. Let $C \subseteq R^n$ be a linear code of rank $rk(C)$ and minimum overweight distance d. Then,*

$$d \leq 2(n - rk(C) + 1).$$

Example 3. *As an example for a code, we can consider $C = \langle (3,6,3,0), (6,6,0,3) \rangle \subset \mathbb{Z}_9^4$, having minimum overweight distance $d = 6$.*

4.2. A Sphere-Packing Bound

The sphere-packing bound as well as the Gilbert–Varshamov bound are *generic* bounds, and we are able to provide them for the overweight in a simple form involving the volume of the balls in the underlying metric space.

We begin by defining balls with respect to the overweight distance.

Definition 9. *For a given radius $r \geq 0$, the* overweight ball *$B_{r,D}(x)$ of radius r centered in x is defined as*

$$B_{r,D}(x) \ := \ \{y \in R^n \mid D(x,y) \leq r\}.$$

Clearly, the volume of such a ball is invariant under translations, i.e.,

$$|B_{r,D}(x)| \ = \ |B_{r,D}(y)|,$$

for all $x, y \in R^n$.

Moreover, setting $u := |R^\times|$ and $v := |R| - 1 - u$, we have the generating function $f_W(z) = 1 + uz + vz^2$ for this weight function, so that the generating function for W on R^n takes the form

$$\begin{aligned} f_W^n(z) &= (1 + uz + vz^2)^n \\ &= \sum_{k_0+k_u+k_v=n} \binom{n}{k_0, k_u, k_v} 1^{k_0} (uz)^{k_u} (vz^2)^{k_v} \\ &= \sum_{k=0}^{n} \sum_{\ell=0}^{n-k} \binom{n}{k} \binom{n-k}{\ell} u^k v^\ell z^{k+2\ell}, \end{aligned}$$

where we have set $k = k_u$ and $\ell = k_v$, and where the condition $k_0 + k_u + k_v = n$ is transformed in $0 \le k \le n, 0 \le \ell \le n - k$. Now, setting $t = k + 2\ell$, we obtain the simplified expression for the generating function

$$f_W^n(z) = \sum_{t=0}^{2n} \sum_{\ell=0}^{\lfloor \frac{t}{2} \rfloor} \binom{n}{t-2\ell} \binom{n-t+2\ell}{\ell} u^{t-2\ell} v^\ell z^t.$$

Lemma 2. *The foregoing implies that the ball of radius e (centered in 0) has volume exactly*

$$|B_{e,D}(0)| = \sum_{t=0}^{e} \sum_{\ell=0}^{\lfloor \frac{t}{2} \rfloor} \binom{n}{t-2\ell} \binom{n-t+2\ell}{\ell} u^{t-2\ell} v^\ell. \tag{1}$$

We thus provided an explicit formula for the cardinality of balls in R^n with respect to the overweight distance.

We now obtain the sphere-packing bound for the overweight distance by combining the previous results. As before, R is a finite ring and $u = |R^\times|$, whereas $v = |R| - 1 - u$ represents the number of non-zero non-units.

Corollary 1 (Sphere-Packing Bound)**.** *Let $C \subseteq R^n$ be a (not necessarily linear) code of length n, and minimum overweight distance $d = 2e + 1$. Then, we have*

$$|C| \le \frac{|R|^n}{|B_{e,D}(0)|},$$

where the cardinality of $B_{e,D}(0)$ is given in Equation (1).

If the minimum distance is even and R is a finite local ring with maximal ideal J, this bound can be adapted as follows.

Corollary 2. *Let R be a local ring with maximal ideal J, $q = |R/J|$ and $C \subseteq R^{n+1}$ be a (not necessarily linear) code of length $n + 1$ and minimum overweight distance $d = 2e + 2$. Then,*

$$|C| \le \frac{|R|^{n+1}}{q|B_{e,D}(0)|},$$

where $B_{e,D}(0)$ is the overweight ball of radius e in R^n, and its volume is given in Equation (1).

