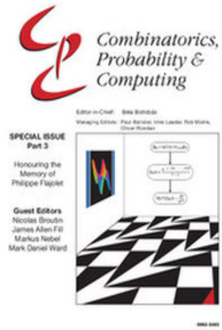

Optimal Codes in the Enomoto-Katona Space

YEOW MENG CHEE, HAN MAO KIAH, HUI ZHANG and XIANDE ZHANG

Combinatorics, Probability and Computing / *FirstView* Article / December 2014, pp 1 - 25
DOI: 10.1017/S0963548314000509, Published online: 09 October 2014

Link to this article: http://journals.cambridge.org/abstract_S0963548314000509

How to cite this article:
YEOW MENG CHEE, HAN MAO KIAH, HUI ZHANG and XIANDE ZHANG Optimal Codes in the Enomoto-Katona Space. Combinatorics, Probability and Computing, Available on CJO 2014 doi:10.1017/S0963548314000509

Combinatorics, Probability and Computing: page 1 of 25.
doi:10.1017/S0963548314000509

Optimal Codes in the Enomoto–Katona Space

YEOW MENG CHEE[1], HAN MAO KIAH[2]†, HUI ZHANG[1] and XIANDE ZHANG[1]‡

[1]Division of Mathematical Sciences, School of Physical and Mathematical Sciences,
Nanyang Technological University, 21 Nanyang Link, Singapore 637371, Singapore
(e-mail: {ymchee,huizhang,XiandeZhang}@ntu.edu.sg)

[2] Coordinated Science Lab, University of Illinois, Urbana-Champaign, 1308 W. Main Street, Urbana, IL 61801, USA
(e-mail: hmkiah@illinois.edu)

Received 11 October 2013; revised 17 March 2014

Coding in a new metric space, called the Enomoto–Katona space, has recently been considered in connection with the study of implication structures of functional dependencies and their generalizations in relational databases. The central problem is the determination of $C(n,k,d)$, the size of an optimal code of length n, weight k, and distance d in the Enomoto–Katona space. The value of $C(n,k,d)$ was known only for some congruence classes of n when $(k,d) \in \{(2,3),(3,5)\}$. In this paper, we obtain new infinite families of optimal codes in the Enomoto–Katona space and verify a conjecture of Brightwell and Katona in certain instances. In particular, $C(n,k,2k-1)$ is determined for all sufficiently large n satisfying either $n \equiv 1 \bmod k$ and $n(n-1) \equiv 0 \bmod 2k^2$, or $n \equiv 0 \bmod k$. We also give complete solutions for $k = 2$ and determine $C(n,3,5)$ for certain congruence classes of n with finite exceptions.

2010 *Mathematics subject classification*: Primary 05C35
Secondary 05C15, 57M25

1. Introduction

The problem we consider is motivated by implication structures of functional dependencies in relational databases.

Let A be a set of n *attributes*. Each attribute $x \in A$ is associated with a set Ω_x, called its *domain*. A *relation* is a finite set R of n-tuples (called *data items*) such that $R \subseteq \times_{x \in A} \Omega_x$. A relation R of m data items may be visualized as an $m \times n$ array (called a *table*), with

† This work was done while H. M. Kiah was a graduate student at the Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore.

‡ Research of the authors is supported in part by the Singapore National Research Foundation under research grant NRF-CRP2-2007-03. X. Zhang is also supported in part by NSFC under grant 11301503.

columns indexed by A, such that each row corresponds to a data item. Denote this table by $R(A)$. Formally, if $R = \{(\mathsf{d}_{i,x})_{x\in A} : 1 \leqslant i \leqslant m\}$, then the cell in $R(A)$ with row index i and column index x has entry $\mathsf{d}_{i,x}$. A *relational database* is a set of tables, where tables may be defined over different attribute sets. The relational database, introduced by Codd [10], was the first database with a rigorous mathematical foundation, and remains the predominant choice for data storage and management today.

For a given table $R(A)$ and $X \subseteq A$, the X-*value* of a data item $\mathsf{d} = (\mathsf{d}_x)_{x\in A}$ in $R(A)$ is the $|X|$-tuple $\mathsf{d}|_X = (\mathsf{d}_x)_{x\in X}$. Let $X \subseteq A$ and $y \in A$ for a given table $R(A)$. We say that y (*functionally*) *depends*[1] on X, written as $X \to y$, if no two rows of $R(A)$ agree in X but differ in y. In other words, if the X-value of a data item is known, then its $\{y\}$-value can be determined with certainty. Identifying functional dependencies is important in relational database design [3–5, 19].

Demetrovics, Katona and Sali [12] generalized functional dependencies as follows.

Definition. Let $X \subseteq A$ and $y \in A$ for a given table $R(A)$. Then, for positive integers $p \leqslant q$, we say that y (p, q)-*depends* on X, written as $X \xrightarrow{(p,q)} y$, if there do not exist $q + 1$ data items (rows) $\mathsf{d}_1, \mathsf{d}_2, \ldots, \mathsf{d}_{q+1}$ of $R(A)$ such that

(i) $|\{\mathsf{d}_i|_{\{x\}} : 1 \leqslant i \leqslant q + 1\}| \leqslant p$ for each $x \in X$, and
(ii) $|\{\mathsf{d}_i|_{\{y\}} : 1 \leqslant i \leqslant q + 1\}| = q + 1$.

Our usual concept of functional dependency is equivalent to $(1, 1)$-dependency. When functional dependencies are not known, (p, q)-dependencies identified in a relational database can still be exploited to improve storage efficiency [11–13, 16].

Let $p \leqslant q$ be positive integers. For a table $R(A)$, define the operation $J^{(p,q)}_{R(A)} : 2^A \to 2^A$ such that for $X \subseteq A$ we have

$$J^{(p,q)}_{R(A)}(X) = \{y \in A : X \xrightarrow{(p,q)} y\}.$$

We call $J^{(p,q)}_{R(A)}$ the (p, q)-*implication structure* of $R(A)$, since it specifies the subsets of attributes that are implied by some (p, q)-dependency of $R(A)$. A function $J : 2^A \to 2^A$ is said to be (p, q)-*representable* if there exists a table $R(A)$ such that $J^{(p,q)}_{R(A)} = J$.

The function $J^{(1,1)}_{R(A)}$ is a closure operator on A. Armstrong [3] showed that the converse is also true: any closure operator $J : 2^A \to 2^A$ is $(1, 1)$-representable. This is, however, not true for general p and q [12]. When a function J is (p, q)-representable, there is interest in determining the table $R(A)$ with the least number of rows such that $J^{(p,q)}_{R(A)} = J$ [11, 13, 16]. Consideration of this problem, particularly for fixed k, the function $J^k_n : 2^A \to 2^A$ that takes the form

$$J^k_n(X) = \begin{cases} X & \text{if } |X| < k, \\ A & \text{otherwise,} \end{cases}$$

led to coding-theoretic problems in a new metric space, called the *Enomoto–Katona space* [14].

[1] By definition, if $y \in X$, then $X \to y$ trivially.

1.1. The Enomoto–Katona space

If X is a finite set, the set of all k-subsets of X is denoted by $\binom{X}{k}$. Let n and k be positive integers such that $2k \leqslant n$ and let X be an n-set. Consider the set

$$\mathcal{E}(X,k) = \left\{ \{A,B\} \subseteq \binom{X}{k} : A \cap B = \varnothing \right\}$$

of all unordered pairs of disjoint k-subsets of X. Elements of $\mathcal{E}(X,k)$ are called *set-pairs*. The function $\mathrm{d}_{\mathcal{E}} : \mathcal{E}(X,k) \times \mathcal{E}(X,k) \to \{0,1,\ldots,2k\}$ given by

$$\mathrm{d}_{\mathcal{E}}(\{A,B\},\{S,T\}) = \min\{|A \setminus S| + |B \setminus T|, |A \setminus T| + |B \setminus S|\}$$

is a metric of $\mathcal{E}(X,k)$, and the finite metric space $(\mathcal{E}(X,k), \mathrm{d}_{\mathcal{E}})$ is called the *Enomoto–Katona space.*

An *Enomoto–Katona code* (or *EK code* for short), is a set $\mathcal{C} \subseteq \mathcal{E}(X,k)$. More specifically, $\mathcal{C}$ is an EK code of *length n*, *weight k*, and *distance d*, or (n,k,d) EK code, if $\mathrm{d}_{\mathcal{E}}(\mathsf{u},\mathsf{v}) \geqslant d$ for all distinct $\mathsf{u},\mathsf{v} \in \mathcal{C}$.

