Metrics Based on Finite Directed Graphs and Coding Invariants

Tuvi Etzion\textsuperscript{*}, Marcelo Firer\textsuperscript{+}, Roberto Assis Machado\textsuperscript{+}

\textsuperscript{*}Computer Science Department, Technion, Israel Institute of Technology, Haifa 32000, Israel

\textsuperscript{+}Department of Mathematics, Unicamp, Campinas, Brasil

etzion@cs.technion.ac.il, mfirer@ime.unicamp.br, robertomachado@ime.unicamp.br

Index Terms

Coding theory, MacWilliams’ type identity, maximum likelihood decoding, graph metrics.

Abstract

Given a finite directed graph with \( n \) vertices, we define a metric \( d_G \) on \( \mathbb{F}_q^n \), where \( \mathbb{F}_q \) is the finite field with \( q \) elements. The weight of a word is defined as the number of vertices that can be reached by a directed path starting at the support of the vector. Two canonical forms, which do not affect the metric, are given to each graph. Based on these forms we characterize each such metric. We further use these forms to prove that two graphs with different canonical forms yield different metrics. Efficient algorithms to check if a set of metric weights define a metric based on a graph are given. We provide tight bounds on the number of metric weights required to reconstruct the metric. Furthermore, we give a complete description of the group of linear isometries of the graph metrics and a characterization of the graphs for which every linear code admits a \( G \)-canonical decomposition. Considering those graphs, we are able to derive an expression of the packing radius of linear codes in such metric spaces. Finally, given a directed graph which determines a hierarchical poset, we present sufficient and necessary conditions to ensure the validity of the MacWilliams Identity and the MacWilliams Extension Property.

I. Introduction

Let \( \mathbb{F}_q^n \) be an \( n \)-dimensional vector space over a finite field \( \mathbb{F}_q \) and let a code be a linear subspace over \( \mathbb{F}_q^n \). In order to gain some tools that may help (mainly in the decoding processes), different metric structures are considered. The most common one, namely the Hamming metric, shares two well known properties:

1) It is determined by a weight, in the sense that \( d(x, y) = w(x - y) \).

2) If we consider two error vectors \( x = (x_1, \ldots, x_n) \) and \( x' = (x'_1, \ldots, x'_n) \) such that \( x_i \neq 0 \) if \( x'_i \neq 0 \), then \( w(x) \leq w(x') \).

This paper was presented in part at IEEE International Symposium on Information Theory 2016 (ISIT 2016) and at IEEE Information Theory Workshop 2016 (ITW 2016).
Both properties are crucial in the context of coding theory. The first one means that the distance \( d(\cdot, \cdot) \) is invariant by translations \( (d(x + z, y + z) = d(x, y)) \), for all \( x, y, z \in \mathbb{F}_q^n \), what turns linear codes to be more manageable in what concerns metric properties. To be more precise, as an example, the well known Syndrome decoding algorithm performs a minimal distance decoding procedure if and only if the metric is invariant by translations. The second property says that the metric is suitable for error detection and correction when considering reasonable channels, in the sense that making errors in a family \( I \) of bits can not be more probable then making errors in both the families \( I \) and \( J \).

Two large (and nearly disjoint) families of metrics on \( \mathbb{F}_q^n \) which satisfy those properties were introduced over the years: the combinatorial metric, introduced by Gabidulin in 1971 [8] and the poset metrics, introduced in the context of Coding Theory by Brualdi et al. [5] in 1995.

Since then, the poset metrics gained attention and has been extensively studied in the literature in the context of coding theory, exploring all its major aspects and invariants: MacWilliams Identity ([6], [12]), perfect and MDS codes [3], [10], [11], duality, packing, and covering problems [13], [15], [16]. Some generalizations of poset metrics were introduced over the years, including its extension to Frobenius and finite principal ideal rings [4],[9] and the poset-block metrics [1].

In this work we introduce a new family of metrics: given a directed graph \( G \), on \( n \) vertices, we define a metric \( d_G \) on \( \mathbb{F}_q^n \). The graph based metrics generalize the poset metrics. It is closely related to the poset-block metrics (introduced in [1]) but it is actually a new family of metrics. As we shall see in Section III, the poset-block metrics may be obtained as a variation of the graph based metrics. In terms of coding theory, in some sense they work on different directions. If we consider, as an example, a situation where we have two blocks of information, blocks with different sizes (in bits), the poset-block values only the number of blocks where errors occur, but not the size of the blocks. The graph based metrics refine the poset-block metrics by considering the errors that compromise larger blocks to be more relevant than errors that damage the information on smaller blocks.

Metrics defined by a graph may be useful to model some specific kind of channels and used to perform bitwise or message wise unequal error protection. The goal of this paper is to study such metrics in this context, and is organized as follows:

In Section II we present the basic definitions of directed graphs and metrics defined by graphs. In Section III, two canonical forms, of directed graphs (for construction of the related metrics), are defined – the expanded canonical form and the reduced canonical form. We prove that these two canonical forms are unique, i.e. for a given directed graph there is a unique expanded canonical form and a unique reduced canonical form. Moreover, we prove that two directed graphs determine the same metric if they have the same canonical forms. In Section IV we consider several computational questions related to the minimum number of metric weights which are required to reconstruct the whole metric. These questions have several variants which will be discussed in this section. Bounds, some of which are tight are presented. From here...
on we move to questions closely related to coding theory. In Section V we discuss the group of linear isometries of the graph metrics and the connection of the work to coding theory. The structure of this group is a tool, which is used in Section VI to determine an interesting decomposition of linear codes - the G-canonical decomposition - in the case when the canonical reduced form of a graph is a hierarchical poset. Finally, in Section VII we present some conditions on a graph that are sufficient for the validity of MacWilliams Identity and the MacWilliams Extension Property.

II. Basic Concepts

A (simple finite) directed graph $G(V,E)$ consists of a finite set of vertices $V = \{v_1, \ldots, v_n\}$ and a set of directed edges $E$ (parallel edges are not allowed), where $e \in E$ is an ordered pair $(u,v) \in V \times V$ with $u \neq v$. The vertex $u$ in this edge is called the tail of the edge and the vertex $v$ is called the head of the edge. A trail of length $k$ is a sequence of edges $(u_0,u_1),(u_1,u_2),\ldots,(u_{k-1},u_k)$ in which all edges are distinct. When $u_0 = u_k$, the trail is called a circuit. In case all the vertices $u_i$'s in a trail are distinct, except for the possibility that $u_0 = u_k$, the trail is called a simple directed path of length $k$. If $u_0 = u_k$ then the path is called a directed cycle.

A complete graph is a graph which contains all the possible $n(n-1)$ edges. Such a graph is also called a clique.

If there is a trail from $u$ to $v$, we say that $u$ dominates $v$, and denote it by $u \to v$. A set $X \subset V$ is called a closed set if $u \in X$ and $u$ dominates $v \in V$ implies that $v \in X$. The closure $\langle X \rangle_G$ of a set $X \subset V$ is the smallest closed subset containing $X$. The set of all closed sets of $G$ is denoted by $\mathcal{I}(G)$. By abuse of notation, if $X = \{v\}$, we denote $\langle \{v\} \rangle_G = \langle v \rangle_G$.

A directed graph $G(V,E)$ will be called L-weighted (denoted by $G(V,E,L)$) if $G(V,E)$ is a directed graph and there is a function $L : V \to \mathbb{N}$, where $\mathbb{N}$ denotes the set of natural numbers (positive integers). The value $L(v)$, $v \in V$, is called the $L$-weight of $v$. Clearly, if for each $v \in V$ we have that $L(v) = 1$ then we can omit the $L$-weights of the vertices, and the L-weighted directed graph is just a directed graph.

Each directed graph $G(V,E)$, where $V = \{v_1, \ldots, v_n\}$, defines a metric $d_G$ as follows: The words of the space are all $n$-tuples over a given alphabet whose size is at least two. Given $x = (x_1,x_2,\ldots,x_n) \in \mathbb{F}_q^n$, the support of $x$ is the set of non-zero coordinates of $x$, i.e. $\text{supp}(x) = \{i; x_i \neq 0\}$. When considering an ordering (arbitrary but fixed) in the set $V = \{v_1,v_2,\ldots, v_n\}$ of vertices of $G$, we have the G-support of $x$, defined as $\text{supp}_G(x) = \{v_i \in V; x_i \neq 0\}$. For simplicity, we shall write $\text{supp}_G(x) = \text{supp}(x)$. We can naturally identify $\text{supp}(x)$ with $[i \in [n]; x_i \neq 0]$ and we shall do so if no confusion may arise. The G-weight, $w_G(x)$, of a word $x$ is the number of distinct vertices in $G$ dominated by the vertices in the support of $x$:

$$w_G(x) = |\langle \text{supp}(x) \rangle_G|,$$

where $|A|$ denotes the cardinality of the set $A$. The $G$-distance between two words $x, y \in \mathbb{F}_q^n$ is defined by $d_G(x,y) = w_G(y - x)$. Since the actual values in the nonzero entries of $x$ and $y$ do not play any role in the
determination of the G-weights, we will assume, without loss of generality, that all the G-weights are given 
only for binary words. 

Theorem 1: If G(V,E) is a directed graph, then d_G is a metric. 

Example 1: The Hamming weight w_H is a particular case of a G-weight where E = 0. On the opposite side, 
if G is a complete graph, then ⟨A⟩G = V for every A ≠ 0 and hence, w_G(x) = n, for each x ≠ 0. 

III. Canonical Forms 

In this section two canonical forms will be introduced for any given finite directed graph G(V,E), one 
reduced canonical form and the other expanded canonical form. Each canonical form can be served to 
characterize a set of graphs which form the same G-metric. These canonical forms are unique and will lead 
to one of the main results of this paper, that two different graphs whose expanded or reduced canonical 
forms are different yield two different metrics. The canonical forms can be defined through a set of edges 
called shortcuts. An edge (u,v) ∈ E is called a shortcut if there exists a simple directed path from u to v 
which contains at least two edges. Adding or removing shortcuts to the graph do not affect the metric as 
the following lemma, whose proof is obvious, states. 

Lemma 1: If G(V,E) and G'(V,E') are two directed graphs which differ in exactly one edge e, and e is a 
shortcut, then d_G = d_G'. 

We remark that shortcuts must be removed one at a time and not simultaneously, in order to avoid the 
situation which we find in the next example: 

Example 2: Let G(V,E) be a graph with V = {u,v,w} and E = {(u,v),(u,w),(v,w),(w,v)}. Then (u,v) and 
(u,w) are shortcuts. By removing any of those shortcuts or all of them we get the sets of edges E_v = 
{(u,w),(v,w),(w,v)}, E_w = {(u,v),(v,w),(w,v)} and E_v,w = {(v,w),(w,v)}, to which corresponds the graphs G_v, 
G_w and G_v,w, respectively. One can directly check that 

⟨u⟩G = ⟨u⟩G_v = ⟨u⟩G_w = {u,v,w} 

⟨u⟩G_v,w = {u}. 

The expanded canonical form of a directed graph G(V,E) is the graph G'(V,E') obtained by adding edges, one 
by one, in a way that each edge added to the graph is a shortcut. Hence, the expanded canonical graph 
G'(V,E') is a graph for which E ⊂ E', E' \ E contains only shortcuts and any possible edge to add in E' is 
not a shortcut. 