Proof. Pick $x_1, \ldots, x_q$ such that the cosets $x_1 + J, \ldots, x_q + J$ form a partition of R. For all $m \in J$, define the set

$$S_m := \{x_1 + m, \ldots, x_q + m\}.$$

Notice that the sets S_m form a partition of R and that all elements of S_m have mutual overweight distance 1. Thus, given $r \in R$, we denote with $S(r)$ the unique set S_m that contains r. Furthermore, let

$$\pi : R^{n+1} \to R^n$$

be the projection that removes the $n+1$'th coordinate and

$$Z(x) := \{z \in R^{n+1} \mid D(\pi(z), \pi(x)) \leq e,\, z_{n+1} \in S(x_{n+1})\}.$$

Now, if $x \neq y \in R^{n+1}$ are two codewords, then $Z(x)$ and $Z(y)$ are disjoint. Indeed, if $z \in Z(x) \cap Z(y)$, then $S(x_{n+1}) = S(y_{n+1})$ as they cannot be disjoint. Hence, $D(x_{n+1}, y_{n+1}) \leq 1$. Furthermore, both $D(\pi(x), \pi(z))$ and $D(\pi(y), \pi(z))$ are less than or equal to e, implying that $D(\pi(x), \pi(y)) \leq 2e$. It follows that $D(x, y) \leq 2e + 1$, which is a contradiction. □

To find non-trivial examples of perfect codes is as notoriously hard as over finite fields in the Hamming metric. Clearly, in the case $R = \mathbb{F}_q$, there are non-trivial perfect codes, as the overweight coincides with the Hamming weight. Examples of such codes can be found in [5] (Section IV). Furthermore, in the case $R = \mathbb{Z}_{p^k}$, linear 1-perfect codes are classified in terms of their parity-check matrix in [6] (Theorem IV.1).

4.3. A Gilbert–Varshamov Bound

With arguments similar to those for the sphere-packing bound, we can also obtain a lower bound on the maximal size of a code with a fixed minimum distance.

Proposition 4 (Gilbert–Varshamov bound)**.** *Let R be a finite ring, n a positive integer and $d \in \{0, \ldots, 2n\}$. Then, there exists a code $C \subseteq R^n$ of minimum overweight distance at least d satisfying*

$$|C| \;\geq\; \frac{|R|^n}{\left|B_{d-1,D}(0)\right|},$$

where the volume is given in Equation (1) *for $e = d - 1$, i.e.,*

$$\left|B_{d-1,D}(0)\right| \;=\; \sum_{t=0}^{d-1} \sum_{\ell=0}^{\lfloor \frac{t}{2} \rfloor} \binom{n}{t-2\ell} \binom{n-t+2\ell}{\ell} u^{t-2\ell} v^{\ell}.$$

Proof. Assume $C \subseteq R^n$ of minimum overweight distance of at least d is a largest code of length n and minimum distance d. Then, the set of balls $B_{d-1,D}(x)$ centered in the codewords $x \in C$ must already cover the space R^n. Since, if this were not the case, one would find an element $y \in R^n$ that is not contained in the ball of radius $d-1$ around any element of C. This word y would have distance at least d to each of the words of C, and thus $C \cup \{y\}$ would be a code of properly larger size with distance at least d, a contradiction to the choice of C.

From the covering argument, we then see that

$$|C| \;\geq\; \frac{|R|^n}{\left|B_{d-1,D}(0)\right|},$$

as desired. □

Let us consider the special case where R is a finite chain ring. Since the overweight is an additive weight, and the conditions of [14] are easily verified, we can use [14] (Theorem 22) to obtain that random linear codes over R^n achieve the (asymptotic) Gilbert–Varshamov bound with high probability.

Example 4. *As an easy example, we can consider $R^n = \mathbb{Z}_8^2$. The maximal minimum overweight distance is given by $d = 2n = 4$. The Gilbert–Varshamov bound states for this example that*

there exists a code C with $|\ C\ | > 1$, as $|\ B_{3,D}(0)\ | = 55$. For example, the code $C = \langle (2,2) \rangle$ has four elements.

4.4. A Plotkin Bound

Over a local ring, we can use methods similar to the ones used for the classical Plotkin bound to obtain an analogue of the Plotkin bound for (not necessarily linear) codes equipped with the overweight.