The following example gives a construction of a table from an EK code (see [13, 20]).

Example. Consider the $(19,3,5)$ EK code $\mathcal{C} = \{c_i : i \in \mathbb{Z}_{19}\}$, where $X = \mathbb{Z}_{19}$ with

$$c_i = \{\{i, i+1, i+4\}, \{i+3, i+8, i+14\}\} \text{ for } i \in \mathbb{Z}_{19}.$$

Let A be a set of nineteen attributes, given by $\mathcal{C}$. We construct a table $R(A)$ with nineteen rows indexed by X, whose implication structure $J^{(1,2)}_{R(A)}$ is precisely J^2_{19}. Each set-pair $\{A,B\}$ constructs a column in the following manner: place 1 at rows indexed by elements of A, place 2 at rows by elements of B and place distinct integers greater than 2 in the remaining rows. Therefore, the table $R(A)$ is as shown in Table 1.

The maximum size of an (n,k,d) EK code is denoted by $C(n,k,d)$. An (n,k,d) EK code of size $C(n,k,d)$ is said to be *optimal*. The central problem is to determine $C(n,k,d)$.

1.2. Problem status

Trivially, $C(n,k,1) = \binom{n}{k}\binom{n-k}{k}/2$ and $C(n,k,2k) = \lfloor n/2k \rfloor$, so we assume $2 \leqslant d \leqslant 2k-1$ for the rest of this paper.

General upper and lower bounds on the size of codes in the Enomoto–Katona space have been obtained by Brightwell and Katona [7]. In particular, they showed for $1 \leqslant d \leqslant 2k \leqslant n$ that

$$C(n,k,d) \leqslant \frac{\prod_{i=n-2k+d}^{n} i}{2\left(\prod_{i=\lceil (d+1)/2 \rceil}^{k} i\right)\left(\prod_{i=\lfloor (d+1)/2 \rfloor}^{k} i\right)}. \tag{1.1}$$

Brightwell and Katona [7] also showed that $C(n,k,d) = \Theta(n^{2k-d+1})$ for fixed k and d. Furthermore, for fixed k and d they conjectured that the bound (1.1) is attained for infinite values of n. Bollobás, Füredi, Kantor, Katona and Leader [6] (see also [20]) subsequently established that the upper bound in (1.1) is asymptotically tight.

Table 1. The table $R(A)$.

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}	c_{13}	c_{14}	c_{15}	c_{16}	c_{17}	c_{18}	c_{19}
0	1	3	3	3	3	2	3	3	3	3	3	2	3	3	3	1	2	3	1
1	1	1	4	4	4	3	2	4	4	4	4	3	2	4	4	3	1	2	3
2	3	1	1	5	5	4	4	2	5	5	5	4	4	2	5	4	3	1	2
3	2	4	1	1	6	5	5	5	2	6	6	5	5	5	2	5	4	4	1
4	1	2	5	1	1	6	6	6	6	2	7	6	6	6	6	2	5	5	4
5	4	1	2	6	1	1	7	7	7	7	2	7	7	7	7	6	2	6	5
6	5	5	1	2	7	1	1	8	8	8	8	2	8	8	8	7	6	2	6
7	6	6	6	1	2	7	1	1	9	9	9	8	2	9	9	8	7	7	2
8	2	7	7	7	1	2	8	1	1	10	10	9	9	2	10	9	8	8	7
9	7	2	8	8	8	1	2	9	1	1	11	10	10	10	2	10	9	9	8
10	8	8	2	9	9	8	1	2	10	1	1	11	11	11	11	2	10	10	9
11	9	9	9	2	10	9	9	1	2	11	1	1	12	12	12	11	2	11	10
12	10	10	10	10	2	10	10	10	1	2	12	1	1	13	13	12	11	2	11
13	11	11	11	11	11	2	11	11	11	1	2	12	1	1	14	13	12	12	2
14	2	12	12	12	12	11	2	12	12	12	1	2	13	1	1	14	13	13	12
15	12	2	13	13	13	12	12	2	13	13	13	1	2	14	1	1	14	14	13
16	13	13	2	14	14	13	13	13	2	14	14	13	1	2	15	1	1	15	14
17	14	14	14	2	15	14	14	14	14	2	15	14	14	1	2	15	1	1	15
18	15	15	15	15	2	15	15	15	15	15	2	15	15	15	1	2	15	1	1

Theorem 1.1 (Bollobás *et al.* [6]).

$$\lim_{n\to\infty} \frac{C(n,k,d)}{n^{2k-d+1}} = \frac{1}{2\left(\prod_{i=\lceil (d+1)/2\rceil}^{k} i\right)\left(\prod_{i=\lfloor (d+1)/2\rfloor}^{k} i\right)}.$$

The best known upper bound is due to Quistorff [18].

Theorem 1.2 (Quistorff bound [18]). *Suppose $k-d+1 \leqslant e \leqslant \min\{k, 2k-d\}$. Then*

$$C(n,k,d) \leqslant \left\lfloor \frac{\binom{n}{e}}{2\binom{k}{e}} \left\lfloor \frac{\binom{n-e}{2k-d-e+1}}{\binom{k}{2k-d-e+1}} \right\rfloor \right\rfloor.$$

Only the following exact values of $C(n,k,d)$ are known, and this verifies the conjecture of Brightwell and Katona for $(k,d) \in \{(2,3),(3,5)\}$.

Theorem 1.3 (Bollobás *et al.* [6]).

$$C(n,2,3) = \frac{n(n-1)}{8}, \quad \textit{if } n \equiv 1 \textit{ or } 9 \bmod 72 \textit{ and}$$
$$C(n,3,5) = \frac{n(n-1)}{18}, \quad \textit{if } n \equiv 1 \textit{ or } 19 \bmod 342.$$

1.3. Contributions

Our contributions in this paper are as follows.

Theorem 1.4. *Let $n \geqslant 2k$.*

(i) *For any fixed $k \geqslant 2$, we have*

$$C(n,k,2k-1) = \left\lfloor \frac{n}{2k} \left\lfloor \frac{n-1}{k} \right\rfloor \right\rfloor$$

for all sufficiently large n satisfying

(a) $n \equiv 1 \bmod k$ *and* $n(n-1) \equiv 0 \bmod 2k^2$, *or*

(b) $n \equiv 0 \bmod k$.

(ii) *When $k = 2$,*

$$C(n,2,d) = \begin{cases} \left\lfloor \frac{n(n-1)}{4} \left\lfloor \frac{n-2}{2} \right\rfloor \right\rfloor & \text{if } d = 2, \\ \left\lfloor \frac{n}{4} \left\lfloor \frac{n-1}{2} \right\rfloor \right\rfloor - 1 & \text{if } d = 3,\ n = 6 \text{ or } n \equiv 5,7 \bmod 8, \\ \left\lfloor \frac{n}{4} \left\lfloor \frac{n-1}{2} \right\rfloor \right\rfloor & \text{if } d = 3,\ n \neq 6 \text{ and } n \not\equiv 5,7 \bmod 8. \end{cases}$$

(iii) *When $k = 3$,*

$$C(n,3,5) = \begin{cases} 2 & \text{if } n = 10, \\ \frac{n(n-1)}{18} & \text{if } n \equiv 1 \bmod 9,\ n \geqslant 19, \text{ except possibly } n \in \{64, 100, 136\}, \\ \frac{n(n-3)}{18} & \text{if } n \equiv 0 \bmod 3,\ n \geqslant 1029. \end{cases}$$

Remark. **(i)** Asymptotic results similar to Theorem 1.4(i) were known only when $k \in \{2, 3\}$. In particular, Theorem 1.4(i) verifies the conjecture of Brightwell and Katona for all $(k,d) = (k, 2k-1)$.

(ii) We determine the exact value of $C(n,2,d)$ completely. Previously, the value of $C(n,2,2)$ was unknown and $C(n,2,3)$ was determined only when $n \equiv 1$ or 9 mod 72.