As an example, the expanded canonical form of a graph is a complete graph if and only if there is 
a circuit containing all the vertices. The expanded canonical form of a graph G(V,E) has all the possible 
shortcuts. Some of these induce complete subgraphs of the expanded canonical form G'(V,E'). To figure out 
the maximal cliques in the expanded canonical form we state the next elementary lemmas, which follow 
straightforward from the definition of G'(V,E').
Lemma 2: Let $G(V,E)$ be a directed graph and $G'(V,E')$ its expanded canonical form. Let $V' \subset V$ be such that $u$ dominates $v$ for any $u,v \in V'$. Then $G'(V,E')$ induces a clique on $V'$.

The clique induced on $V'$, by the expanded canonical form $G'(V,E')$, is called maximal if there is no $v \in V \setminus V'$ such that $G'(V,E')$ induces a clique on $V' \cup \{v\}$. It is clear that if $V'$ and $V''$ are maximal cliques, then either $V' = V''$ or $V' \cap V'' = \emptyset$. Finally, we have that the expanded canonical form is unique.

Lemma 3: Any directed graph has a unique expanded canonical form. In other words, the expanded canonical form is well-defined.

Theorem 2: Two different graphs with the same expanded canonical form yield the same metric.

Proof: Let $G(V,E)$ be a graph and $G'(V,E')$ its canonical expanded form. Since we can move from the expanded graph $G'$ to the original graph $G$ by removing one by one the shortcuts in $E' \setminus E$, Lemma 1 ensures that $d_G = d_{G'}$.

Next, we want to define a second canonical form for a graph which defines a metric. Succinctly, the reduced canonical form is an $L$-weighted acyclic graph $G'(V',E',L')$ obtained from a graph $G(V,E)$ by contracting each strongly-connected component to a vertex, and the weight of this vertex is the number of vertices in the strongly-connected component, i.e., assuming that $G(V,E)$ is in the expanded canonical form (Lemma 3, we obtain $G'(V',E',L')$ in the following way:

1) Each maximal clique in $G(V,E)$ is associated to a vertex $u$ in $V'$. We denote by $\pi : V \to V'$ the map that associates to a vertex of $V$ the vertex of $V'$ that represents its clique.
2) The weight $L' : V' \to \mathbb{N}$ is defined as $L'(u) = |\pi^{-1}(u)|$.
3) The projection $\pi$ also defines the structure of edges: we start defining a set of vertices $E'$ where $(u_1,u_2) \in E'$ if, and only if, there are $v_1 \in \pi^{-1}(u_1)$ and $v_2 \in \pi^{-1}(u_2)$ such that $v_2 \in \langle v_1 \rangle_G$. We remark that $G' = (V',E')$ has no circuits, so that it actually determines a partial order on $V'$: $u_1 \leq u_2$ if $u_1 \in \langle u_2 \rangle_G$.

We denote this poset by $P_G = (V',\leq)$. It follows that, on $G'$, the closure of a set $A$ is just the ideal generated by $A$ in the usual meaning in the context of posets.
4) Finally, we obtain $E'$ by removing all shortcuts from $E'$ and get the canonical reduced form $G'(V',E',L')$.

We remark the following:

- The reduced closure of a subset $A \subset V$ is the $G'$ closure under the projection $\pi$, i.e., $\langle A \rangle_{G'E'} := \langle \pi(A) \rangle_{G'}$. By doing so, given $x \in \mathbb{F}_q^n$ we can determine its $G$-weight by considering the $L$-weighted reduced form of $G$:

$$w_G(x) = \sum_{u \in \text{supp}(x)G'} L(u).$$

We remark that, if we define a constant weight $L_1(u) = 1$, for all $u \in V'$, we get the poset-block metric as defined in [1].
There may be many shortcuts with tail $u$ and the heads $\{u_1, u_2, \ldots, u_r\}$ contained in a single clique. In the reduced form, all these edges are replaced by a single edge with tail $u$ and head in the vertex that replaces the clique.

Fig. 1. On the left a directed graph with 5 vertices and one single clique $\{u_4, u_5\}$. On the right, its reduced canonical form. The clique is replaced by a vertex $v$ with $L'(v) = 2$. Whenever we removed the shortcut $s_1$ or $s_2$ is immaterial, since they are replaced by a single directed edge $(u_1, v)$.

**Proposition 1:** Any directed graph has a unique reduced canonical form, i.e., it is well-defined.

*Proof:* The structure of maximal cliques in $G$ is unique and so the graph $G'$ is uniquely determined by $G$. Moreover, the $L'$-weight is determined exclusively by the projection $\pi : V \rightarrow V'$, hence it is also uniquely determined. $G'$ is a poset and $G'$ may be viewed as its Hasse diagram. It is well known that the Hasse diagram (considered as a directed graph) of a poset is uniquely determined by the poset and it follows that the reduced canonical form is well-defined.

Proposition 1 leads to the two main results of this section. Before we state and prove them, we need a simple and useful lemma:

**Lemma 4:** Let $G_1(V, E_1)$ and $G_2(V, E_2)$ be two graphs such that $d_{G_1} = d_{G_2}$. Then, given $u, v \in V$, $v \in \langle u \rangle_{G_1}$ if and only if $v \in \langle u \rangle_{G_2}$.

*Proof:* Suppose that $v \in \langle u \rangle_{G_1}$ and $v \notin \langle u \rangle_{G_2}$. This would imply that $\langle u \rangle_{G_1} = \langle u, v \rangle_{G_1}$ while $\langle u \rangle_{G_2} \not= \langle \{v, u\} \rangle_{G_2}$ and hence a word $x \in \mathbb{F}_q^u$ such that $\text{supp}(x) = \{u, v\}$ will have different weights: $w_{G_1}(x) < w_{G_2}(x)$, contradicting the assumption that $d_{G_1} = d_{G_2}$.

**Theorem 3:** If two directed graphs $G_1(V, E_1)$ and $G_2(V, E_2)$ induces metrics such that $d_{G_1} = d_{G_2}$ then the expanded canonical forms of $G_1$ and $G_2$ are equal, and the same is true for their reduced canonical forms.

*Proof:* The proof is done by induction on $n = |V|$. The proof of the base $n = 1$ is trivial. Let $0 \neq e_v \in \mathbb{F}_q^u$ be a word with $\text{supp}(e_v) = \{v\}$ and with minimal $w_{G_1}$ and $w_{G_2}$ weight. If $w_{G_1}(e_v) = w_{G_2}(e_v) = n$, then both $G_1$ and $G_2$ contain a cycle of length $n$ hence they expanded canonical form is a complete graph with $n$ vertices and the reduced canonical form is the graph with a unique vertex $u$ and $L_1(u) = L_2(u) = n$.

We assume that $w_{G_i}(e_v) < n$. We consider the projections $u_1 = \pi_1(v)$ and $u_2 = \pi_2(v)$ where $\pi_i : G_i \rightarrow G'_i$ is the projection onto the canonical reduced form, for $i = 1, 2$. The minimality of $w_{G_i}(v)$ ensures that $u_i$ is a minimal element in the poset $G'_i$ and $w_{G_i}(v) = L_i(u_i) = |\pi_i^{-1}(u_i)|$, for $i = 1, 2$. It follows that the maximal cycle
containing \( v \) in \( G_1 \) and \( G_2 \) \((\pi_1^{-1}(u_1) \) and \( \pi_2^{-1}(u_2) \), respectively) have the same cardinality and if \((v,w) \in E_i\), then \( w \in \pi_i^{-1}(u_i) \). We consider now the graphs \( G'_i \) obtained from \( G_i \) by adding all the possible edges to the maximal cycle \( \pi_i^{-1}(u_i) \). It is clear that \( d_{G_i} = d_{G'_i} \), for \( i = 1,2 \). Let us consider the graph \( \Gamma_i \) obtained from \( G'_i \) by removing the vertex \( v \) and all the edges of \( G'_i \) that have \( v \) either as a tail or as a head. It follows that

\[
|\langle w \rangle_{\Gamma_i}| = \begin{cases} 
|\langle w \rangle_{\Gamma_i}| & \text{if } v \notin \langle w \rangle_{\Gamma_i} \\
|\langle w \rangle_{\Gamma_i}| - 1 & \text{if } v \in \langle w \rangle_{\Gamma_i}.
\end{cases}
\]

Lemma 4 ensures that \( v \in \langle w \rangle_{\Gamma_i} \) if and only if \( v \in \langle w \rangle_{G_i} \) and it follows that \( d_{\Gamma_i} = d_{\Gamma_2} \). By the induction hypothesis, we may assume that the reduced and the extended canonical forms of \( \Gamma_1 \) and \( \Gamma_2 \) are the same.

Let \( \Gamma'_1 = \Gamma_2' = (V',E',L') \) be the reduced form of both \( \Gamma_1 \) and \( \Gamma_2 \). If \( w_{G_i}(v) > 1 \) then the canonical reduced form of \( G_i \) is the \( L \)-weighted graph \((V',E',L')\), which differ from \( \Gamma'_i \) only in the \( L \)-weight:

\[
L(u) = \begin{cases} 
L'(u) & \text{if } u \in \pi_i(v) \\
L'(u) + 1 & \text{if } u = \pi_i(v).
\end{cases}
\]

If \( w_{G_i}(v) = 1 \), we have that canonical reduced form of \( G_i \) is the \( L \)-weighted graph \( G''_i = (V''_i,E''_i,L''_i) \) where

\[
V''_i = V' \cup \{u_i\},
\]

\[
E''_i = E' \cup \{(\pi_i(w),u_i) ; v \in \langle w \rangle_{\Gamma_i}\},
\]

\[
L''_i(u) = \begin{cases} 
L'(u) & \text{for } u \neq u_i \\
1 & \text{for } u = u_i.
\end{cases}
\]

recalling that \( u_i = \pi_i(v) \). Again, Lemma 4 ensures \( v \in \langle w \rangle_{\Gamma_1} \) if, and only if, \( v \in \langle w \rangle_{G_2} \) and it follows that, up to renaming \( u = u_1 = u_2 \) we have that \( E''_1 = E''_2 \) hence \( G''_1 = G''_2 \).

For the extended canonical form we proceed in exactly the same way, making the induction step by excluding a vertex \( v \in V \) such that \( |\langle v \rangle_{\Gamma_i}| = |\langle v \rangle_{\Gamma_2}| \) is minimal.

**Corollary 1:** Two graph metrics are isomorphic if and only if their related graphs in expanded or reduced canonical forms are isomorphic.

**Proof:** It follows from Theorem 3, Theorem 2 and from the fact that

\[
w_G(x) = \sum_{u \in \text{supp}(x)\cap\Gamma} L(u).
\]

**IV. Metric Reconstruction**

In this section we will consider the minimum number of \( G \)-weights which are required to reconstruct the whole metric. This question has at least three variants.

1) Let \( S(n) \) be a set of words of length \( n \) such that for any given metric based on a directed graph with \( n \) vertices, the \( G \)-weights of the words in \( S(n) \) are sufficient to recover the whole metric. Let \( D(n) \) be the minimum size of \( S(n) \). What is the value of \( D(n) \)?
2) Let $d_G$ be a metric based on a directed graph with $n$ vertices. Let $M_G(n)$ be minimum number of G-weights required to recover the whole metric. Let $G$ be the expanded canonical graph for which this number is the smallest among all graphs, and let $M^{\text{min}}(n)$ be the value of $M_G(n)$ for this graph. Let $G$ be the expanded canonical graph for which this number is the largest among all graphs, and let $M^{\text{max}}(n)$ be the value of $M_G(n)$ for this graph. What is the value of $M^{\text{min}}(n)$? What is the value of $M^{\text{max}}(n)$?