For the rest of this section, R is a finite local ring with maximal ideal J. The notation stems from the Jacobson radical of the ring R. Note that the factor ring R/J is a finite field, whose cardinality will be denoted by q.

Similarly to the Hamming case, for a subset $A \subseteq R$, we will denote by

$$\overline{W}(A) = \frac{\sum_{a \in A} W(a)}{|A|}$$

the average weight of the subset A.

Lemma 3. *Let $I \subseteq R$ be a left or right ideal. Then,*

$$\overline{W}(I) = \begin{cases} \frac{|R|+|J|-2}{|R|} & \text{if } I = R, \\ 2\left(1 - \frac{1}{|I|}\right) & \text{if } \{0\} \subsetneq I \subsetneq R, \\ 0 & \text{else.} \end{cases}$$

Proof. Note that the last case is trivial as $I = \{0\}$. If $\{0\} \subsetneq I \subsetneq R$, then all non-zero elements of I have weight 2, so this case follows as well.

Finally, if $I = R$, then there are $|R \setminus J| = |R| - |J|$ elements of weight 1 and $|J| - 1$ elements of weight 2. Hence, the total weight is $|R| - |J| + 2(|J| - 1)$ and dividing by $|R|$ yields the claim. □

Corollary 3. *Let R be a local ring with maximal ideal J and assume that $|J| \geq 2$. Then, we have that $\overline{W}(J) \geq \overline{W}(I)$ for all left or right ideals $I \subseteq R$.*

Proof. We immediately see that $\overline{W}(J) \geq \overline{W}(I)$ for all $I \subseteq J$. Now, consider the case $I = R$. We have that

$$\begin{aligned} \overline{W}(R) &= \frac{|R| + |J| - 2}{|R|} = \frac{|R \setminus J|}{|R|} + 2\frac{|J| - 1}{|R|} \\ &= \frac{|R \setminus J|}{|R|} + 2\frac{|J| - 1}{|J|} \cdot \frac{|J|}{|R|} \\ &\leq 2\frac{|J| - 1}{|J|} \cdot \frac{|R \setminus J|}{|R|} + 2\frac{|J| - 1}{|J|} \cdot \frac{|J|}{|R|} \\ &= 2\frac{|J| - 1}{|J|} = \overline{W}(J), \end{aligned}$$

where we used that $2\frac{|J|-1}{|J|} \geq 1$. □

To ease the notation, let us denote by η the following

$$\eta = \overline{W}(J) = 2\left(1 - \frac{1}{|J|}\right).$$

In what follows, we provide a Plotkin bound for the overweight over a local ring R with maximal ideal J. The case $|J| = 1$ is already well studied, since, in this case, R is a field and D is simply the Hamming distance. Hence, we will assume that $|J| \geq 2$.

We start with a lemma for the Hamming weight. The proof of it follows the idea of the classical Plotkin bound, which can be found in [15], and for the homogeneous weight in [10].

Lemma 4. *Let $I \subseteq R$ be a subset and P be a probability distribution on I. Then, we have that*

$$\sum_{x\in I}\sum_{y\in I} w_H(x-y)P(x)P(y) \leq 1 - \frac{1}{|I|}.$$

Proof. We have that

$$\sum_{x\in I}\sum_{y\in I} w_H(x-y)P(x)P(y) = \sum_{x\in I} P(x)(1-P(x)) = \sum_{x\in I} P(x) - \sum_{x\in I} P(x)^2.$$

If we apply the Cauchy–Schwarz inequality to the latter sum, we obtain that

$$\sum_{x\in I} P(x) - \sum_{x\in I} P(x)^2 \leq 1 - \frac{1}{|I|}\left|\sum_{x\in I} P(x)\right|^2 = 1 - \frac{1}{|I|}.$$

From which we can conclude. □

We are now ready for the most important step of the Plotkin bound. As before, R is a local ring with non-zero maximal ideal J and $\eta = \overline{W}(J)$.