(iii) The exact value of $C(n,3,5)$ is determined for n belonging to a set of density 4/9. Previously, the exact value of $C(n,3,5)$ was known only for $n \equiv 1$ or 19 mod 342, a set of density 1/171.

These results are obtained by constructing EK codes (or their equivalent combinatorial objects) whose sizes meet the Quistorff bound. This paper was presented in part at the *2013 IEEE International Symposium on Information Theory* [9].

2. EK packings and designs

Our approach is based on combinatorial design theory. In this section, we introduce necessary concepts and establish connections to EK codes.

Throughout the rest of this paper, X denotes a set of size n. For a positive integer m, $[m]$ denotes the set of integers $\{1,2,\dots,m\}$, while $\mathbb{Z}_{\geqslant m}$ denotes the set of integers at least m and $\mathbb{Z}_m$ denotes the integers modulo m. The set of all ordered k-tuples of X with distinct components is denoted by $\overline{\binom{X}{k}}$. We use angled brackets $\langle$ and $\rangle$ for multisets. We sometimes use the exponential notation to describe multisets so that a multiset where an element g_i appears s_i times, $i \in [t]$, is denoted by $g_1^{s_1} g_2^{s_2} \cdots g_t^{s_t}$.

A *set system* is a pair $\mathfrak{S} = (X, \mathcal{A})$, where X is a finite set of *points* and $\mathcal{A} \subseteq 2^X$. Elements of $\mathcal{A}$ are called *blocks*. The *order* of $\mathfrak{S}$ is the number of points in X, and the *size* of $\mathfrak{S}$ is the number of blocks in $\mathcal{A}$. Let $K \subseteq \mathbb{Z}_{\geqslant 0}$. The set system $(X, \mathcal{A})$ is said to be K-*uniform* if $|A| \in K$ for all $A \in \mathcal{A}$.

Let $2 \leqslant t < 2k$ and $0 \leqslant e \leqslant \lfloor t/2 \rfloor$. We say that the tuple

$$(x_1, x_2, \dots, x_t) \in \overline{\binom{X}{t}}$$

is (e,t)-*contained* in a set-pair $\{A, B\} \in \mathcal{E}(X,k)$ if we have either

$$\{x_1, x_2, \dots, x_e\} \subseteq A \quad \text{and} \quad \{x_{e+1}, x_{e+2}, \dots, x_t\} \subseteq B,$$

or

$$\{x_1, x_2, \dots, x_e\} \subseteq B \quad \text{and} \quad \{x_{e+1}, x_{e+2}, \dots, x_t\} \subseteq A.$$

Let $\mathcal{C} \subseteq \mathcal{E}(X,k)$. Then $(X, \mathcal{C})$ is an EK packing of *strength* t, or more precisely a t-(n,k) *EK packing*,[2] if, for every $0 \leqslant e \leqslant \lfloor t/2 \rfloor$, every t-tuple in $\overline{\binom{X}{t}}$ is (e,t)-contained in at most one set-pair in $\mathcal{C}$. A t-(n,k) *EK design* is a t-(n,k) EK packing satisfying the condition that for $e = \lfloor t/2 \rfloor$, every t-tuple in $\overline{\binom{X}{t}}$ is (e,t)-contained in exactly one set-pair in $\mathcal{C}$. It is easy to see that if $(X, \mathcal{C})$ is a t-(n,k) EK design, then

$$\begin{aligned} |\mathcal{C}| &= \frac{n(n-1)\cdots(n-t+1)}{2k(k-1)\cdots(k-\lfloor t/2 \rfloor+1)k(k-1)\cdots(k-\lceil t/2 \rceil+1)} \\ &= \frac{\dfrac{n!}{(n-t)!}}{2\dfrac{k!}{(k-\lfloor t/2 \rfloor)!}\dfrac{k!}{(k-\lceil t/2 \rceil)!}} = \frac{\dbinom{n}{t}\dbinom{t}{\lfloor t/2 \rfloor}}{2\dbinom{k}{\lfloor t/2 \rfloor}\dbinom{k}{\lceil t/2 \rceil}}. \end{aligned}$$

EK packings of strength t are equivalent to EK codes of distance $2k-t+1$, while EK designs of strength t give rise to optimal EK codes of distance $2k-t+1$.

Proposition 2.1. *Let $\mathcal{C} \subseteq \mathcal{E}(X,k)$. Then $(X, \mathcal{C})$ is a t-(n,k) EK packing if and only if $\mathcal{C}$ is an $(n,k,2k-t+1)$ EK code. Furthermore, if $(X, \mathcal{C})$ is a t-(n,k) EK design, then $\mathcal{C}$ is an optimal $(n,k,2k-t+1)$ EK code.*

[2] Note that $\mathcal{C} \subseteq \mathcal{E}(X,k)$, while $\mathcal{A} \subseteq 2^X$.

Proof. Suppose $(X, \mathcal{C})$ is a t-(n,k) EK packing and $\{A,B\}$, $\{S,T\} \in \mathcal{C}$. Then we claim that $\mathrm{d}_{\mathcal{E}}(\{A,B\},\{S,T\}) \geqslant 2k-t+1$. Suppose otherwise. Then without loss of generality, $|A \setminus S| + |B \setminus T| \leqslant 2k-t$ and there exists a non-negative $e \leqslant \lfloor t/2 \rfloor$, $I \in \binom{X}{e}$, $J \in \binom{X}{t-e}$ such that $I \subseteq A \cap S$ and $J \subseteq B \cap T$. If $I = \{x_1, x_2, \ldots, x_e\}$ and $J = \{x_{e+1}, x_{e+2}, \ldots, x_t\}$, we see that $(x_1, x_2, \ldots, x_t)$ is (e,t)-contained in $\{A,B\}$ and $\{S,T\}$, contradicting the fact that $(X, \mathcal{C})$ is a t-(n,k) EK packing.

Conversely, suppose $\mathcal{C}$ is an $(n,k,2k-t+1)$ EK code. If $(X,\mathcal{C})$ is not a t-(n,k) EK packing, then there exists a non-negative $e \leqslant \lfloor t/2 \rfloor$, $(x_1, x_2, \ldots, x_t) \in \overline{\binom{X}{t}}$, and $\{A,B\}$, $\{S,T\} \in \mathcal{C}$ such that $(x_1, x_2, \ldots, x_t)$ is (e,t)-contained in $\{A,B\}$ and $\{S,T\}$. Without loss of generality, suppose $\{x_1, x_2, \ldots, x_e\} \subseteq A \cap S$ and $\{x_{e+1}, x_{e+2}, \ldots, x_t\} \subseteq B \cap T$. Hence, $|A \setminus S| + |B \setminus T| \leqslant 2k - (e+t-e) = 2k-t$, and consequently $\mathrm{d}_{\mathcal{E}}(\{A,B\},\{S,T\}) \leqslant 2k-t$, contradicting the fact that $\mathcal{C}$ is an $(n,k,2k-t+1)$ EK code.

Finally, when $(X,\mathcal{C})$ is a t-(n,k) EK design, $\mathcal{C}$ is an optimal $(n,k,2k-t+1)$ EK code, since $|\mathcal{C}|$ meets the Quistorff bound with $e = \lfloor t/2 \rfloor$. □

In view of Proposition 2.1, our strategy in constructing optimal EK codes (and hence determining $C(n,k,d)$) is to construct equivalent EK packings and designs of sizes meeting the Quistorff bound. However, to construct the corresponding combinatorial objects is technical and complex. Here we outline the general strategy.

Section 3 determines $C(n,k,2k-1)$ for sufficiently large n when k is fixed. We used a method developed by Lamken and Wilson [17] to construct the necessary EK packings and designs. Sections 4 to 6 then address $C(n,k,d)$ for specific values of k and d. To do so, we introduce some auxiliary designs and develop a set of recursive constructions in Section 4 to build EK designs from smaller ones. In Section 5, direct methods are used to construct a large enough set of small designs on which the recursions can work to generate all larger designs.

Next we introduce EK group divisible designs and their connections to EK codes.