3) Given a metric $d_G$, we can ask queries, each query on the G-weight value of a specific word, in a sequence, where each query is based on the values given in previous answers that we got. Let $Q(n)$ be the minimum number of queries which are required to recover any given such metric. What is the value of $Q(n)$?

We note on the difference between the second and the third variants. For the second variant we assume that for the given metric we receive the smallest number of metric weights to recover all the G-weights. In the third variant we have to find a general strategy to ask queries in a way that the total number of queries will be small.

The following lemma is readily verified.

Lemma 5: $M^{\text{min}}(n) \leq M^{\text{max}}(n) \leq Q(n) \leq D(n)$.

We start with the fundamental question of constructing the graph from the G-weights of the words in the space. Given a set of all G-weights we would like to find the related graph, i.e. its reduced canonical form or its expanded canonical form. To this end we first need the following lemma which is verified from the definitions.

Lemma 6: For a given graph metric $d_G$, of a graph $G(V,E)$, $w_G(e_i + e_j) = w_G(e_i)$, $i \neq j$, if and only if there exists a path from $v_i$ to $v_j$, i.e. $v_i$ dominates $v_j$.

Theorem 4: The G-weights of all words with Hamming weights one or two are sufficient to determine the expanded canonical form of the related graph.

Proof: By Lemma 6, we can determine for each pair of vertices $\{v_i, v_j\}$, $i \neq j$, if the edges $v_i \rightarrow v_j$ and $v_j \rightarrow v_i$ exist or not, by observing G-weights only of words with Hamming weight one or two. □

Corollary 2: If the graph $G$ with the metric $d_G$ has $n$ vertices, then the $n + \binom{n}{2}$ G-weights of words with Hamming weights one or two are sufficient to determine the whole metric.

Corollary 3: $D(n) \leq n + \binom{n}{2}$.

On the other hand it is not difficult to see that there are metrics in which all the weights, except for one, are not sufficient to determine the whole metric. Consider the following two graphs $G_1(V, E_1)$ and $G_2(V, E_2)$, where $V = \{v_1, \ldots, v_n\}$. Assume that $G_1$ is the complete graph, while in $G_2$, $v_1, \ldots, v_{n-1}$ form a clique on $n-1$ vertices, and there is an edge $v_i \rightarrow v_n$, for each $i$, $1 \leq i \leq n-1$. The G-weights of all words, except one word $e_n$, in these two metrics are the same ($w_{G_1}(x) = w_{G_2}(x) = n$ for each word $x \in \mathbb{F}_q^n$, except for $w(G_1)e_n = n$ and $w_{G_2}(e_n) = 1$).
Corollary 4: If $d_G$ is a metric based on a graph $G$, then each one of the $G$-weights of words with Hamming weight one might be required to determine the whole metric, i.e. $e_i \in S(n)$ for each $1 \leq i \leq n$.

Corollary 3 can be slightly improved as follows.

Theorem 5: $D(n) \leq \lceil \frac{n}{2} \rceil + \binom{n}{2}$.

Proof: Let $V = \{v_1, \ldots, v_n\}$ be the set of vertices of $G$. We claim that the $G$-weights of all the words of Hamming weight one and the $G$-weights of all the words with Hamming weights two, excluding $w_G(e_1+e_2)$, $w_G(e_3+e_4)$, $w_G(e_5+e_6)$, ..., are sufficient to recover the $G$-weights of the whole metric. To prove this we only have to show that we can find the weights of $w_G(e_1+e_2)$, $w_G(e_3+e_4)$, $w_G(e_5+e_6)$, ..., and apply Corollary 2. By Lemma 6, we have that by observing the set $\{w_G(e_1+e_i) : 3 \leq i \leq n\}$ we can find how many (and which) vertices (excluding $v_2$) are dominated by $v_1$. Let $\delta_1$ be the number of vertices that $v_1$ dominates (including $v_1$ and excluding $v_2$). If $w(e_1) = \delta_1$ then clearly $v_1$ does not dominate $v_2$ and if $w_G(e_1) = \delta_1 + 1$ then clearly $v_1$ dominates $v_2$. Similarly, let $\delta_2$ be the number of vertices that $v_2$ dominate (including $v_2$ and excluding $v_1$) and similarly we can find if $v_2$ dominates $v_1$ or not. Now, if $v_1$ dominates $v_2$ then $w_G(e_1+e_2) = w_G(e_1)$ and if $v_2$ dominates $v_1$ then $w_G(e_1+e_2) = w_G(e_2)$. Finally, if $v_1$ does not dominate $v_2$ and $v_2$ does not dominate $v_1$ then $w_G(e_1+e_2) = \delta_1 + \delta_2 - t$, where $t$ is the number of vertices dominated by both $v_1$ and $v_2$ (and $t$ can be found by observing the vertices dominated by $v_1$ and observing the vertices dominated by $v_2$). Similarly, we can find $w_G(e_3+e_4)$ and so on. To conclude, we apply Corollary 2 to complete the proof.

Theorem 6: Given any set with less than $\lceil \frac{n}{2} \rceil + \binom{n}{2}$ $G$-weights, there exists a directed graph $G$ with $n$ vertices whose graph metric cannot be reconstructed only from these $G$-weights.

Proof: Consider the following two directed graphs $G_1(V, E_1)$ and $G_2(V, E_2)$, where $V = \{v_1, \ldots, v_n\}$. $G_1$ has a clique on the set of vertices $\{v_4, \ldots, v_n\}$; for each $i$, $4 \leq i \leq n$, $G_1$ has the edge $v_i \rightarrow v_j$, $1 \leq j \leq 3$, and it also has the edge $v_1 \rightarrow v_2$. $G_2$ has a clique on the set of vertices $\{v_4, \ldots, v_n\}$; for each $i$, $4 \leq i \leq n$, $G_2$ has the edge $v_i \rightarrow v_j$, $1 \leq j \leq 3$, and it also has the edge $v_1 \rightarrow v_3$. This means that the two graphs differ only in the edges $v_1 \rightarrow v_2$ and $v_1 \rightarrow v_3$. The related metrics differ only in two $G$-weights. In $G_1$, $w_{G_1}(e_1 + e_2) = 2$ and $w_{G_1}(e_1 + e_3) = 3$, while in $G_2$, $w_{G_2}(e_1 + e_2) = 3$ and $w_{G_2}(e_1 + e_3) = 2$. Hence, if these two $G$-weights are not given we won’t be able to distinguish between the two metrics. Therefore, from the $G$-weight of words with Hamming weight two only a set of $G$-weights which correspond to vertex-disjoint edges can be omitted. Thus, from all the $G$-weights of words with Hamming weight two we are required to have at least $\binom{n}{2} - \lceil \frac{n}{2} \rceil$ words. By Corollary 4, the $G$-weights of all the words with Hamming weight one are also required and the lemma follows.

Corollary 5: $D(n) \geq \lceil \frac{n}{2} \rceil + \binom{n}{2}$.

Now, by Theorem 5 and Corollary 5 we have that

Corollary 6: $D(n) = \lceil \frac{n}{2} \rceil + \binom{n}{2}$.

The values of $M^{\text{min}}(n)$ and $M^{\text{max}}(n)$ given in the next two theorems.

Theorem 7: $M^{\text{min}}(n) = n$. 

May 2, 2017 DRAFT
Proof: Given the Hamming metric which is represented by a graph $G$ with $n$ isolated vertices, it is trivial to see that the set of weights $\{w_G(v_i) : i \in [n]\}$ is sufficient to determine the metric. Hence, $M^{\text{min}}(n) \leq n$. $M^{\text{min}}(n) \geq n$ from information theory arguments.

Lemma 7: $M^{\text{max}}(n) \geq 2n - 4.$

Proof: If $n = 1$ then no queries are required and if $n = 2$ then two queries are necessary. For $n = 3$ the graph with exactly one edge requires four queries.

Now, we have to present a graph with $n \geq 4$ vertices that requires at least $2n - 4$ queries to reveal all its edges. The graph will consists of at least four vertices. One vertex $v_{\text{max}}$ dominates all the other vertices. One vertex $v_{\text{min}}$ which does not dominate any other vertex. Any other vertex of the $n - 2$ vertices dominates $v_{\text{min}}$. We must know $w_G(v_{\text{max}})$ to reveal it dominates all the other vertices. For any vertex $v$ which dominates $v_{\text{min}}$ we must know $w_G(v)$, but this is not enough. For example, assume that $w_G(v_1) = w_G(v_2) = 2$. If we have no other information on $v_1$ and $v_2$ they can either dominate $v_{\text{min}}$ or be together in a clique of size two. Hence, we need two queries for each one of the vertices that dominates $v_{\text{min}}$ with a possible exception of one. Revealing all this information implies that $w_G(v_{\text{min}}) = 1$, so there is no need to have a query on $v_{\text{min}}$. Thus, the total number of required queries is $2n - 4$.

Lemma 8: $M^{\text{max}}(n) \leq 2n - 1.$

Proof: For a given vertex $u$, that dominates the vertices $v_1, v_2, \ldots, v_r$, the two G-weights $w_G(u)$ and $w_G(u,v_1,v_2,\ldots,v_r)$, which are equal to $r + 1$ are enough to know all the edges from $u$ to the other vertices of the graph. This proves that $M^{\text{max}}(n) \leq 2n - 2$. Hence, we only have to show that we can omit one G-weight. For any vertex $u$ which does not dominate any other vertex, the G-weight $w_G(v)$ is sufficient and if there exists one such vertex then the claim of the lemma is proved. If there is no such vertex, then the vertex of minimal G-weight is in a clique of size $r \geq 2$, which consists of the vertices $v_1, v_2, \ldots, v_r$. The $r + 1$ G-weights $w_G(v_1), w_G(v_2), \ldots, w_G(v_r)$, and $w_G(v_1,v_2,\ldots,v_r)$ are sufficient to find this clique and hence $2n - 1$ G-weights are sufficient.

Thus, $M^{\text{max}}(n) \leq 2n - 1.$

The gap between the lower bound of Lemma 7 and the upper bound of Lemma 8 can be reduced, but we omit it as the proof is too tedious and we leave it to the reader. For the value of $Q(n)$, we managed only slightly to bridge on the gap between the upper and lower bounds derived from Lemma 5. Hence, the intriguing problem to determine the value of $Q(n)$ is left as a research problem for the reader.

After the expanded canonical form graph was constructed based on the G-weights of the words with Hamming weights one and two, we would like to know whether the given G-weights of the other words are consistent with the G-weights of the words with Hamming weights one or two. The most simple way is to consider each $r$-subsets of vertices and use a search algorithm for all the vertices dominated by this $r$-subset of vertices. We leave other variants to the interested reader and for future research (hs). Some of the variants are related to the well-known question concerning graph isomorphism, until recently the only
fundamental computational problem whose answer is completely unknown [18] (before the most recent result of László Babai [2]).