Proposition 5. *Let P be a probability distribution on R. Then, it holds that*

$$\sum_{x\in R}\sum_{y\in R} W(x-y)P(x)P(y) \leq \eta.$$

Proof. Let $q = |R/J|$ and pick $x_1, \ldots, x_q$ such that $x_i + J \neq x_j + J$ if $i \neq j$. Then, it follows that the cosets $\overline{x_i} := x_i + J$ form a partition of R. For all $k \in \{1, \ldots, q\}$, we denote by

$$P_k = \sum_{x\in\overline{x_k}} P(x).$$

It follows that $\sum_{k=1}^{q} P_k = 1$. By rewriting the initial sum as sum over all cosets, we obtain that

$$\begin{aligned}
&\sum_{x\in R}\sum_{y\in R} W(x-y)P(x)P(y)\\
&= \sum_{k=1}^{q}\sum_{x\in\overline{x_k}}\sum_{y\in R} W(x-y)P(x)P(y)\\
&= \sum_{k=1}^{q}\sum_{x\in\overline{x_k}}\left(\sum_{y\in\overline{x_k}} 2w_H(x-y)P(x)P(y) + \sum_{z\in R\setminus\overline{x_k}} w_H(x-z)P(x)P(z)\right)\\
&= \sum_{k=1}^{q}\left(2\sum_{x\in\overline{x_k}}\sum_{y\in\overline{x_k}} w_H(x-y)P(x)P(y) + \sum_{x\in\overline{x_k}}\sum_{z\in R\setminus\overline{x_k}} P(x)P(z)\right)\\
&= \sum_{k=1}^{q}\left(2\sum_{x\in\overline{x_k}}\sum_{y\in\overline{x_k}} w_H(x-y)P(x)P(y) + \sum_{x\in\overline{x_k}} P(x)(1-P_k)\right).
\end{aligned}$$

If $P_k \neq 0$, then $\tilde{P}(x) := P(x)/P_k$ defines a probability distribution on $\overline{x_k}$. In this case, we apply Lemma 4 to obtain that

$$\begin{aligned}&\sum_{x\in\overline{x_k}}\sum_{y\in\overline{x_k}} w_H(x-y)P(x)P(y)\\ =&P_k^2\left(\sum_{x\in\overline{x_k}}\sum_{y\in\overline{x_k}} w_H(x-y)\frac{P(x)P(y)}{P_k^2}\right)\\ \leq&P_k^2\left(1-\frac{1}{|J|}\right).\end{aligned}$$

Note that the same inequality also trivially holds if $P_k = 0$. Applying this and using that $\sum_{x\in\overline{x_k}} P(x) = P_k$, we obtain that

$$\begin{aligned}&\sum_{k=1}^{q}\left(2\sum_{x\in\overline{x_k}}\sum_{y\in\overline{x_k}} w_H(x-y)P(x)P(y)+\sum_{x\in\overline{x_k}} P(x)(1-P_k)\right)\\ \leq&\sum_{k=1}^{q}\left(P_k^2\cdot 2\left(1-\frac{1}{|J|}\right)+P_k(1-P_k)\right)\\ \leq&\sum_{k=1}^{q} P_k\cdot 2\left(1-\frac{1}{|J|}\right)=2\left(1-\frac{1}{|J|}\right)=\eta,\end{aligned}$$

where we used that $2\left(1-\frac{1}{|J|}\right)\geq 1$ since $|J|\geq 2$ in the last inequality. □

To complete the Plotkin bound for the overweight, we now follow the steps in [10]. Using Proposition 5, we obtain the following result:

Proposition 6. *Let $C \subseteq R^n$ be a (not necessarily linear) code of minimum overweight distance d. Then,*

$$|C|(|C|-1)d \leq \sum_{x\in C}\sum_{y\in C} D(x,y) \leq |C|^2 n\,\eta.$$

Proof. The first inequality follows since the distance between all distinct pairs of C is at least d.

For the second inequality, let $p_i : R^n \to R$ be the projection onto the ith coordinate. Note that

$$P_i(z) := \frac{|p_i^{-1}(z)\cap C|}{|C|}$$

defines a probability distribution on R for all $i \in \{1,\ldots,n\}$. Using Proposition 5, we obtain that

$$\begin{aligned}\sum_{x\in C}\sum_{y\in C} D(x,y) &= \sum_{i=1}^{n}\sum_{x\in C}\sum_{y\in C} W(x_i-y_i)\\ &= \sum_{i=1}^{n}\sum_{r\in R}\sum_{s\in R} W(r-s)P_i(r)P_i(s)|C|^2\\ &\leq |C|^2\sum_{i=1}^{n}\eta = |C|^2 n\eta.\end{aligned}$$

Thus, we obtain the claim. □

From this inequality, we obtain a Plotkin bound for the overweight distance. As before, R is a local ring with non-zero maximal ideal J and $\eta = 2\left(1-\frac{1}{|J|}\right)$.