2.1. EK group divisible designs

Let $\mathcal{G} = \{G_1, G_2, \ldots, G_s\}$ be a partition of an n-set X and $\mathcal{C} \subseteq \mathcal{E}(X,k)$. Then $(X,\mathcal{G},\mathcal{C})$ is an *EK group divisible design* (or EKGDD for short) if, for all $(x,y) \in \overline{\binom{X}{2}}$ such that $\{x,y\} \nsubseteq G_i$ for all $i \in [s]$, we have

(i) (x,y) is $(1,2)$-contained in exactly one set-pair $\{A,B\}$ and
(ii) (x,y) is $(0,2)$-contained in at most one set-pair $\{A,B\}$.

In addition, $|G_i \cap (A \cup B)| \leqslant 1$ for all $i \in [s]$ and $\{A,B\} \in \mathcal{C}$. Such an EKGDD is more precisely called a (T,k) EKGDD, where $T = \langle |G_i| : i \in [s] \rangle$.

A 2-(n,k) EK design can be regarded as a $(1^n,k)$ EKGDD, where each group contains just a single point. Furthermore, a $(g_1 g_2 \cdots g_s, k)$ EKGDD can be regarded as a 2-$\left(\sum_{i=1}^{s} g_i, k\right)$ EK packing, and therefore also as a $\left(\sum_{i=1}^{s} g_i, k, 2k-1\right)$ EK code. In addition, as the following shows, certain classes of EKGDDs give optimal EK codes.

Proposition 2.2. *If there exists a (k^s,k) EKGDD $(X,\mathcal{G},\mathcal{C})$, then $\mathcal{C}$ is an optimal $(ks,k,2k-1)$ EK code.*

Proof. Observe that $\mathcal{C}$ is a $(ks,k,2k-1)$ EK code since $(X,\mathcal{C})$ is a 2-(ks,k) EK packing. There are $(ks)\cdot(ks-k)$ ordered pairs $(x,y)\in\overline{\binom{X}{2}}$ where $\{x,y\}$ does not belong to any group. In addition, we have $2k^2$ ordered pairs in $\overline{\binom{X}{2}}$ that are $(1,2)$-contained in each set-pair. Hence, the code $\mathcal{C}$ is of size $s(s-1)/2$, which meets the Quistorff bound. □

3. $C(n,k,2k-1)$ for sufficiently large n

We show that a 2-(n,k) EK design and a (k^n,k) EKGDD exist when n belongs to certain congruence classes, provided n is sufficiently large. Our proof is an application of decompositions of edge-coloured directed graphs (digraphs).

An *edge-coloured directed graph* is a triple $G=(V,C,E)$, where V is a finite set of *vertices*, C is a finite set of *colours* and E is a subset of $\overline{\binom{V}{2}}\times C$. Members of E are called *edges*. The *complete edge-coloured digraph* on n vertices with r colours, denoted by $K_n^{(r)}$, is the edge-coloured digraph (V,C,E), where $|V|=n$, $|C|=r$, and $E=\overline{\binom{V}{2}}\times C$. An edge-coloured digraph (V,C,E) is *simple* if there is at most one directed edge between any two ordered distinct vertices. In other words, for any ordered pair $(x,y)\in\overline{\binom{V}{2}}$, the set $\{((x,y),c)\in E : c\in C\}$ has at most one element. In particular, the complete edge-coloured digraph is not simple when $r>1$.

A family $\mathcal{F}$ of edge-coloured subgraphs of an edge-coloured digraph K is a *decomposition* of K if every edge of K belongs to exactly one member of $\mathcal{F}$. Given a family $\mathcal{G}$ of edge-coloured digraphs, a decomposition $\mathcal{F}$ of K is a *$\mathcal{G}$-decomposition of* K if each edge-coloured digraph in $\mathcal{F}$ is isomorphic to some digraph $G\in\mathcal{G}$.

Lamken and Wilson [17] studied the existence of $\mathcal{G}$-decompositions of $K_n^{(r)}$ and showed that for fixed $\mathcal{G}$ and r, a $\mathcal{G}$-decomposition exists for sufficiently large n under certain conditions. To state the theorem, we require the following concepts.

Suppose $\mathcal{G}$ is a family of edge-coloured digraphs which share the same r-colour set C. Consider an edge-coloured digraph $G=(V,C,E)\in\mathcal{G}$. Let $((u,v),c)\in E$ denote a directed edge from u to v, coloured by c. For any vertex u and colour c, define the *indegree* and *outdegree* of u with respect to c and G as follows:

$$\mathrm{in}_c(u,G) := |\{v : ((v,u),c)\in E\}|,$$
$$\mathrm{out}_c(u,G) := |\{v : ((u,v),c)\in E\}|.$$

Then for vertex u, we define the *degree vector* of u in G, denoted by $\boldsymbol{\delta}(u,G)$, to be the vector of length $2r$, such that $\boldsymbol{\delta}(u,G)=(\mathrm{in}_c(u,G),\mathrm{out}_c(u,G))_{c\in C}$. Define $\mathbf{A}^+(\mathcal{G})$ to be the set of all finite non-negative integral linear combinations of vectors in $\{\boldsymbol{\delta}(u,G) : G\in\mathcal{G}, u\in V\}$, that is,

$$\mathbf{A}^+(\mathcal{G}) := \left\{\sum_{G\in\mathcal{G}}\sum_{u\in V}\mu_{u,G}\boldsymbol{\delta}(u,G) : \mu_{u,G}\in\mathbb{Z}_{\geqslant 0}\right\}.$$

Define $A(\mathcal{G})$ to be the set of all positive integers t such that $(t,t,\ldots,t)\in\mathbf{A}^+(\mathcal{G})$. If $A(\mathcal{G})$ is non-empty, we define $\alpha(\mathcal{G})$ to be the greatest common divisor of $A(\mathcal{G})$.

On the other hand, for each $G\in\mathcal{G}$, let $m_c(G)$ be the number of edges with colour c in G and we define the *edge vector* of G, denoted by $\boldsymbol{\epsilon}(G)$, to be the vector of length r, such that $\epsilon(G) = (m_c(G))_{c\in C}$. Similarly, define $\mathbf{B}^+(\mathcal{G})$ to be the set of all finite non-negative integral linear combinations of vectors in $\{\epsilon(G) : G \in \mathcal{G}\}$. Here,

$$\mathbf{B}^+(\mathcal{G}) := \left\{\sum_{G\in\mathcal{G}} \mu_G \epsilon(G) : \mu_G \in \mathbb{Z}_{\geqslant 0}\right\}.$$

Define $B(\mathcal{G})$ to be the set of all positive integers t such that $(t, t, \ldots, t) \in \mathbf{B}^+(\mathcal{G})$. If $B(\mathcal{G})$ is non-empty, we say that $\mathcal{G}$ is *admissible*, and define $\beta(\mathcal{G})$ to be the greatest common divisor of $B(\mathcal{G})$.

Finally, we state the following theorem due to Lamken and Wilson [17].

Theorem 3.1 (Lamken and Wilson [17, Theorem 1.2]). *Let $\mathcal{G}$ be an admissible family of simple edge-coloured digraphs defined on a common set of r colours. Then there exists a constant n_0 such that a $\mathcal{G}$-decomposition of $K_n^{(r)}$ exists for all $n \geqslant n_0$ satisfying both*

$$n - 1 \equiv 0 \bmod \alpha(\mathcal{G}) \quad \textit{and} \quad n(n-1) \equiv 0 \bmod \beta(\mathcal{G}).$$

In the rest of this section, we construct families $\mathcal{G}$ of r-edge-coloured digraphs so that a $\mathcal{G}$-decomposition of $K_n^{(r)}$ yields an EK design or an EKGDD.

3.1. $C(n, k, 2k-1)$ when $n \equiv 1 \bmod k$ and $n(n-1) \equiv 0 \bmod 2k^2$

Fix $k \geqslant 2$ and define the edge-coloured digraph $G_k = (V_k, C_k, E_k)$, where

$$\begin{aligned} V_k &= \{i_j : i \in [k], j \in [2]\}, \\ C_k &= \{\circ, \bullet\} \text{ and} \\ E_k &= \{((i_r, j_s), \circ) : i, j \in [k], (r, s) \in \{(1,2), (2,1)\}\} \\ &\quad \cup \left\{((i_r, j_r), \bullet) : (i, j) \in \overline{\binom{[k]}{2}}, r \in [2]\right\}. \end{aligned}$$

In addition, define $e_\bullet := 1 \longleftrightarrow 2$, that is, the graph consisting of an edge coloured by $\bullet$. Then, define $\mathcal{G}_k := \{G_k, e_\bullet\}$.