V. ISOMETRIES AND CODING

As seen in Section III, the reduced canonical form $G'(V', E', L')$ of $G$ may be considered as a poset $P_G = (V', \leq)$, and we say that $v \in X$ is maximal in $X \subset V$ if $\pi(v)$ is maximal in $\pi(X) \subset V'$, considered as a subposet of $P_G$. We denote by $\text{Max}_G(X)$ the set of maximal elements of $X$ and $\text{Max}_G(x) := \text{Max}_G(\text{supp}(x))$. The cleared out form of $x \in \mathbb{F}_q^n$ is the vector $\tilde{x} = (\tilde{x}_1, \ldots, \tilde{x}_n)$ such that $\tilde{x}_i = x_i$ if $i \in \text{Max}_G(x)$ and $\tilde{x}_i = 0$ otherwise. A subset $\tilde{X} \subset X \subset V$ is said to be a minimal set of generators (MSG) of $X$ if $\langle \tilde{X} \rangle = \langle X \rangle_G$ and $\tilde{X}$ is minimal with this property. A MSG is not unique, but it is easy to see that $\tilde{X} \subset \text{Max}_G(X)$. A minimal set of generators (MSG) of $X \subset V$ is a subset $\tilde{X} \subset X$ such that $\langle \tilde{X} \rangle = \langle X \rangle_G$ and $\tilde{X}$ is minimal with this property. A MSG is not unique, but it is easy to see that $\tilde{X} \subset \text{Max}_G(X)$.

Given a directed graph $G$ we denote by $\widetilde{G}$ its expanded canonical form of $G$, $\text{Aut} (\widetilde{G})$ the group of automorphisms of $\widetilde{G}$ and $GL(n, G)_q$ the group of linear isometries of $\left( \mathbb{F}_q^n, d_G \right)$, i.e.,

$$GL(n, G)_q = \{ T : \mathbb{F}_q^n \rightarrow \mathbb{F}_q^n; \ T \text{ linear and } d_G(x, y) = d_G(T(x), T(y)), \forall x, y \in \mathbb{F}_q^n \}.$$ 

We first remark is that $GL(n, G)_q = GL(n, \widetilde{G})_q$, so an automorphism $\phi \in \text{Aut} (\widetilde{G})$ induces an isometry $T_\phi \in GL(n, G)_q$, acting on $\mathbb{F}_q^n$ by permutation of the coordinates: $T_\phi ((x_1, x_2, \ldots, x_n)) = (x_{\phi(1)}, x_{\phi(2)}, \ldots, x_{\phi(n)})$.

We say that a linear map $T : \mathbb{F}_q^n \rightarrow \mathbb{F}_q^n$ respects domination if $T(e_i) = \sum^n_{j=1} \alpha_{ij} e_j$ satisfies the conditions: (i) $\alpha_{ii} \neq 0$ for every $i \leq n$; (ii) $\alpha_{ij} \neq 0$ implies that $v_j \in \langle v_i \rangle_G$. We denote $N(G)$ as the set of all linear maps respecting domination.

**Lemma 9:** $\text{Aut} (\widetilde{G})$ and $N(G)$ are subgroups of $GL(n, G)_q$.

**Proof:** It is clear that if $T_\phi \in \text{Aut} (\widetilde{G})$, then $T_\phi \in GL(n, G)_q$. To conclude that $\text{Aut} (\widetilde{G})$ is a subgroup, note that $T_\phi T^{-1}_\psi = T_{\phi \psi^{-1}}$. By construction, a linear map $T$ belongs to $N(G)$ if, and only if, $\text{Max}_G(x) = \text{Max}_G(T(x))$ for every $x \in \mathbb{F}_q^n$, and it follows that $N(G)$ is a subgroup of $GL(n, G)_q$.

**Lemma 10:** The group $\text{Aut} (\widetilde{G}) \cap N(G)$ is the semidirect product $\text{Aut} (\widetilde{G}) \rtimes N(G)$.

**Proof:** It is clear that $\text{Aut} (\widetilde{G}) \cap N(G) = \{ Id \}$, so all is needed is to prove that $N(G)$ is normal. This follows straightforward from the definition of the action of $\text{Aut} (\widetilde{G})$, noting that $\text{Max}(T \circ T \circ T^{-1}(x)) = \text{Max}(T(x))$.

We claim that actually $GL(n, G)_q = \text{Aut} (\widetilde{G}) \rtimes N(G)$. In order to prove it, we first need to prove some preliminary results.

**Lemma 11:** Given $T \in GL(n, G)_q$ and $e_i \in \beta$, there are $v_{j(i)} \in \text{supp}(T(e_i))$ and $S_i \in N(G)$ such that: (i) $\langle \text{supp}(T(e_i)) \rangle_G = \langle v_{j(i)} \rangle_G$; (ii) $\text{supp}(S; T(e_i))$ is a MSG for $\text{supp}(T(e_i))$.

**Proof:** (i) Since $T \in GL(n, G)_q$ it follows there is $v_{j(i)} \in \text{supp}(T(e_i))$ such that $v_i \in \text{supp}(T^{-1}(e_{j(i)}))$, and since $T^{-1} \in GL(n, G)_q$ it follows that $\langle \text{supp}(T(e_i)) \rangle_G = \langle v_{j(i)} \rangle_G$. (ii) Let $\overline{X}$ be a MSG of $\text{supp}(T(e_i))$. Item (i) ensures
that $\tilde{X} = \{v_{j(i)}\}$. The linear isometry defined by $S_i(e_j) = e_j$ if $j \neq j(i)$ and $S(T(e_i)) = e_{j(i)}$, satisfies the desired conditions. 

The previous Lemma ensures the existence of a map $\phi_T : V \rightarrow V$, where $\phi_T(i)$ is defined to be a vertex such that $\langle \text{supp}(T(e_i)) \rangle_G = \langle v_{\phi_T(i)} \rangle_{G'}$, for each $T \in GL(n,G)$. This map is not necessarily unique, since $\phi_T(i)$ may be exchanged by another vertex in the same clique. This amounts to the choice of a MSG made in the proof of item (i) of the previous Lemma to determine the map $S_i$. Considering the family $S = [S_i;1 \leq i \leq n]$, we get a well defined map $S_{\phi_T}$:

**Proposition 2:** The map $\phi_T \in Aut(\tilde{G})$ and $S_{\phi_T} = S_n \circ \cdots \circ S_2 \circ S_1 \circ T$.

**Theorem 8:** $GL(n,G) = Aut(\tilde{G}) \ltimes N(G)$.

**Proof:** It follows from Proposition 2 and the fact that $S_{\phi_T} \in Aut(\tilde{G})$, each $S_i \in N(G)$ and $T = S_n^{-1} \circ \cdots \circ S_1^{-1} \circ S_{\phi_T}$.

VI. G-Canonical Decomposition of linear codes for hierarchical graphs and its packing radius

As mentioned in Section III, the reduced canonical form $G'(V',E',L')$ of a graph $G(V,E)$ determines a partial order over $V'$, a poset, which we denote by $P_G = (V', \leq_G)$. We may write $\leq$ instead of $\leq_G$, if no confusion may arise.

In the case the graph has no circuit, that is, in case $G$ defines a poset, the so-called hierarchical posets play an exceptional role, since many of the known properties of codes, including MacWilliams Identity and Extension properties, hold for a poset metric if and only if the poset is hierarchical (see [12], [13], [6]). Many of those results, originally proved for the usual Hamming metric, depends essentially on the action of the group of linear isometries being transitive on spheres centered at 0. We shall derive many similar properties for graphs for which the reduced canonical form determines a hierarchical poset. We will generalize the $P$-canonical decomposition for hierarchical posets, (see [7]), proving that each linear code $C \subset \mathbb{F}_q^n$ is equivalent - up to linear isometry - to a code that may be expressed as a product of codes, each one having its support contained in a different level of the poset.

We start giving proper definitions. Let $G' = G'(V',E',L')$ be the reduced canonical form of $G(V,E)$. Considering $G'$ as a poset, it is naturally decomposed into levels. A chain in $G'$ is a subset $I \subset V'$ such that given $u,v \in I$, either $u \leq v$ or $v \leq u$. The length of a chain is just its cardinality. The height $h_{G'}(v)$ of an element $v \in V'$ is the maximal length of a chain that has $v$ as a maximal element. The height of $G'$ is $h_{G'} = \max\{h_{G'}(v) ; v \in V'\}$. The $i$-th level $V'_i$ of $G'$ is the set of elements of height $i$: $V'_i = \{v \in V' ; h_{G'}(v) = i\}$. The level structure of $G'$ induces a similar structure in the original graph $G$: $h_G(v) := h_{G'}(\pi(v))$ and $V_i = \{v \in V ; h_G(v) = i\}$.

The set $V'$ is decomposed as a disjoint union of its levels, $V' = V'_1 \cup \cdots \cup V'_{h_{G'}}$, called the level structure of the graph $G'$. A similar decomposition holds for the original graph $G$: $V = V_1 \cup \cdots \cup V_{h_{G}}$, where the union

May 2, 2017

DRAFT
is disjoint and \( h_G = h_{G'} \). The graph \( G \) and the reduced \( G' \) are hierarchical if \( u \in V'_i, i > 1 \) and \( v \in V'_{i-1} \), then \( u \rightarrow v \).

Two linear codes \( C, C' \in \mathbb{F}_q^n \) are said to be \( G \)-equivalent if there is \( T \in \text{GL}(n, G)_q \) such that \( T(C) = C' \).

**Definition 1:** Let \( G(V, E) \) be a graph with \( h_C \) levels. A linear code \( C \subset \mathbb{F}_q^n \) admits a \( G \)-canonical decomposition if \( C \) is \( G \)-equivalent to a linear code \( \bar{C} = C_1 \oplus \cdots \oplus C_{h_C} \), where \( \text{supp}(C_i) \subset V_i \).

**Lemma 12:** Let \( G(V, E) \) be a hierarchical graph with \( h_G \) levels. Let \( C \subset \mathbb{F}_q^n \) be a linear code with \( \text{supp}(C) \subset V_i \), for some \( i \le h_G \). Consider \( x \in \mathbb{F}_q^n \) such that \( \text{Max}_C(x) \subset V_i \) and \( \bar{x} \notin C \). Then, \( C \oplus \text{span}[x] \) and \( C \oplus \text{span}[\bar{x}] \) are \( G \)-equivalent.

**Proof:** Let \( \{f_1, \ldots, f_k\} \) be a basis of \( C \). Since \( \bar{x} \notin C \), then \( \alpha = \{\bar{x}, f_1, \ldots, f_k\} \) is a basis of \( C \oplus \text{span}[\bar{x}] \). We extend \( \alpha \), by canonical vectors, to a basis \( \beta = \{\bar{x}, f_1, \ldots, f_k, e_{j_1}, \ldots, e_{j_l}\} \) of \( \mathbb{F}_q^n \). Since \( \text{Max}_C(x) \subset T_G^i \) and \( \bar{x} \notin C \), we have that also \( x \notin C \) and consequently \( \alpha' = \{x, f_1, \ldots, f_k, e_{j_1}, \ldots, e_{j_l}\} \) of \( \mathbb{F}_q^n \). Let \( T : \mathbb{F}_q^n \rightarrow \mathbb{F}_q^n \) be the linear map defined by \( T(x) = \bar{x}, T(f_i) = f_i \) and \( T(e_{j_k}) = e_{j_k} \). Clearly \( T(C \oplus \text{span}[x]) = T(C \oplus \text{span}[\bar{x}]) \). By construction, \( \text{Max}(y) = \text{Max}(T(y)) \) for every \( y \in \mathbb{F}_q^n \), hence \( T \) is an isometry. 