Theorem 3 (Plotkin bound for the overweight distance). *Let $C \subseteq R^n$ be a (not necessarily linear) code of minimum overweight distance $d = D(C)$ and assume that $d > n\eta$. Then,*

$$|C| \leq \frac{d}{d - n\eta}.$$

Proof. We divide both sides of the inequality in Proposition 6 by $|C|$ to obtain that

$$|C|(d - n\eta) \leq d.$$

The result then follows from the assumption that $d - n\eta > 0$. □

By rearranging the same inequality, we also obtain the following version of the Plotkin bound, which does not require the assumption that $d > n\eta$.

Corollary 4. *Let $C \subseteq R^n$ be a (not necessarily linear) code with $\mid C \mid \geq 2$ and let $d = D(C)$. Then,*

$$d \leq \frac{|C|n\eta}{|C| - 1}.$$

Proof. We obtain this by dividing both sides of the inequality in Proposition 6 by $|C|(|C| - 1)$, which is non-zero by assumption. □

Remark 2. *Note that W is a homogeneous weight on $\mathbb{Z}_4$, and thus our bound coincides with the bound from [10] for the homogeneous weight on $\mathbb{Z}_4$.*

Example 5. *If we consider codes over $\mathbb{Z}_9$ and fix $|C| = 9$, $n = 3$. We obtain that $d \leq 9/2$ and hence by the integrality that $d \leq 4$. The linear code*

$$C = \langle (1, 1, 3) \rangle$$

attains this bound.

5. A Johnson Bound for the Homogeneous Weight

Another interesting bound is the Johnson bound due to its relation with list-decodability. In the classical form, the Johnson bound gives an upper bound on the largest size $A_q(n, d, w)$ of a constant-weight w code over $\mathbb{F}_q$ of length n and minimum Hamming distance d. However, for the list-decodability of a code, we are interested in codes having codewords of weight *at most* w. In fact, if the largest size of such a code $A'_q(n, d, w)$ is small, e.g., at most a constant L, then every ball of radius w contains at most L codewords and hence one can decode a list of a size at most L. In more detail, the Johnson bound for list-decodability in the Hamming metric states that, if

$$\frac{w}{n} < (1 - \frac{1}{q})\left(1 - \sqrt{1 - \frac{q}{q-1}\delta}\right) = J_q(\delta),$$

where δ denotes the relative minimum distance, then $A'_q(n, d, w) \leq n(d - 1)$.

This famous bound is still missing for the well-studied homogeneous weight, which is, like the overweight, a generalization of the Lee weight over $\mathbb{Z}_4$. In this section, we prove a Johnson bound for the homogeneous weight from Definition 6, denoted by wt and let γ be its average weight on R. By abuse of notation, we denote with wt also the extension of wt to R^n, that is,

$$wt(x) = \sum_{i=1}^{n} wt(x_i).$$

Note that wt does not necessarily satisfy the triangle inequality. In [7] (Theorem 2), it is shown that the homogeneous weight on $\mathbb{Z}_m$ satisfies the triangle inequality if and only if m is not divisible by 6.

We define the ball of radius r with respect to a homogeneous weight wt to be the set of all elements having distance less than or equal to r.

Definition 10. *Let $y \in R^n$ and $r \in \mathbb{R}_{\geq 0}$. The ball $B_{r,wt}(y)$ of radius r centered in y is defined as*

$$B_{r,wt}(y) := \{x \in R^n \mid wt(x-y) \leq r\}.$$

Our aim is to provide a Johnson bound for the homogeneous weight over Frobenius rings. Thus, we begin by defining list-decodability.