Example. The edge-coloured graph G_2 is given by

$$\begin{tikzcd} 1_1 \arrow[r, leftrightarrow] \arrow[d, leftrightsquigarrow] \arrow[dr, leftrightsquigarrow] & 2_1 \arrow[d, leftrightsquigarrow] \arrow[dl, leftrightsquigarrow] \\ 1_2 \arrow[r, leftrightarrow] & 2_2 \end{tikzcd}$$

where $\longleftrightarrow$ denotes two directed edges of colour $\bullet$ (one in each direction), and $\leftrightsquigarrow$ denotes two directed edges of colour $\circ$ (one in each direction). Hence $\mathcal{G}_2$ consists of G_2 and $e_\bullet$.

Proposition 3.2. *If a $\mathcal{G}_k$-decomposition of $K_n^{(2)}$ exists, then a 2-(n, k) EK design exists.*

Proof. Let $\mathcal{F}$ be a $\mathcal{G}_k$-decomposition of $K_n^{(2)}$. For a subgraph $G \in \mathcal{F}$ that is isomorphic to G_k, consider the graph isomorphism $\phi_G : G_k \to G$. Define

$$A_G = \{\phi_G(i_1) : i \in [k]\} \text{ and } \quad B_G = \{\phi_G(i_2) : i \in [k]\}.$$

Let X be the vertex set of $K_n^{(2)}$ and

$$\mathcal{C} = \{\{A_G, B_G\} : G \in \mathcal{F} \text{ and } G \text{ isomorphic to } G_k\}.$$

We claim that $(X, \mathcal{C})$ is a 2-(n,k) EK design. Counting the number of edges coloured by $\circ$, we have $|\mathcal{C}| = n(n-1)/2k^2$. Hence, it suffices to check that for $e \in \{0,1\}$, each $(x,y) \in \overline{\binom{X}{2}}$ is $(e,2)$-contained in at most one set-pair in $\mathcal{C}$.

Suppose otherwise. Then there exist $(x,y) \in \overline{\binom{X}{2}}$, $G, H \in \mathcal{F}$ and $e \in \{0,1\}$ such that (x,y) is $(e,2)$-contained in $\{A_G, B_G\}$ and $\{A_H, B_H\}$.

If $e = 0$, then assume that $\{x,y\} \subset A_G \cap A_H$. So, the edge $((x,y),\bullet)$ belongs to both G and H, contradicting the fact that $\mathcal{F}$ is a $\mathcal{G}_k$-decomposition of $K_n^{(2)}$.

If $e = 1$, then assume that $x \in A_G \cap A_H$ and $y \in B_G \cap B_H$. Hence, the edge $((x,y),\circ)$ belongs to G and H, contradicting the fact that $\mathcal{F}$ is a $\mathcal{G}_k$-decomposition of $K_n^{(2)}$. $\square$

Finally, we compute $\alpha(\mathcal{G}_k)$ and $\beta(\mathcal{G}_k)$. Observe that

$$\begin{aligned}\boldsymbol{\delta}(1_1, G_k) + \boldsymbol{\delta}(1, e_\bullet) &= (k,k,k-1,k-1) + (0,0,1,1) = (k,k,k,k), \quad \text{and}\\ \boldsymbol{\epsilon}(G_k) + k\boldsymbol{\epsilon}(e_\bullet) &= (2k^2, 2k(k-1)) + k(0,2) = (2k^2, 2k^2).\end{aligned}$$

Then $A(\mathcal{G}_k) = \{kt : t \in \mathbb{Z}_{\geqslant 1}\}$ and $B(\mathcal{G}_k) = \{2k^2 t : t \in \mathbb{Z}_{\geqslant 1}\}$. Hence,

$$\alpha(\mathcal{G}_k) = k \quad \text{and} \quad \beta(\mathcal{G}_k) = 2k^2.$$

The following is then immediate from Propositions 2.1 and 3.2, and Theorem 3.1.

Theorem 3.3. *Fix $k \geqslant 2$. Then*

$$C(n,k,2k-1) = \frac{n(n-1)}{2k^2}$$

for all sufficiently large n satisfying $n \equiv 1 \bmod k$ and $n(n-1) \equiv 0 \bmod 2k^2$.

3.2. $C(n,k,2k-1)$ when $n \equiv 0 \bmod k$

Fix $k \geqslant 2$ and define the edge-coloured digraph $H_k = (V_k, C_k, E_k)$, where

$$\begin{aligned}V_k &= \{i_j : i \in [k], j \in [2]\},\\ C_k &= \big([k] \times [k] \times \{\circ\}\big) \cup \left(\overline{\binom{[k]}{2}} \times \{\bullet\}\right) \text{ and}\\ E_k &= \{((i_r, j_s), (i,j,\circ)) : i,j \in [k], (r,s) \in \{(1,2),(2,1)\}\}\\ &\quad \cup \left\{((i_r, j_r), (i,j,\bullet)) : (i,j) \in \overline{\binom{[k]}{2}}, r \in [2]\right\}.\end{aligned}$$

Then, define $\mathcal{H}_k := \{H_k\}$.

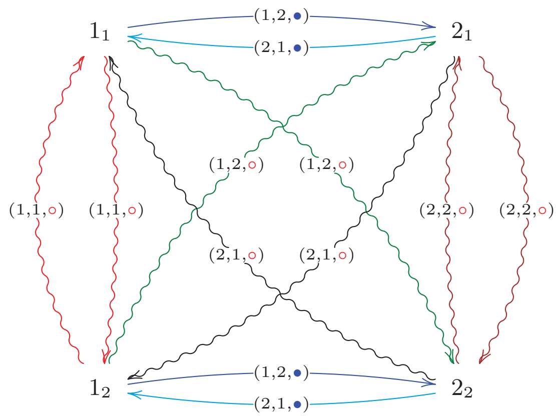

Figure 1. (Colour online) An auxiliary edge-colored digraph H_2. An $\{H_2\}$-decomposition of $K_n^{(6)}$ implies the existence of a $(2^n, 2)$ EKGDD.

Example. The family $\mathcal{H}_2$ comprises only the digraph H_2, given in Figure 1.

Proposition 3.4. *If an $\mathcal{H}_k$-decomposition of $K_n^{(2k^2-k)}$ exists, then a (k^n, k) EKGDD exists.*

Proof. Let $\mathcal{F}$ be an $\mathcal{H}_k$-decomposition of $K_n^{(2k^2-k)}$. Suppose V is the vertex set of $K_n^{(2k^2-k)}$ and we define the point set X to be $\{v_i : v \in V, i \in [k]\}$.

Then, for a subgraph $H \in \mathcal{F}$, let $\phi_H : H_k \to H$ be a graph isomorphism and let

$$A_H = \{\phi_H(i_1)_i : i \in [k]\}, \quad B_H = \{\phi_H(i_2)_i : i \in [k]\}.$$

Then define

$$\mathcal{G} = \{\{v_i : i \in [k]\} : v \in V\},$$
$$\mathcal{C} = \{\{A_H, B_H\} : H \in \mathcal{F}\},$$

and we claim that $(X, \mathcal{G}, \mathcal{C})$ is a (k^n, k) EKGDD. Suppose otherwise. Since $|\mathcal{C}| = n(n-1)/2$, it suffices to consider the following two cases.

(i) Suppose a group in $\mathcal{G}$ intersect a set-pair in $\mathcal{C}$ at least two points. In other words, there exist $v \in V$ and $H \in \mathcal{F}$ such that $|\{v_i : i \in [k]\} \cap (A_H \cup B_H)| \geqslant 2$. This contradicts the fact that H is isomorphic to $\underline{H_k}$.

(ii) Suppose there exist $(x, y) \in \overline{\binom{V}{2}}$, $G, H \in \mathcal{F}$ and $e \in \{0, 1\}$ such that (x_i, y_j) is $(e, 2)$-contained in $\{A_G, B_G\}$ and $\{A_H, B_H\}$.

If $e = 0$, then assume that $\{x_i, y_j\} \subset A_G \cap A_H$. Hence, the edge $((x, y), (i, j, \bullet))$ belongs to both G and H, contradicting the fact that $\mathcal{F}$ is an $\mathcal{H}_k$-decomposition.