**Remark 1:** The \( G \)-isometry \( T \), constructed in the previous lemma, satisfies \( \text{supp}(T(C \oplus \text{span}[x])) \subset V_i \) and \( T(y) = y \), for \( \text{supp}(y) \subset [n] \setminus V_i \). These properties will be used in the proof of Theorem 9.

Given \( x \in \mathbb{F}_q^n \), the \( i \)-th projection \( \bar{x}^i \in \mathbb{F}_q^n \) is defined by \( \bar{x}^i_j = x_j \) if \( j \in V_i \) and \( \bar{x}^i_j = 0 \) otherwise.

**Theorem 9:** Let \( G(V, E) \) be a graph. The poset \( P_G \) is hierarchical if, and only if, any linear code \( D \) admits a \( G \)-canonical decomposition.

**Proof:** First, suppose that \( P_G \) is a hierarchical poset with \( h_G \) levels. If \( \text{dim}(D) = 1 \), it is enough to use Lemma 12, considering \( C = \{0\} \).

Suppose the result holds for linear codes with dimension smaller than \( k \) and let \( D = \text{span}[x_1] \oplus \text{span}[x_2, \ldots, x_k] \) be a \( k \)-dimensional code. The induction hypothesis ensures that, for \( D' = \text{span}[x_2, \ldots, x_k] \), there is a linear isometry \( T' \) such that \( T'([D']) = \bigoplus_{i=1}^i [D'_i] \) and \( \text{supp}([D'_i]) \subset V_i \). Since \( T'(x_1) \notin T'(D') \), then \( \text{supp}(T'(x_1)) \cap T'(D') = \{0\} \) and there exists a level \( i \) such that \( T'(x_1) \notin D'_i \). Denote by \( i_0 \) the maximal level with this property and let \( y \) be defined by \( y_i = T'(x_i) \) if \( i \in V_j \) with \( j \le i_0 \) and \( y_i = 0 \) otherwise. Thus we find that \( T'(D) = \text{span}([y]) \oplus T'(D') \).

Considering \( C = C_{b_1}^i \), Lemma 12 ensures there is \( T \in \text{GL}(n, G)_q \) such that \( (D'_i) = D'_i \) for \( i \neq i_0 \) and \( \text{supp}(T(\text{span}[y] \oplus D'_{b_i})) = \text{supp}(\text{span}[y] \oplus D'_{b_i}) \subset V_{b_i} \). Therefore, if \( D_i = D'_i \) for \( i \neq i_0 \) and \( D_{b_0} = \text{span}[y] \oplus D'_{b_0} \), then \( T(T'(D)) = \bigoplus_{i=1}^i D_i \) is a linear code \( G \)-equivalent to \( C \).

On the other hand, suppose that \( P_G \) is not hierarchical and let \( i \in [h_G] \) be the lowest level of \( P_G \) for which there are \( a \in V'_i \) and \( b \in V'_{i+1} \) such that \( a \notin b \). Consider \( j \in \pi^{-1}(a) \) and \( k \in \pi^{-1}(b) \). The linear code \( C = \text{span}[e_j + e_k] \) cannot be \( G \)-equivalent to a canonically decomposed code \( \bar{C} \). Indeed, it follows from Proposition 2 that any linear \( G \)-isometry \( T \in \text{GL}(n, G)_q \) induces an automorphism \( \Phi_T : V \rightarrow V \). Thus, the closures \( \langle \text{supp}(T(e_j)) \rangle_G \) and \( \langle \text{supp}(T(e_k)) \rangle_G \) are generated by \( \Phi_T(j) \in V \) and \( \Phi_T(k) \in V \) respectively. Moreover, since \( \pi(j) \notin \pi(k) \), we have that \( \pi(\Phi_T(j)) \notin \pi(\Phi_T(k)) \). It follows that \( \text{Max}_G(T(\text{span}[e_j + e_k])) \supset \{\Phi_T(j), \Phi_T(k)\} \) is
not contained in a single level. Since \( \text{dim}(C) = 1 \) and \( T \in GL(n, G) \) is taken arbitrarily, we show that \( C \) does not admit a \( G \)-canonical decomposition.

**Remark 2:** In the proof of the previous theorem, we constructed a map \( T \), considering a basis \( \beta = \{x_1, \ldots, x_k\} \) of \( D \). For the purpose of this Theorem, the choice of the basis is immaterial. For future purpose (Lemma 17), it is worth to note that the choice may be done in such a way that the linear isometry \( T \) (which maps \( D \) into its \( G \)-canonical decomposition), restricted to \( \beta \) is defined by \( T(x_i) = x_i \). It follows that if \( P_C \) is hierarchical, given a code \( C \) it is possible to find a basis \( \tilde{\beta} = \{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_k\} \) such that, when considering only the maximal components of each \( x_i \), we get a basis \( \tilde{\beta} = \{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_k\} \) such that supp\((\tilde{x}_i)\) is contained in a single level of \( P_C \) and \( \tilde{\beta} \) generates a code \( D \) that is a \( G \)-canonical decomposition of \( C \).

Given a code \( C \subset \mathbb{F}_q^n \), we define its minimal \( G \)-distance as \( d(C) := \min \{d_C(x, y) ; x, y \in C, x \neq y\} \). It is clear that for a linear code the minimal distance is the minimal weight \( \min \{w_C(x) ; x \in \mathbb{F}_q^n, x \neq 0\} \). The packing radius \( R(C) \) of the code is the maximal radius of disjoint balls centered at codewords:

\[
R(C) := \max\{r \in [n] \cup \{0\} ; B_r(x) \cap B_r(y) = \emptyset, \forall x, y \in C, x \neq y\}.
\]

In general, the packing radius is not determined by the minimal distance. Consider, for example, the graph with vertices set \( V = \{v_1, v_2, v_3\} \) with the single edge \( (v_3, v_1) \). Considering the one-dimensional codes \( C_1 = \{000, 110\} \) and \( C_2 = \{000, 001\} \) we get that \( d(C_1) = d(C_2) = 2 \), while \( R(C_1) = 0 \) and \( R(C_2) = 1 \).

Let us consider the particular case where \( G' \) is hierarchical and \( L' : V' \rightarrow \mathbb{N} \) is constant on each level \( V'_i \), let us say assuming the value \( L'(i) \). In this situation, we have that \( R(C) = R(d(C)) \), i.e., the packing radius \( R(C) \) of a linear code \( C \) is determined by its minimal distance \( d(C) \) as follows:

**Proposition 3:** Suppose that the reduced canonical form \( G'(V', E', L') \) of \( G \) is hierarchical and that \( L' \) is constant and equal to \( L'(i) \) on each level \( V'_i \). Let \( x \in C \) be a codeword of minimal \( G \)-weight and let \( k_0 \) be the minimal level of \( V' \) intersected by \( C \) : \( k_0 = \min \{i ; \text{supp}(x) \cap V'_i \neq \emptyset, x \in C\} \). Then,

\[
R(C) = \left\lfloor \frac{|\pi(\text{supp}(x))| - 1}{2} \right\rfloor L'(k_0) + \sum_{i=1}^{k_0-1} |V'_i|L'(i),
\]

where \( X \) is a MSG for \( \langle \text{supp}(x) \rangle_G \).

**Proof:** Let us assume that \( C \) is \( G \)-canonical decomposed as \( C_{k_0} \oplus C_{k_0+1} \oplus \cdots \oplus C_{k_n} \). Given \( x \in C_{k_0} \), since \( L' \) is assumed to be constant on each level, it is immediate to check that

\[
w_C(x) = |\pi(\text{supp}(x))| L'(k_0) + \sum_{i=1}^{k_0-1} |V'_i|L'(i).
\]

Considering the \( G \)-canonical the \( G \)-canonical decomposition, we have that any codeword \( x \in C_{k_0} \) is a cleared codeword hence the minimal distance is attained by such a vector \( x_0 \) with \( |\pi(\text{supp}(x_0))| \) minimal. Let us denote \( r_0 = |\pi(\text{supp}(x_0))| \). Given two codewords \( x, y \in C_{k_0} \), the minimality of \( r_0 \) ensures that \( |\pi(\text{supp}(x)) \cap \pi(\text{supp}(y))| \geq r_0 \) and hence \( d_C(x, y) \geq r_0 L'(k_0) + \sum_{i=1}^{k_0-1} |V'_i|L'(i) \). To show that this radius can not be
increased, we choose a subset \( I \subset \pi(\operatorname{supp}(x_0)) \) with \(|I| = |r_0| + 1\) \( w_C(x) \leq \) and define \( y = (y_1, \ldots, y_n) \in C_{k_0} \) by 
\[ y_j = x_j \text{ if } j \in J \text{ and } y_j = 0 \text{ otherwise.} \]
It follows that
\[ d_G(x_0, y) = (\delta - 1) L'(k_0) + \sum_{i=1}^{k_0-1} |V_i'|L'(i) \quad \text{and} \quad w_C(y) = (\delta + 1) L'(k_0) + \sum_{i=1}^{k_0-1} |V_i'|L'(i), \]
where \( \delta = \left\lfloor \frac{|\pi(\operatorname{supp}(y))| - 1}{2} \right\rfloor. \)
Denoting \( \rho = (\delta + 1) L'(k_0) + \sum_{i=1}^{k_0-1} |V_i'|L'(i) \) we have that \( d_G(x_0, y) < \rho \) and \( w_C(y) = \rho \), that is, \( y \in B_\rho(0) \cap B_\rho(x_0) \), hence \( R(C) = \delta L'(k_0) + \sum_{i=1}^{k_0-1} |V_i'|L'(i) \).

From Proposition 3, \( R(C) = R(d_G(C)) \). We remark that this situation, whether the packing radius is determined by the minimal distance, is unusual but very helpful, since even for codes of dimension 1, determining the packing radius is, in general, an NP-hard problem (see [15, Section 4]).

VII. The MacWilliams Identity and Extension Theorem

In this section, we introduce the MacWilliams Identity and the MacWilliams Extension Property in the context of graph-metrics. The condition established in Theorem 9 for a graph \( G(V, E) \) to admit a \( G \)-decomposition is not sufficient to ensure the MacWilliams properties. Indeed, consider the metric induced by the graph \( G([4], E = \{(3, 4), (4, 3)\}) \). Let \( C_1 = \{0000, 1100\} \) and \( C_2 = \{0000, 0011\} \) be two one-dimensional binary codes. It is clear that \( C_1 \) is isometric to \( C_2 \), since \( w_C(1100) = w_C(0011) = 2 \). The dual codes \( C_1^\perp \) and \( C_2^\perp \) are generated by the sets \( \{1100, 1110, 1101\} \) and \( \{0011, 1011, 0111\} \) respectively. Direct calculations shows that \( W^G_{C_1}(X) = 1 + 4X^2 + 3X^4 \) and \( W^G_{C_2}(X) = 1 + 2X + 2X^2 + 2X^3 + X^4 \). Therefore, the MacWilliams Identity does not hold in full generality. It is also easy to check (considering the description of \( GL(n, G)_q \) given in Section V) that no linear isometry can map 1100 into 0011, i.e., MacWilliams Extension Property also does not hold in the general case. Our aim now is to find sufficient conditions to ensure the existence of a MacWilliams identity. Surprisingly, as we shall see, the Extension Property is strictly stronger than the MacWilliams Identity.

We start with a definition that will be crucial for both the cases.