Definition 11. *Let R be a finite ring. Given $\rho \in \mathbb{R}_{\geq 0}$, a code $C \subseteq R^n$ is called (ρ, L) list-decodable (with respect to wt) if, for every $y \in R^n$, it holds that*

$$|B_{\rho n,wt}(y) \cap C| \leq L.$$

Over Frobenius rings, the following result holds, which will play an important role in the proof of the Johnson bound.

Proposition 7 ([10] (Corollary 3.3))**.** *Let R be a Frobenius ring, $C \subseteq R^n$ a (not necessarily linear) code of minimum distance d and $\omega = \max\{wt(c) \mid c \in C\}$. If $\omega \leq \gamma n$, then*

$$|C|(|C|-1)d \leq \sum_{x,y \in C} wt(x-y) \leq 2|C|^2\omega - \frac{|C|^2\omega^2}{\gamma n}.$$

With this, we obtain an analogue of the Johnson bound for the homogeneous weight.

Theorem 4. *Let R be a Frobenius ring and $C \subseteq R^n$ be a (not necessarily linear) code of minimum distance d. Assume that $\rho \leq \gamma$. Then, it holds that C is $(\rho, d\gamma n)$ list-decodable if one of the following conditions is satisfied:*

(i) *We have that $\gamma\, n(d - \gamma\, n) \geq 1$.*

(ii) *It holds that $\rho \leq \gamma - \sqrt{(\gamma - \frac{d}{n})\gamma + \frac{1}{n^2}}$.*

Proof. Assume that $e \leq \rho n$ and let $y \in R^n$. We have to show that, under the given conditions, $|B_{e,wt}(y) \cap C| \leq d\gamma n$.

Note first that we may assume that $y = 0$; otherwise, simply consider the translate

$$C' = \{c - y \mid c \in C\}.$$

Assume that $x_1, \ldots, x_N$ are in $B_{e,wt}(0) \cap C$. We have that $wt(x_i - x_j) \geq d$ for $i \neq j$, thus using Proposition 7 and $wt(x-y) = wt(y-x)$, we obtain that

$$N(N-1)d \leq 2\sum_{i<j} wt(x_i - x_j) \leq 2N^2 e - \frac{N^2 e^2}{\gamma n}.$$

Hence, it follows that

$$N(d\gamma n - 2e\gamma n + e^2) \leq d\gamma n.$$

It holds that

$$d\gamma n - 2e\gamma n + e^2 = (n\gamma - e)^2 - n\gamma(n\gamma - d).$$

If we assume that $n\gamma(n\gamma - d) \leq -1$, then we clearly have

$$(n\gamma - e)^2 - n\gamma(n\gamma - d) \geq 1.$$

If this is not the case, we see that $\sqrt{(\gamma - \frac{d}{n})\gamma + \frac{1}{n^2}}$ is well-defined. Thus, if

$$\frac{e}{n} \leq \gamma - \sqrt{\left(\gamma - \frac{d}{n}\right)\gamma + \frac{1}{n^2}},$$

then

$$n\gamma - e \geq \sqrt{(n\gamma - d)n\gamma + 1},$$

and hence

$$(n\gamma - e)^2 - n\gamma(n\gamma - d) \geq 1.$$

It follows that $N \leq d\gamma n$. □

Remark 3. *Note that the second condition already forces $\rho \leq \gamma$.*

Example 6. *As an easy example, we can consider the code $C = \langle (1,1), (4,0) \rangle \subset \mathbb{Z}_8^2$ of minimum homogeneous distance 2 and $\gamma = 1$. The second condition of Theorem 4 is clearly satisfied by choosing $\rho = 1/2$ since*

$$\gamma - \sqrt{\left(\gamma - \frac{d}{n}\right)\gamma + \frac{1}{n^2}} = \frac{1}{2},$$

implying that the code is $(1/2, 4)$ list decodable. For example, when setting $y = (1,2)$, we see that

$$B_{1,wt}(y) \cap C = \{(1,1), (2,2), (1,5), (6,2)\},$$

so the bound is attained.

6. Open Problems

We conclude this paper with some interesting open questions for the newly defined overweight that we have encountered.