Similarly, if $e = 1$, then assume that $x_i \in A_G \cap A_H$ and $y_j \in B_G \cap B_H$. Hence, the edge $((x, y), (i, j, \circ))$ belongs to both G and H, contradicting the fact that $\mathcal{F}$ is an $\mathcal{H}_k$-decomposition of $K_n^{(2k^2-k)}$. $\square$

As before, we observe that

$$\sum_{i\in[k]} \delta(i_1, H_k) = (1,1,\ldots,1), \quad \text{and}$$
$$\epsilon(H_k) = (2,2,\ldots,2).$$

Then $A(\mathcal{H}_k) = \{t : t \in \mathbb{Z}_{\geqslant 1}\}$ and $B(\mathcal{H}_k) = \{2t : t \in \mathbb{Z}_{\geqslant 1}\}$. Hence,

$$\alpha(\mathcal{G}_k) = 1 \quad \text{and} \quad \beta(\mathcal{G}_k) = 2.$$

From Propositions 2.2, 3.4, and Theorem 3.1, we have the following.

Theorem 3.5. *Fix $k \geqslant 2$. Then*

$$C(nk, k, 2k-1) = \frac{n(n-1)}{2}$$

for all sufficiently large n.

Theorems 3.3 and 3.5 combine to give Theorem 1.4(i).

4. Recursive constructions

This section introduces certain auxiliary designs and gives the necessary recursive constructions so that we can build bigger EK designs and EKGDDs from small ones.

Proposition 4.1 (Filling in groups). *If a $(g_1 g_2 \cdots g_s, k)$ EKGDD exists and a $(1^{g_i} t, k)$ EKGDD exists for every $i \in [s]$, then a $\left(1^{\sum_{i=1}^s g_i} t, k\right)$ EKGDD exists.*

Proof. Let $(X, \{G_1, G_2, \ldots, G_s\}, \mathcal{C})$ be a $(g_1 g_2 \cdots g_s, k)$ EKGDD and

$$H = \{\infty_1, \infty_2, \ldots, \infty_t\}.$$

For $i \in [s]$, let $(G_i \cup H, \{\{x\} : x \in G_i\} \cup \{H\}, \mathcal{C}_i)$ be a $(1^{g_i} t, k)$ EKGDD.
Consider

$$\begin{aligned} X^* &= X \cup H, \\ \mathcal{G}^* &= \{\{x\} : x \in X\} \cup \{H\} \text{ and} \\ \mathcal{C}^* &= \mathcal{C} \cup \left(\bigcup_{i=1}^{s} \mathcal{C}_i\right). \end{aligned}$$

Then $(X^*, \mathcal{G}^*, \mathcal{C}^*)$ is a $(1^{\sum_{i=1}^s g_i} t, k)$ EKGDD. □

When $t = 1$, we obtain an EK design.

Corollary 4.2. *If a $(g_1 g_2 \cdots g_s, k)$ EKGDD exists and a 2-$(1 + g_i, k)$ EK design exists for all $i \in [s]$, then a 2-$(1 + \sum_{i=1}^s g_i, k)$ EK design exists.*

A similar construction for EK packings holds.

Proposition 4.3. *Suppose a $(g_1g_2\cdots g_st,k)$ EKGDD exists. If a 2-$(1+g_i,k)$ EK design exists for $i\in[s]$ and a 2-$(1+t,k)$ EK packing of size $\lfloor(1+t)/2k\lfloor t/k\rfloor\rfloor-1$ exists, then there exists a 2-(N,k) EK packing of size $\lfloor N/2k\lfloor(N-1)/k\rfloor\rfloor-1$, where $N=1+t+\sum_{i=1}^{s}g_i$. If $g_1=g_2=\cdots=g_s=1$ and there exists a 2-(t,k) EK packing of size $\lfloor t/2k\lfloor(t-1)/k\rfloor\rfloor-1$, then there exists a 2-$(N,k)$ EK packing of size $\lfloor N/2k\lfloor(N-1)/k\rfloor\rfloor-1$, where $N=s+t$.*

For the above propositions to be useful, we require large classes of EKGDDs. In the next two subsections, we discuss the recursive constructions for EKGDDs.

4.1. *t*-wise balanced designs, group divisible designs and transversal designs

Instead of defining new combinatorial objects, we introduce classical combinatorial objects to aid in our recursive constructions. The existence results of the latter objects are well known and we build on these results to construct our desired EK designs and EKGDDs.

Definition. A *t-wise balanced design*, or a t-BD(v,K), is a K-uniform set system $(X,\mathcal{A})$ of order v such that every t-subset of X is contained in exactly one block of $\mathcal{A}$.

The following existence results hold for t-BDs [1,15].

Theorem 4.4. *Let $BD_t(K)$ denote the set of positive integers v such that there exists a t-BD(v,K). Then we have*

(i) $\mathrm{BD}_2(\{4,5,6\})\supseteq\mathbb{Z}_{\geqslant 13}\setminus\{14,15,18,19,23\}$,
(ii) $\mathrm{BD}_2(\{7,8,9\})\supseteq\mathbb{Z}_{\geqslant 343}$ *and*
(iii) $\mathrm{BD}_3(\{4,6\})\supseteq 2\mathbb{Z}_{\geqslant 2}$.

Definition. Let $(X,\mathcal{A})$ be a set system and let $\mathcal{G}=\{G_1,G_2,\ldots,G_s\}$ be a partition of X into subsets, called *groups*. The triple $(X,\mathcal{G},\mathcal{A})$ is a *group divisible design* (GDD) when every 2-subset of X not contained in a group appears in exactly one block, and $|A\cap G|\leqslant 1$ for all $A\in\mathcal{A}$ and $G\in\mathcal{G}$.

We denote a GDD $(X,\mathcal{G},\mathcal{A})$ by K-GDD if $(X,\mathcal{A})$ is K-uniform. The *type* of a GDD $(X,\mathcal{G},\mathcal{A})$ is the multiset $\langle|G|:G\in\mathcal{G}\rangle$. A 2-BD$(v,K)$ can be regarded as a K-GDD of type 1^v, where each group contains a single point.

Definition. A *transversal design* TD(k,m) is a $\{k\}$-GDD of type m^k.

The following result on the existence of transversal designs (see [2]) is used without explicit reference in this paper.

Theorem 4.5. *Let* TD(k) *denote the set of positive integers m such that there is a* TD(k,m). *Then we have*

(i) TD$(5)\supseteq 4\mathbb{Z}_{\geqslant 1}$,

(ii) $\mathrm{TD}(8) \supseteq 3\mathbb{Z}_{\geqslant 23}$ *and*
(iii) $\mathrm{TD}(k) \supseteq \{q : q \geqslant k-1$ *is a prime power*$\}$.

In addition, we require transversal designs with disjoint blocks, and the following proposition (see [2]) is useful.

Proposition 4.6. *Suppose there exists a* $\mathrm{TD}(k+1, m)$. *Then there exists a* $\mathrm{TD}(k, m)$ *with* m *disjoint blocks.*

4.2. Recursive constructions for EKGDDs
In this subsection, we give four recursive constructions for EKGDDs.

Proposition 4.7 (Inflation). *Suppose there exist a* (T, k) *EKGDD and a* $\mathrm{TD}(2k, m)$. *Then an* (mT, k) *EKGDD exists.*

Proof. Let $(X, \mathcal{G}, \mathcal{C})$ be a (T, k) EKGDD. For $C = \{\{a_1, a_2, \ldots, a_k\}, \{b_1, b_2, \ldots b_k\}\} \in \mathcal{C}$, let $X_C = \{a_1, a_2, \ldots, a_k, b_1, b_2, \ldots b_k\}$ and $(X_C \times [m], \mathcal{A}_C)$ be a $\mathrm{TD}(2k, m)$ with groups $\{\{x\} \times [m] : x \in X_C\}$. Given a block

$$A = \{(a_1, i_1), (a_2, i_2), \ldots, (a_k, i_k), (b_1, j_1), (b_2, j_2), \ldots, (b_k, j_k)\} \quad \text{in } \mathcal{A}_C,$$

construct the set-pair

$$\phi(A, C) = \{\{(a_1, i_1), (a_2, i_2), \ldots, (a_k, i_k)\}, \{(b_1, j_1), (b_2, j_2), \ldots, (b_k, j_k)\}\}.$$

Let $\mathcal{A}_C^* = \{\phi(A, C) : A \in \mathcal{A}_C\}$ and consider

$$\begin{aligned} X^* &= X \times [m], \\ \mathcal{G}^* &= \{G \times [m] : G \in \mathcal{G}\} \text{ and} \\ \mathcal{C}^* &= \bigcup_{C \in \mathcal{C}} \mathcal{A}_C^*. \end{aligned}$$

Then $(X^*, \mathcal{G}^*, \mathcal{C}^*)$ is an (mT, k) EKGDD. □

Wilson's Fundamental Construction for GDDs [21] can also be modified to give the following recursions for EKGDDs. This construction is described in Table 2.