Definition 2: Let \( G'(V', E') \) be the reduced canonical form of \( G \) and let \( V'_1 \cup \cdots \cup V'_{k_G} \) be the level decomposition of \( V' \). We say that \( G \) satisfies the Unique Decomposition Property (UDP) if \( S, S' \subset V'_i \) are sets satisfying \( \sum_{a \in S} L'(a) = \sum_{b \in S'} L'(b) \), then there is a bijection \( g : S \to S' \) such that \( L'(a) = L'(g(a)) \) for all \( a \in S \).

Proposition 4: Suppose that the reduced canonical form \( G' = (V', E', L') \) of \( G \) is a hierarchical poset. The group of linear isometries \( GL(n, G)_q \) acts transitively on the spheres of \( (\mathbb{F}_q^n, d_G) \) if, and only if, \( G \) satisfies the UDP.

Proof: Given \( x \in \mathbb{F}_q^n \), let \( \tilde{X} \) be a MSG for \( \langle \operatorname{supp}(x) \rangle \). Since \( G' \) is hierarchical, all elements of \( \tilde{X} \) belong to the same level, let us say \( \tilde{X} \subset V'_{k_0} \). Then we find that
\[ w_G(x) = \sum_{v \in \tilde{X}} L'(v) + \sum_{i=1}^{k_0-1} \sum_{v \in V'_i} L'(v), \]
and the result follows from Proposition 2. 

May 2, 2017 DRAFT
A. The MacWilliams Identity

We denote $A_i^G(C) = |\{c \in C : w_G(c) = i\}|$ and define the $G$-weight enumerator of a code $C$ as the polynomial

$$W^G_C(X) = \sum_{i=0}^{n} A_i^G(C)X^i.$$ 

Given a graph $G(V,E)$, let us denote by $\overline{G}(V,\overline{E})$ its reverse graph. It is easy to see that a closure $\bar{f} \in I(G)$ if, and only if, its complement ${\bar{f}} \in I(\overline{G})$. Furthermore, $A \subset V$ is a circuit in $G$ if, and only if, $A$ is also a circuit in $\overline{G}$. Therefore, the reduced canonical form $\overline{P}(\overline{G}(V',\leq_{\overline{G}}))$ of $\overline{G}$ is the dual of $P_G(V',\leq_{P_G})$.

Definition 3: (The MacWilliams Identity) A graph $G(V,E)$ admits a MacWilliams Identity if for every linear code $C \subset \mathbb{F}^n_q$, the $G$-weight enumerator $W^G_C(X)$ of $C$ determines the $G$-weight enumerator $\overline{W}^G_{C^\perp}(X)$ of the dual code $C^\perp$.

In order to characterize the graphs that admit the MacWilliams Identity, we briefly introduce some facts about additive characters that will be used in a manner similar to MacWilliams’ original approach, later adapted by Choi et al [14] to the poset case. An additive character $\chi$ of $\mathbb{F}_q$ is a nontrivial homomorphism of the additive group $\mathbb{F}_q$ into the multiplicative group of complex numbers with 1-norm. The next lemma is well known.

Lemma 13: Let $C \subset \mathbb{F}_q^n$ be a linear code and let $\chi$ of $\mathbb{F}_q$ be an additive character. Then, $\sum_{x \in C} \chi(x \cdot y) = C$ if $y \in C^\perp$ and 0 otherwise.

Lemma 14: Let $G(V,E)$ be a finite graph. Given a linear code $C$ of $\mathbb{F}_q^n$, then

$$A_i^G(C) = \frac{1}{|C|} \sum_{1 \leq i \leq n} \sum_{x \in C^*} \sum_{j \in S_i} \chi(x \cdot y),$$

where $S_i$ and $\overline{S}_j$ are the spheres of radii $i$ and $j$ considering the metric induced by the graphs $G$ and $\overline{G}$, respectively, i.e., $S_i = \{x \in \mathbb{F}_q^n : w_G(x) = i\}$ and $\overline{S}_j = \{x \in \mathbb{F}_q^n : w_{\overline{G}}(x) = j\}$.

Proof: First, note that $A_i^G(C) = |C \cap S_i| = \sum_{x \in C \cap S_i} 1$. Lemma 13 implies that

$$A_i^G(C) = \sum_{y \in S_i} \frac{1}{|C|} \sum_{x \in C^*} \chi(x \cdot y) = \frac{1}{|C|} \sum_{1 \leq i \leq n} \sum_{x \in C^*} \chi(x \cdot y) = \frac{1}{|C|} \sum_{1 \leq i \leq n} \sum_{y \in S_i} \chi(x \cdot y).$$

Lemma 15: A directed graph $G(V,E)$ admits the MacWilliams Identity if, given $i, j \in [n]$, and $x, x' \in S_i$, then

$$\sum_{y \in S_j} \chi(x \cdot y) = \sum_{y \in S_j} \chi(x' \cdot y).$$

Proof: Let $C_1, C_2 \subset \mathbb{F}_q^n$ be linear codes such that $W^G_{C_1}(X) = W^G_{C_2}(X)$. Since $G$ satisfies the first condition we have that $p_{ij} := \sum_{y \in S_j} \chi(x \cdot y)$ does not depend on the choice of $x \in S_i$. Therefore, $W^G_{C_1}(X) = W^G_{C_2}(X)$, since, from Lemma 14,

$$A^\perp_i(C_k^\perp) = \frac{1}{|C_k|} \sum_{1 \leq i \leq n} \sum_{x \in C_k \cap S_i} p_{ij} = \frac{1}{|C_k|} \sum_{1 \leq i \leq n} A^G_i(C_k)p_{ij}, \text{ for } k \in \{1, 2\}. $$

May 2, 2017 DRAFT
Lemma 16: Let $G(V, E)$ be a graph with $h_G = 1$ and let $P_G = (V', \leq)$ be its reduced canonical form (considered as a poset). Given $I, J \in I(G)$, let $S_I = \{x \in \mathbb{P}^n_\delta : \langle \text{supp}(x) \rangle_G = I\}$. Then, for $x \in S_I$ we have that,
\[
\sum_{y \in S_{P'}} \chi(x \cdot y) = (-1)^{|I \cap P'|} \prod_{i \in \pi(P \cap P')} (q^{L(i)} - 1) \prod_{j \in \pi(P')} (q^{|L(j)|} - 1)^{|P'| - 1}.
\]

Proof: Given $x \in \mathbb{P}^n_\delta$, the vector $\vec{x} \in \mathbb{P}^n_\delta$ is defined by $\vec{x}_j = x_j$ if $j \in \pi^{-1}(i)$ and $\vec{x}_j = 0$ otherwise. In order to simplify the notation, we assume that $V' = \{1, \ldots, m\}$, for an integer $m \leq |V|$. Thus,
\[
\sum_{y \in S_{P'}} \chi(x \cdot y) = \sum_{y \in S_{P'}} \chi\left(\sum_{i=1}^m \vec{x}_i \cdot \vec{y}_i\right) = \sum_{y \in S_{P'}} \chi\left(\vec{x} \cdot \vec{y}\right) = \prod_{i \in \pi(P')} (q^{L(i)} - 1) \sum_{\vec{y} \in \mathbb{P}^n_\delta | \vec{y} \neq 0} \chi\left(\vec{x} \cdot \vec{y}\right),
\]
where in the first equality we consider the fact that $\chi$ is a group homomorphism and we separate the product in the $J$ and $J^c$ parts and the second follows from the fact that $\vec{x} \cdot \vec{y} = 0$ for $i \in \pi(J)$ and $\chi(0) = 1$. We note that, given $\vec{y}^1 + \cdots + \vec{y}^n \in S_{P'}$, we have that $\vec{y}^i \in \mathbb{P}^n_\delta | \vec{y}^i \neq 0 \subset \mathbb{P}^n_\delta$ if $i \in \pi(J^c)$ and $\vec{y}^i = 0$, otherwise. It follows that
\[
\sum_{\vec{y}^i \in \mathbb{P}^n_\delta | \vec{y}^i \neq 0} \chi\left(\vec{x} \cdot \vec{y}^i\right) = \sum_{\vec{y}^i \in \mathbb{P}^n_\delta | \vec{y}^i \neq 0} \chi\left(\vec{x} \cdot \vec{y}^i\right) = \prod_{i \in \pi(J^c)} (q^{L(i)} - 1) \sum_{\vec{y} \in \mathbb{P}^n_\delta | \vec{y} \neq 0} \chi\left(\vec{x} \cdot \vec{y}\right).
\]
Furthermore, from Lemma 13, we have that $\sum_{\vec{y} \in \mathbb{P}^n_\delta | \vec{y} \neq 0} \chi\left(\vec{y} \cdot z\right) = q^{L(i)} - 1$ if $z = 0$ and equals $-1$ otherwise. Hence,
\[
\sum_{y \in S_{P'}} \chi(x \cdot y) = (-1)^{|I \cap \pi(P') \cap P'|} \prod_{i \in \{j \in \pi(P' \cap P') : \pi^{-1}(i) \subset \pi(F)\}} (q^{L(i)} - 1) \prod_{j \in \pi(P')} (q^{|L(j)|} - 1)^{|P'| - 1} = (-1)^{|I \cap P'|} \prod_{i \in \pi(P')} (q^{L(i)} - 1) \prod_{j \in \pi(P')} (q^{|L(j)|} - 1)^{|P'| - 1},
\]
where the second equality follows from the fact that $|\{i \in V' : \text{supp}(x) \cap P' \neq \emptyset\}| = |\{i \in V' : \pi^{-1}(i) \subset I \cap P'\}| = |\pi(I \cap P')|$. ■

In the next theorem, we will use a variation of the previous lemma, considering not only the closed set $J^c$, but the family $\{J\}$ of all closed sets with the same cardinality of $J$: Given $I, J \in I(G)$ and $x \in S_I$, then
\[
\sum_{y \in S_{P'}} \chi(x \cdot y) = \sum_{K \in \mathbb{P}^n_\delta} \sum_{y \in S_{K}} \chi(x \cdot y) = \sum_{K \in \mathbb{P}^n_\delta} (-1)^{|I \cap \pi(K)|} \prod_{i \in \pi(P \cap K)} (q^{L(i)} - 1) \prod_{j \in \pi(K)} (q^{|L(j)|} - 1)^{|K| - 1}.
\]

Theorem 10: Let $G(V, E)$ be a graph and suppose that its reduced canonical form $P_G = (V', \leq)$ is an antichain. The graph $G$ admits the MacWilliams Identity if, and only if, $G$ satisfies the UDP.

Proof: It is clear to see that $|I| = |I| \iff |P| = |P'|$. Given $x, x' \in S_I$, consider the closures $I_1 = \langle \text{supp}(x) \rangle_G$ and $I_2 = \langle \text{supp}(x') \rangle_G$. Thus, $\sum_{i \in \pi(I_1)} L(i) = \sum_{i \in \pi(I_2)} L(i)$. If $G$ satisfies the UDP, then there is a bijection $g : \pi(I_1) \rightarrow \pi(I_2)$ such that $L(i) = L(g(i))$. Furthermore, since $\sum_{i \in \pi(I_1)} L(i) = \sum_{i \in \pi(I_2)} L(i)$, then $g$ may be extended to $V'$. It implies that there is a $G$-automorphism $\varphi \in Aut(G)$ such that $\varphi(I_1) = I_2$.