Problem 1. *Classify the codes that attain the bounds derived in this paper.*

Problem 2. *Give a Griesmer bound, an Elias-Bassalygo and a Johnson bound for the overweight.*

Proving an analogue of a Griesmer, Elias-Bassalygo and Johnson bound poses a difficult challenge over rings and in particular for the overweight, due to the necessity of an effective upper bound on the sum of the distances.

Author Contributions: All authors contributed to the content of this article. Conceptualization, N.G., M.G., J.R. and V.W.; methodology, N.G., M.G., J.R. and V.W.; validation, N.G., M.G., J.R. and V.W.; formal analysis, N.G., M.G., J.R. and V.W.; investigation, N.G., M.G., J.R. and V.W.; data curation, N.G., M.G., J.R. and V.W.; writing—original draft preparation, N.G., M.G., J.R. and V.W.; writing—review and editing, N.G., M.G., J.R. and V.W.; visualization, N.G., M.G., J.R. and V.W. All authors have read and agreed to the published version of the manuscript.

Funding: The first and third author are supported by armasuisse Science and Technology (Project Nr.: CYD C-2020010) and were supported in part by the Swiss National Science Foundation Grant No. 188430. The fourth author is supported by the Swiss National Science Foundation Grant No. 195290 and by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 899987.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kerdock, A.M. A class of low-rate nonlinear binary codes. *Inf. Control.* **1972**, *20*, 182–187. [CrossRef]
2. Preparata, F.P. A class of optimum nonlinear double-error-correcting codes. *Inf. Control.* **1968**, *13*, 378–400. [CrossRef]
3. Nechaev, A.A. Kerdock's code in cyclic form. *Diskret. Mat.* **1989**, *1*, 123–139. [CrossRef]
4. Hammons, A.R., Jr.; Kumar, P.V.; Calderbank, A.R.; Sloane, N.J.A.; Solé, P. The $\mathbb{Z}_4$-linearity of Kerdock, Preparata, Goethals, and related codes. *IEEE Trans. Inform. Theory* **1994**, *40*, 301–319. [CrossRef]
5. Krotov, D.S. On $\mathbb{Z}_{2^k}$-Dual Binary Codes. *IEEE Trans. Inf. Theory* **2007**, *53*, 1532–1537. [CrossRef]
6. Shi, M.; Wu, R.; Krotov, D.S. On $\mathbb{Z}_{p^k}$-Additive Codes and Their Duality. *IEEE Trans. Inf. Theory* **2019**, *65*, 3841–3847. [CrossRef]
7. Constantinescu, I.; Heise, W. A metric for codes over residue class rings. *Probl. Peredachi Informatsii* **1997**, *33*, 22–28.
8. Greferath, M.; Schmidt, S.E. Gray isometries for finite chain rings and a nonlinear ternary $(36, 3^{12}, 15)$ code. *IEEE Trans. Inform. Theory* **1999**, *45*, 2522–2524. [CrossRef]
9. Nechaev, A.A.; Honold, T. Weighted modules and representations of codes. *Probl. Peredachi Informatsii* **1999**, *35*, 18–39.
10. Greferath, M.; O'Sullivan, M.E. On bounds for codes over Frobenius rings under homogeneous weights. *Discret. Math.* **2004**, *289*, 11–24. [CrossRef]
11. Bachoc, C. Applications of coding theory to the construction of modular lattices. *J. Comb. Theory Ser. A* **1997**, *78*, 92–119. [CrossRef]
12. Dougherty, S.T.; Kim, J.L.; Kulosman, H. MDS codes over finite principal ideal rings. *Des. Codes Cryptogr.* **2009**, *50*, 77–92. [CrossRef]
13. Norton, G.H.; Salagean, A. On the Hamming distance of linear codes over a finite chain ring. *IEEE Trans. Inf. Theory* **2000**, *46*, 1060–1067. [CrossRef]
14. Byrne, E.; Horlemann, A.L.; Khathuria, K.; Weger, V. Density of Free Modules over Finite Chain Rings. *arXiv* **2021**, arXiv:2106.09403.
15. van Lint, J. *Introduction to Coding Theory*; Graduate Texts in Mathematics; Springer: Berlin/Heidelberg, Germany, 1982.