Proposition 4.8 (Fundamental Construction). *Suppose a (master) GDD* $(X, \mathcal{G}, \mathcal{A})$ *of type* T *exists and let* $w : X \to \mathbb{Z}_{\geqslant 0}$ *be a weight function. If, for each* $A \in \mathcal{A}$, *there exists an (ingredient)* $(\langle w(a) : a \in A\rangle, k)$ *EKGDD, then there exists a* $(\langle \sum_{x \in G} w(x) : G \in \mathcal{G}\rangle, k)$ *EKGDD.*

Proof. The Fundamental Construction in Table 2 constructs the desired EKGDD from the master GDD and ingredient EKGDD. □

Proposition 4.8 admits the following specializations.

Table 2. Fundamental Construction for EKGDDs.

Input:	(master) GDD $\mathcal{D} = (X, \mathcal{G}, \mathcal{A})$;
	weight function $w : X \to \mathbb{Z}_{\geqslant 0}$;
	(ingredient) (T_A, k) EKGDD $(X_A, \mathcal{G}_A, \mathcal{C}_A)$ for each $A \in \mathcal{A}$, where
	$T_A = \langle w(x) : x \in A \rangle$,
	$X_A = \cup_{x \in A}(\{x\} \times [w(x)])$,
	$\mathcal{G}_A = \{\{x\} \times [w(x)] : x \in A\}$,
Output:	$(\langle \sum_{x \in G} w(x) : G \in \mathcal{G} \rangle, k)$ EKGDD $(X^*, \mathcal{G}^*, \mathcal{C}^*)$, where
	$X^* = \cup_{x \in X}(\{x\} \times [w(x)])$,
	$\mathcal{G}^* = \{\cup_{x \in G}(\{x\} \times [w(x)]) : G \in \mathcal{G}\}$,
	$\mathcal{C}^* = \cup_{A \in \mathcal{A}} \mathcal{C}_A$, and
Note:	By convention, for $x \in X$, $\{x\} \times [w(x)] = \varnothing$ if $w(x) = 0$.

Proposition 4.9 (PBD closure). *Let $K \subset \mathbb{Z}_{\geqslant 1}$. Suppose a 2-BD$(v, K)$ exists and a (g^t, k) EKGDD exists for all $t \in K$. Then there exists a (g^v, k) EKGDD.*

Proof. Let $(X, \mathcal{A})$ be a 2-BD(v, K). Consider the 2-BD as a (master) GDD of type 1^v and weight function $w(x) = g$ for all $x \in X$. Now apply the Fundamental Construction. □

Proposition 4.10 (EKGDD from truncated TD). *Let r, s be non-negative integers. Suppose a* TD$(u + r + s, m)$ *exists with r disjoint blocks and $g_1, g_2, \ldots, g_s$ are non-negative integers at most $m - r$. If a (g^t, k) EKGDD exists for each $t \in \{u, u + 1, \ldots, u + r + s\}$, then a (T, k) EKGDD exists, where $T = (g(m - r))^{u+r}(gg_1)(gg_2)\cdots(gg_s)$.*

Proof. Delete the points in the r disjoint blocks from the point set of TD$(u + r + s, m)$ so that each group is of size $m - r$. Subsequently, for $i \in [s]$, delete points from the ith group of the TD$(u + r + s, m)$ so that g_i points remain. Observe that the remaining $u + r$ groups are of size $m - r$. This therefore results in a $\{u, u + 1, \ldots, u + r + s\}$-GDD of type

$$(m - r)^{u+r} g_1 g_2 \cdots g_s.$$

Use this as the master GDD and apply the Fundamental Construction with weight function w that assigns weight g to all points. □

4.3. Recursive constructions for EK designs of strength t

The above constructions in general yield EK designs of strength two. The following construction gives t-(n, k) EK designs for general t using t-BDs.

Proposition 4.11 (Filling in t-BDs). *Let $K \subset \mathbb{Z}_{\geqslant 1}$ and suppose a t-BD(v, K) $(X, \mathcal{A})$ exists. If there exists a t-(h, k) EK design for all $h \in K$, then there exists a t-(v, k) EK design.*

Proof. Let $(X, \mathcal{A})$ be a t-BD(v, K). For $A \in \mathcal{A}$, let $(A, \mathcal{C}_A)$ be a t-$(|A|, k)$ EK design. Let

$$\mathcal{C} = \bigcup_{A \in \mathcal{A}} \mathcal{C}_A,$$

and then $(X, \mathcal{C})$ is a t-(v, k) EK design. □

5. Direct constructions

We construct some small EK designs and EKGDDs that are required to seed the recursive constructions in the previous section. In general, the objects are constructed via *direct computer search* or *method of differences*. In summary, the following objects are constructed in this section.

Proposition 5.1. *The following objects exist:*

(i) *a* 3-$(n,2)$ *EK design for* $n \in \{4,6\}$,
(ii) *a* 2-$(n,2)$ *EK design for* $n \in \{9,17,25\}$, *and*
a 2-$(n,3)$ *EK design for* $n \in \{19,28,37,46,55,73,82,91,109,118,127,145,163,181,199\}$,
(iii) *a* $(2^u,2)$ *EKGDD for* $u \in \{4,5,6,7,8,9,10,11,12,14,15,18,19, 23\}$,
a $(1^ut,2)$ *EKGDD for* $(u,t) \in \{(8,3),(16,3),(16,7),(24,3),(24,5),(24,7)\}$,
an $(8^ut,2)$ *EKGDD for* $u \in \{6,7\}$, $t \in \{4,6\}$,
a $(3^u,3)$ *EKGDD for* $u \in \{7,8,9,10,11\}$, *and*
an $(18^u27,3)$ *EKGDD for* $u \in \{7,8\}$,
(iv) *a* 2-$(4,2)$ *EK packing of size one, and*
a 2-$(n,2)$ *EK packing of size* $\lfloor n(n-1)/8\rfloor - 1$ *for* $n \in \{5,6,7,13,15,21\}$.

For small values of n or u, we construct the following combinatorial objects via a direct computer search. The corresponding objects are recorded in [8].

Proposition 5.2. *The following objects exist:*

(i) *a* 3-$(n,2)$ *EK design for* $n \in \{4,6\}$,
(ii) *a* $(2^4,2)$ *EKGDD, and*
a $(1^ut,2)$ *EKGDD for* $(u,t) \in \{(8,3),(16,3),(16,7)\}$,
(iii) *a* 2-$(4,2)$ *EK packing of size one, and*
a 2-$(n,2)$ *EK packing of size* $\lfloor n(n-1)/8\rfloor - 1$ *for* $n \in \{5,6,7,13\}$.

5.1. Method of differences

In general, when n or u is large, a direct computer search is unable to find a combinatorial object of the required size. Hence, we impose a certain algebraic structure on our combinatorial object and then search for a *smaller* set of set-pairs satisfying a new set of requirements. In design theory, this method is known as the *method of differences*.