Given $I \in I(G)$, consider $j = |I|$. From Lemma 16 it follows that $\sum_{y \in S_J} \chi(x \cdot y) = \sum_{y \in S_J} \chi(x' \cdot y)$. Since $h_G = 1$, the conditions 1) and 2) of Lemma 15 are equivalent and, then $G$ admits the MacWilliams Identity.
On the other hand, let $S, S' \subset V'$ be minimal sets such that: (i) $\sum_{i \in S} L'(i) = \sum_{i \in S'} L'(i)$; (ii) there is no bijection $g : S \to S'$ such that $L'(i) = L'(g(i))$. Consider the linear codes $C_1 = \text{span}[x]$ and $C_2 = \text{span}[y]$, where $x = \sum_{i \in S} e_i$ and $y = \sum_{i \in S'} e_i$. By construction, the $G$-weight enumerators of $C_1$ and $C_2$ are equal. Furthermore, by the minimality of $S$ and $S'$, $\text{supp}(C_1) \cap \text{supp}(C_2) = \emptyset$ and $L(i) \neq L'(j)$, for any $i \in S$ and $j \in S'$. It follows that there exists $i_0 \in S \cup S'$ such that $L'(i_0) \leq L'(j)$, for each $j \in S \cup S'$. Suppose $i_0 \in S$, then there is $z \in \mathbb{F}_q \setminus C_1^\perp$ such that $\text{supp}(z) \subset \pi^{-1}(i_0)$, what implies $A_{w_G(e)}(C_1^\perp) < A_{w_G(e)}(C_2^\perp)$.

**Lemma 17:** Let $G(V, E)$ be a graph and suppose that $P_G$ is a hierarchical poset with $l$ levels. Let $C$ be a linear code and $C_1 \oplus \cdots \oplus C_l$ its $G$-canonical decomposition. Let $D_i = \{ y \in C_i^\perp : \text{supp}(y) \subset V_i \}$ and $D = D_1 \oplus \cdots \oplus D_l$. Then, $C^\perp$ and $D$ are $G$-equivalent.

**Proof:** Let $\alpha = [x_1, x_2, \ldots, x_r]$ be a basis of $C$ such that $[\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_r]$ generates a $G$-canonical decomposition $C_1 \oplus \cdots \oplus C_r$, as ensured by Remark 2. Since $P_G$ is assumed to be hierarchical, so is $P_{\tilde{G}}$ and it follows, again by Remark 2, that $C^\perp$ admits a basis $\beta = [y_1, y_2, \ldots, y_{n-r}]$ such that $\tilde{\beta} = [\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_{n-r}]$ generates a code $\tilde{G}$-equivalent to $C^\perp$, that is a $G$-canonical decomposition of $C^\perp$. We claim that this code is actually $D$. Indeed, to conclude that, we need to prove $\tilde{x}_j \cdot \tilde{y}_k = 0$, for all $1 \leq j \leq r$ and $1 \leq k \leq n - r$. Actually, since all those vectors have the support in a single level, it is enough to prove $\tilde{x}_j \cdot \tilde{y}_k = 0$ for $\tilde{x}_j \in \tilde{\alpha}$ and $\tilde{y}_k \in \tilde{\beta}$ such that $\text{supp}(\tilde{x}_j) \cup \text{supp}(\tilde{y}_k) \subset V_i$. Note that, if $\text{supp}(\tilde{x}_j) \subset V_i$, then $\text{supp}(x_j) \subset V_1 \cup V_2 \cup \cdots \cup V_i$. Analogously, if $\text{supp}(\tilde{y}_k) \subset V_i$, then $\text{supp}(y_k) \subset V_1 \cup V_{i+1} \cup \cdots \cup V_i$, since the metric is induced by the reverse graph $\overline{G}$. This implies that $x_j \cdot y_k = \tilde{x}_j \cdot \tilde{y}_k$ and since $x_j \cdot y_k = 0$ we get that also $\tilde{x}_j \cdot \tilde{y}_k = 0$.

**Theorem 11:** Let $G(V, E)$ be a graph and suppose that $P_G$ is a hierarchical poset. The graph $G$ admits the MacWilliams Identity if, and only if, $G$ satisfies the UDP.

**Proof:** Suppose that the graph $G(V, E)$ has $h_G$ levels with $n_i$ elements in the $i$-th level. From Lemma 17, we may assume, without loss of generality, that a linear code $C \subset \mathbb{F}_q^n$ is already in a $G$-canonical form $C_1 \oplus \cdots \oplus C_{h_G}$ and its dual is equivalent to $C^\perp = D_1 \oplus D_2 \oplus \cdots \oplus D_l$, where $D_l$ is defined as in Lemma 17.

Furthermore, since the graph $G$ satisfies the Unique Decomposition Property, the graph induced on $V_i'$ also satisfies the Unique Decomposition Property and so, Theorem 10 ensures that $W_{C_i}(X)$ determines $W_{D_i}(X)$. Since $W_{\tilde{G}}(X)$ can be expressed in terms of the $W_{C_i}(X)$’s (and similarly for $W_{C_i^\perp}(X)$), we conclude the proof $W_{\tilde{G}}(X)$ is completely determined by $W_{\tilde{G}}(X)$.

**Remark 3:** As we saw, a poset-block metric can be obtained as a variation of a graph based metric by imposing $L_1(v) = 1$ for every vertex $v$ in the canonical reduced form. In this situation, the UDP condition is equivalent to demand that every block has the same size. For this fact, the condition we establish for the existence of a MacWilliams identity may be seen as a generalization of the conditions established in [17] for the case of poset-block metrics. We also remark that the UDP condition is necessary and sufficient only by assuming that the graph is hierarchical. That is not the case for poset-block metrics.
B. The MacWilliams Extension Theorem

Definition 4: (The MacWilliams Extension Property) A metric space \((\mathbb{F}_q^n, d_G)\) satisfies the MacWilliams Extension Property if for any pair of linear codes \(C\) and \(C'\) and any linear map \(t : C \to C'\) preserving the \(G\)-weight, there is a \(G\)-isometry \(T \in GL(n, G)_q\) such that \(T|_C = t\).

The same conditions on the graph \(G\) that ensured the MacWilliams Identity (the canonical reduced form being an hierarchical poset and the UDP) are not sufficient to characterize those metric spaces \((\mathbb{F}_q^n, d_G)\) that satisfy the MacWilliams Extension Property. In fact, consider the graph \(G([6], \text{ with set of edges } E = \{(1, 2), (2, 1), (3, 4), (4, 3), (5, 6), (6, 5)\})\). We consider the space \((\mathbb{F}_2^n, d_G)\) and the linear codes \(C_1 = \{000000, 100010, 101000, 001010\}\) and \(C_2 = \{000000, 010010, 110011, 100001\}\). The linear map \(t : C_1 \to C_2\) defined by \(t(100010) = 010010\) and \(t(101000) = 110011\) preserves the \(G\)-weight but any possible linear extension does not preserve the \(G\)-weight of the vector 100000; hence \(t\) cannot be extended to \(T \in GL(n, G)_2\).

This situation can be avoided by adding an extra condition.

Definition 5: (Condition \(\Omega\)) Let \(G(V, E)\) be a directed graph and \(G'(V', E', L)\) be its reduced canonical form. We say that \(G\) and \(G'\) satisfy the condition \(\Omega\) if, given an integer \(k > 1\), there are at most two elements \(i, j \in V'\) such that \(L(i) = L(j) = k\).

From here on, we assume that the field \(\mathbb{F}_q\) is binary, that is, \(q = 2\).

Let us assume that the reduced canonical form \(G'\) of \(G\) is a poset with a single level, that is, \(E' = \emptyset\). Given a linear code \(C \subset \mathbb{F}_2^n\). We define \(I_j(C) = \{k \in \pi(\text{supp}(C)); L(k) = j\}\). When considering two codes \(C_1\) and \(C_2\), we shall write \(I_j(C_i) = I_j^i\).

Lemma 18: Let \(G\) be a directed graph satisfying the UDP and condition \(\Omega\). Let \(C_1\) and \(C_2\) be linear codes in \((\mathbb{F}_2^n, d_G)\) and \(t : C_1 \to C_2\) a linear map preserving the \(G\)-weight. Then, \(|I_j^1| = |I_j^2|\) for all \(j \leq r = \max\{L(k); k \in V'\}\).

Proof: Let us decompose the support of the codes as \(\text{supp}(C_1) = I_1^1 \cup I_1^2 \cup \cdots \cup I_r^1\). We remark that some of the \(I_j^1\) may be empty; however, once \(I_j^1 = \emptyset\), we must have that also \(I_j^2 = \emptyset\). Indeed, suppose that \(I_j^1 \neq \emptyset\). This means there is \(x \in C_1\) such that \(\pi(\text{supp}(x)) \cap I_j^1 \neq \emptyset\). The UDP ensures that \(\pi(\text{supp}(t(x))) \cap I_j^2\) is also nonempty.

Let us consider \(2 \leq j \leq r\). Suppose that \(|I_j^1| \leq |I_j^2|\), since \(|I_j^1| \leq 2\) by hypothesis, then the possible values for \((|I_j^1|, |I_j^2|)\) are \((0, 0), (0, 1), (0, 2), (1, 1), (1, 2)\) and \((2, 2)\). The cases \((0, 1)\) and \((0, 2)\) cannot occur, since, as we just saw, \(I_j^1 = \emptyset\) iff \(I_j^2 = \emptyset\). We need to discard the case \((|I_j^1|, |I_j^2|) = (1, 2)\). Let \(x \in C_1\) be a vector such that \(\pi(\text{supp}(x)) = I_j^1\). The UDP ensures there is no \(y \in C_2\) such that \(\pi(\text{supp}(y)) = I_j^2\). However, there must be \(y, z \in C_2\) such that \(\pi(\text{supp}(y)) \cup \pi(\text{supp}(z)) = I_j^2\) and then we have that either \(\pi(\text{supp}(y)) = I_j^2\), or \(\pi(\text{supp}(z)) = I_j^2\), or \(\pi(\text{supp}(y + z)) = I_j^2\), a contradiction.

Now, we need to prove that \(I_1^1 = I_1^2\). We consider a set \(X = \{x_1, x_2, \ldots, x_s\}\) such that \(\bigcup_{i=1}^s (\pi(\text{supp}(x_i)) \cap I_1^1) = I_1^1\), with \(s\) minimal. A careful use of the inclusion-exclusion principle and the UDP ensures that
\[
\left| \bigcup_{i=1}^s (\pi(\text{supp}(x_i)) \cap I_1^1) \right| = \left| \bigcup_{i=1}^s (\pi(\text{supp}(t(x_i))) \cap I_1^2) \right| \leq |I_1^2|.
\]
and we get that $|I^1_l| \leq |I^2_l|$. A similar reasoning for the inverse $t^{-1}$ ensures that $|I^1_l| = |I^2_l|$. 

**Proposition 5:** Let $G(V,E)$ be a directed graph and suppose that its canonical reduced form is a poset with a single level $(h_G = 1)$. The metric space $(\mathbb{F}_2^n, d_G)$ satisfies the MacWilliams Extension Property if, and only if, $G$ satisfies the UDP and the Condition $\Omega$.

**Proof:** First of all, we shall prove that the two stated conditions are sufficient. Let $C_1$ and $C_2$ be two linear codes and let $t : C_1 \rightarrow C_2$ be a linear map that preserves the $G$-weight. Given $x, y \in C_1$, the UDP ensures the existence of the bijections $g_1 : \pi(\text{supp}(x)) \rightarrow \pi(\text{supp}(t(x)))$ and $g_2 : \pi(\text{supp}(y)) \rightarrow \pi(\text{supp}(t(y)))$. We claim that it is possible to choose $g_1$ and $g_2$ in such a way that, if $i \in \pi(\text{supp}(x)) \cap \pi(\text{supp}(y))$, then $g_1(i) = g_2(i)$. Indeed, suppose that $g_1(i) \neq g_2(i)$ and consider the linear codes $C = \text{span}(x,y)$ and $C' = \text{span}(t(x), t(y))$. The only obstructions to $g_1$ and $g_2$ to satisfy this condition would be if either $|I_{L(i)}(C)| < |I_{L(i)}(C')|$ or $|I_{L(i)}(\text{span}(x))| < |I_{L(i)}(\text{span}(t(x)))|$, contradicting Lemma 18. Hence, there is a bijection $\phi_1 : \pi(\text{supp}(C_1)) \rightarrow \pi(\text{supp}(C_2))$ such that $\phi_1 : \pi(\text{supp}(x)) \rightarrow \pi(\text{supp}(t(x)))$ is a bijection preserving the $L$-weight. Since $G$ satisfies the UDP and

$$\sum_{i \in V \setminus \pi(\text{supp}(C_1))} L(i) = \sum_{i \in V \setminus \pi(\text{supp}(C_2))} L(i),$$

then $\phi_1$ can be extended to $\phi : V' \rightarrow V'$. Given $x \in \mathbb{F}_2^n$, we write $x = x_1 + \cdots + x_m$, where $\pi(\text{supp}(x_i)) \subset \{v_i\} \subseteq V'$. It follows that $C_j \subseteq C_{j1} \oplus \cdots \oplus C_{jm}$, where $C_{ji} = \{x_i : x \in C_j\}$, for $j = 1, 2$. Let $\beta_j = \{x_j^1, \ldots, x_j^k\}$ be a basis of $C_{1j}$ and $\alpha_{\psi(i)} = \{t(x)_{\psi(i)j1}, \ldots, t(x)_{\psi(i)jk}\}$ be a basis of $C_{2\psi(i)}$. We consider $W_j$ to be the subspace of $\mathbb{F}_2^n$ isomorphic to $\mathbb{F}_2^{l(\psi(i))}$ such that $\pi(\text{supp}(e_j)) = \psi(j)$ and we extend the basis $\beta_j$ of $C_{1j}$ to a basis $\beta_j' = \{x_j^1, \ldots, x_j^k, e_j^1, \ldots, e_j^r\}$ of $W_j$. It follows that, $\mathbb{F}_2^n = \bigoplus_{i=1}^m W_j$. Analogously, we can also extend $\alpha_{\psi(i)}$ to $\alpha'_{\psi(i)} = \{t(x)_{\psi(i)j1}, \ldots, t(x)_{\psi(i)jk}, f_{\psi(i)j1}, \ldots, f_{\psi(i)jr}\}$ to another basis of $W_{\psi(i)}$. Let $T : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ be the linear map defined by $T(x_{\psi(i)}) = t(x)_{\psi(i)j1}$ and $T(e_{\psi(i)}) = f_{\psi(i)jr}$. By construction, $T$ is a linear $G$-isometry that extends $t$.

Now we prove that the two stated conditions are necessary.

**Case 1:** Suppose that $G$ does not satisfy the UDP. Consider the sets $S, S' \subset V'$ and the linear codes $C_1 = \text{span}(x)$ and $C_2 = \text{span}(y)$ as defined in the proof of Theorem 10. By construction, there is a linear map $t : C_1 \rightarrow C_2$ that preserves the $G$-weight. Suppose that $t$ can be extended to $T \in GL(n, G)$. Then, Lemma 12 ensures that $T$ defines a $G$-automorphism $\phi_G$ such that $\pi(\phi_G(\pi^{-1}(S))) = S'$, an absurd, since $\Gamma(S)$ and $\Gamma(S')$ are assumed to be non-isomorphic $L$-weighted subgraphs of $G'$.

**Case 2:** Suppose there are different elements $v_1, v_2, v_3 \in V'$ such that $L(v_1) = L(v_2) = L(v_3) = l > 1$. For $i = 1, 2, 3$, let $e_{a_i}$ and $e_{b_i}$ be two different vectors in $\mathbb{F}_2^n$ such that $\pi(\text{supp}(e_{a_i})) = \pi(\text{supp}(e_{b_i})) = v_i$, which existence is ensured by the fact that $L(v_i) \geq 2$. Let us define the codes $C_1 = \text{span}(e_{a_1}, e_{a_2}, e_{a_3})$ and $C_2 = \text{span}(e_{a_1} + e_{a_2}, e_{a_1} + e_{a_3}, e_{a_2} + e_{a_3}) \subseteq \mathbb{F}_2^n$. Since all the codewords have the same weight $2l$, any linear isomorphism $t$ between the codes preserves the $G$-weight. However, $3 = |I^1_l| - |I^2_l| = 3$ so $t$ cannot be extended by an isometry $T \in GL(n, G)$.

**Theorem 12:** Let $G(V,E)$ be a graph and suppose that $P_G$, the poset determined by its reduced canonical form, is a hierarchical poset. Hence, $(\mathbb{F}_2^n, d_G)$ satisfies the MacWilliams Extension Property if, and only if, $G$
satisfies the UDP and the graph $\Gamma(V_i)$ satisfies the Condition $\Omega$, for each $1 \leq i \leq h_G$.

Proof: The canonical decomposition (Theorem 9) ensures we may assume that $C = C_1 \oplus \cdots \oplus C_{h_G}$ and $C' = C'_1 \oplus \cdots \oplus C'_{h_G}$. For each $c \in C_i$, we have $t(c) = t_i(c) + F_i(c)$, where $F_i : C_i \rightarrow \sum_{j<i} C'_j$ and $t_i : C_i \rightarrow C'_i$ are both linear maps, and $w_C(c) = w_{C'}(t_i(c))$. Then, it is easy to verify that $t_i : C_i \rightarrow C'_i$ is also a linear isometry. Since supp($C_i$), supp($C'_i$) $\subset V_i$, we can consider $C_i, C'_i \subset \mathbb{F}_q^n$ to be equipped with the metric based on graph $\Gamma(V_i)$ on $\mathbb{F}_q^n$, where $\Gamma(V_i)$ is the graph induced by $G$ on $V_i$. The previous proposition ensures that each $t_i$ admits an extension $T_i$ to $\mathbb{F}_q^n$, it means that, there is a linear $V_i$-isometry $T_i \in GL(n_i, \Gamma(V_i))_2$ of $\mathbb{F}_q^n$ into itself and $T_i|_{C_i} = t_i$. Since every linear map between linear codes can be extended to the entire space, it follows that the linear map $T(x_1 + \cdots + x_{h_G}) = (T_1 + F_1)(x_1) + \cdots + (T_{h_G} + F_{h_G})(x_{h_G})$ is a $G$-isometry.

Since we are assuming that $G'$ is hierarchical, the UDP and the Condition $\Omega$ are checked at each level of the graph. If both the conditions holds up to the level $V_{j-1}$ of $G$ but one of them fails on level $V_j$, it is possible to construct codes $C_1$ and $C_2$ as in Proposition 5, just taking the care to consider the generator vectors $x, y$ (in case the UDP does not hold) and the vectors $e_{ai}, e_{bi}$ (in case condition $\Omega$ does not hold) to have the support on the $j$-level.

\section*{Acknowledgment}

This work was supported in part by the São Paulo Research Foundation (FAPESP) under grants no. 2014/10745-6, 2013/25977-7 and 2015/11286-8.

\section*{References}

[1] M. M. S. Alves, L. Panek, and M. Firer, Error-block codes and poset metrics, \textit{Advances in Mathematics of Communications}, vol. 2, no. 1, pp. 95111, 2008.

[2] László Babai, “Graph isomorphism in quasipolynomial time,” [extended abstract], in \textit{Proceedings of the forty-eighth annual ACM Symposium on Theory of Computing} (STOC ’16). ACM, New York, NY, USA, 684-697, 2016.

[3] A. Barg and P. Purkayastha, “MDS poset codes and distributions”, in \textit{Error-Correcting Codes, Finite Geometries, and Cryptography, AMS Series: Contemporary Mathematics} vol. 523 pp. 135–147, 2010.

[4] A. Barra and H. Gluesing-Luerssen, MacWilliams extension theorems and the local-global property for codes over Frobenius rings, \textit{Journal of Pure and Applied Algebra}, vol. 219, no. 4, pp. 703–728, 2015.

[5] R. A. Brualdi, J. S. Graves, and K. M. Lawrence, “Codes with a poset metric”, \textit{Discrete Mathematics}, vol. 147 pp. 57–72, 1995.

[6] S. Choi, J. Y. Hyun, H. K. Kim, and D. Y. Oh, “MacWilliams-type equivalence relations,” arXiv:orgabs/1205.1090v2, 2013.

[7] L. V. Felix and M. Firer, Canonical-systematic form for codes in hierarchical poset metrics, \textit{Advances in Mathematics of Communications}, vol. 6, no. 3, pp. 315328, 2012.

[8] Gabidulin E. M., “Combinatorial metrics in coding theory,” in \textit{Proc. of the 2-nd ISIT}, Budapest, 1971, 169-176.

[9] M. Greferath, T. Honold, C. M. Fadden, J. A. Wood, and J. Zumbragel, MacWilliams extension theorem for bi-invariant weights over finite principal ideal rings, \textit{Journal of Combinatorial Theory, Series A}, vol. 125, pp. 177–193, 2014.

[10] J. Y. Hyun and Y. Lee, “MDS poset-codes satisfying the asymptotic Gilbert-Varshamov bound in Hamming weights”, \textit{IEEE Trans. Inform. Theory}, vol. IT-57, pp. 8021–8026, 2011.

[11] H. K. Kim and D. S. Krotov, “The poset metrics that allow binary codes of codimension $m$ to be $m$-, $(m-1)$-, or $(m-2)$-perfect”, \textit{IEEE Trans. Inform. Theory}, vol. IT-54, pp. 5241–5246, 2008.
[12] H. K. Kim and D. Y. Oh, “A classification of posets admitting the MacWilliams identity”, *IEEE Trans. Inform. Theory*, vol. IT-51, pp. 1424–1431, 2005.

[13] R. Machado, J. Pinheiro and M. Firer, “Characterization of metrics induced by hierarchical posets”, in *IEEE Trans. Inform. Theory*, vol. PP, no. 99, pp. 1-1.

[14] S. Choi, J.Y. Hyun, H.K. Kim and D.Y. Oh “MacWilliams-type equivalence relations” *arXiv preprint arXiv:1205.1090*, 2012.

[15] R. G. L. D’Oliveira and M. Firer, “The packing radius of a code and partitioning problems: The case for poset metrics on finite vector spaces”, *Discrete Mathematics*, vol. 338, pp. 2143–2167, 2015.

[16] A. de Oliveira Moura and M. Firer, “Duality for poset codes”, *IEEE Trans. Inform. Theory*, vol. IT-56, pp. 3180–3186, 2010.

[17] J. A. Pinheiro and M. Firer, "Classification of Poset-Block Spaces Admitting MacWilliams-Type Identity," *IEEE Trans. Inform. Theory*, vol. 58, no. 12, pp. 7246-7252, 2012.

[18] J. Torán, “On the hardness of graph isomorphism”, *SIAM J. Computing*, vol. 33, pp. 1093–1108, 2004.