Let Γ be an additive abelian group. Given $\mathcal{D} \subseteq \mathcal{E}(\Gamma,k)$, the *outer difference list* of $\mathcal{D}$ is the multiset

$$\Delta^{\text{outer}}\mathcal{D} = \langle x-y : x \in A, y \in B, \text{ or } x \in B, y \in A, \text{ and } \{A,B\} \in \mathcal{D}\rangle,$$

while the *inner difference list* of $\mathcal{D}$ is the multiset

$$\Delta^{\text{inner}}\mathcal{D} = \langle x-y : x,y \in A \text{ or } B,\ x \neq y \text{ and } \{A,B\} \in \mathcal{D}\rangle.$$

In addition, let t be a positive integer. For $\mathcal{D} \subseteq \mathcal{E}(\Gamma \times [t], k)$, we define for $i, j \in [t]$

$$\Delta_{ij}^{\text{outer}}\mathcal{D} = \langle x - y : x_i \in A, y_j \in B, \text{ or } x_i \in B, y_j \in A, \text{ and } \{A, B\} \in \mathcal{D}\rangle,$$
$$\Delta_{ij}^{\text{inner}}\mathcal{D} = \langle x - y : x_i, y_j \in A \text{ or } B,\ x_i \neq y_j \text{ and } \{A, B\} \in \mathcal{D}\rangle.$$

The multiset is called a list of *pure outer/inner differences* when $i = j$, and called a list of *mixed outer/inner differences* when $i \neq j$. Note that when $t = 1$, the difference lists $\Delta_{11}^{\text{outer}}$, $\Delta_{11}^{\text{inner}}$ are the same as Δ^{outer}, Δ^{inner}, respectively. Hence, we use the latter notation and also Γ, instead of $\Gamma \times [1]$.

In this section, we construct $\mathcal{D} \subseteq \mathcal{E}(\Gamma \times [t], k)$ whose difference lists satisfy certain conditions. Then we obtain EK designs and EKGDDs $(\Gamma \times [t], \mathcal{C})$ by *developing* $\mathcal{D}$ over Γ. We state this formally in the subsequent sections.

Adopt the following notations. Let $\{A, B\} \in \mathcal{E}(\Gamma \times [t], k)$ and $\gamma \in \Gamma$. Then

$$\{A, B\} + \gamma := \{\{(x + \gamma)_i : x_i \in A\}, \{(y + \gamma)_j : y_j \in B\}\}.$$

Sometimes we include *infinite elements* so that $X = (\Gamma \times [t]) \cup (\{\infty\} \times [s])$, where $s > 0$. Then, for $\infty_i \in X$,

$$\infty_i + \gamma = \infty_i \quad \text{for all } \gamma \in \Gamma.$$

If $\mathcal{D} \subseteq \mathcal{E}(X, k)$, then

$$\Delta_{ij}^{\text{inner}}\mathcal{D} = \langle x - y : x, y \in \Gamma,\ x_i, y_j \in A \text{ or } B,\ \{A, B\} \in \mathcal{D}\rangle \quad \text{for } i, j \in [t].$$

The difference list $\Delta_{ij}^{\text{outer}}\mathcal{D}$ is similarly defined.

5.2. Direct constructions for EK designs

Definition. Let $k > 0$ and Γ be an abelian group of odd size m. Let $t > 0$ such that $t(mt - 1) \equiv 0 \bmod 2k^2$. Suppose $\mathcal{D} \subseteq \mathcal{E}(\Gamma \times [t], k)$ with size $t(mt - 1)/2k^2$. Then $(\Gamma \times [t], \mathcal{D})$ is a $(\Gamma \times [t])$-*base-set for a* 2-(mt, k) *EK design* if the following conditions hold.

(i) For $i, j \in [t]$,

$$\Delta_{ij}^{\text{outer}}\mathcal{D} = \begin{cases} \Gamma \setminus \{0\} & \text{if } i = j, \\ \Gamma & \text{otherwise.} \end{cases}$$

(ii) Fix $i, j \in [t]$. If $i = j$, then each element in $\Gamma \setminus \{0\}$ appears at most once in $\Delta_{ij}^{\text{inner}}\mathcal{D}$. Otherwise, each element in Γ appears at most once in $\Delta_{ij}^{\text{inner}}\mathcal{D}$.

Proposition 5.3. *If a* $(\Gamma \times [t])$-*base-set for a* 2-(mt, k) *EK design exists, then a* 2-(mt, k) *EK design exists.*

Proof. Let $X = \Gamma \times [t]$, $(X, \mathcal{D})$ be the X-base-set and

$$\mathcal{C} = \{D + \gamma : D \in \mathcal{D}, \gamma \in \Gamma\}.$$

Then $(X, \mathcal{C})$ is a 2-(mt, k) EK design. □

Example. Let $X = \mathbb{Z}_{19}$ and $k = 3$. Suppose $\mathcal{D} = \{\{0,1,4\},\{3,8,14\}\}$. We check that

$$\Delta^{\text{outer}}\mathcal{D} = X \setminus \{0\} \text{ and}$$
$$\Delta^{\text{inner}}\mathcal{D} = \langle 1,3,4,5,6,8,11,13,14,15,16,18\rangle.$$

Hence, $(X,\mathcal{D})$ is a $\mathbb{Z}_{19}$-base-set for a 2-(19,3) EK design. In fact, developing $\mathcal{D}$ over $\mathbb{Z}_{19}$ yields the EK code $\mathcal{C}$ given in the first example.

Corollary 5.4. *A* 2-$(n,2)$ *EK design exists for* $n \in \{9,17,25\}$.

Proof. For $m \in \{9,17,25\}$, the required $\mathbb{Z}_m$-base-sets for 2-$(m,2)$ EK designs are given in [8]. □

Corollary 5.5. *A* 2-$(n,3)$ *EK design exists for* $n \in \{19,28,37,46,55,73,82,91,109,118,127,145,163,181,199\}$.

Proof. The required $\mathbb{Z}_m$-base-sets for 2-$(m,3)$ EK designs for $m \in \{19,37,55,73,91,109,127,145,163,181,199\}$, $(\mathbb{Z}_m \times [2])$-base-sets for 2-$(2m,3)$ EK designs for $m \in \{23,41,59\}$ and a $(\mathbb{Z}_7 \times [4])$-base-set for a 2-(28,3) EK design are given in [8]. □

5.3. Direct constructions for EKGDDs and EK-packings

Definition. Let $k > 0$ and Γ be an abelian group of odd size m. Suppose $\mathcal{D} \subseteq \mathcal{E}(\Gamma \times [k], k)$ with size $(m-1)/2$. Then $(\Gamma \times [k], \mathcal{D})$ is a $(\Gamma \times [k])$-*base-set for a* (k^m, k) *EKGDD* if the following conditions hold.

(i) Fix $i,j \in [k]$. Then $\Delta_{ij}^{\text{outer}}\mathcal{D} = \Gamma \setminus \{0\}$.
(ii) Fix $i,j \in [k]$. Then each element in Γ appears at most once in $\Delta_{ij}^{\text{inner}}\mathcal{D}$ and the zero element does not appear in $\Delta_{ij}^{\text{inner}}\mathcal{D}$.

Proposition 5.6. *If a* $(\Gamma \times [k])$-*base-set for a* (k^m,k) *EKGDD exists, then a* (k^m,k) *EKGDD exists.*

Proof. Let $(\Gamma \times [k], \mathcal{D})$ be the $(\Gamma \times [k])$-base-set and

$$\mathcal{G} = \{\{\gamma\} \times [k] : \gamma \in \Gamma\},$$
$$\mathcal{C} = \{D + \gamma : D \in \mathcal{D}, \gamma \in \Gamma\}.$$

Then $(\Gamma \times [k], \mathcal{G}, \mathcal{C})$ is a (k^m,k) EKGDD. □

Definition. Let $k > 0$ and Γ be an abelian group of odd size m. Suppose $\mathcal{D} \subseteq \mathcal{E}((\Gamma \cup \{\infty\}) \times [k], k)$ of size $(m+1)/2$. Then $((\Gamma \cup \{\infty\}) \times [k], \mathcal{D})$ is a $((\Gamma \cup \{\infty\}) \times [k])$-*base-set for a* (k^{m+1},k) *EKGDD* if the following conditions hold.

(i) $\mathcal{D}$ contains k set-pairs, $\{A_1,B_1\}$, $\{A_2,B_2\}$, ... , $\{A_k,B_k\}$, such that A_i contains ∞_i and $\{j : b_j \in B_i\} = [k]$ for $i \in [k]$.
(ii) Fix $i,j \in [k]$. Then $\Delta_{ij}^{\text{outer}}\mathcal{D} = \Gamma \setminus \{0\}$.

(iii) Fix $i, j \in [k]$. Then each element in Γ appears at most once in $\Delta_{ij}^{\text{inner}}\mathcal{D}$ and the zero element does not appear in $\Delta_{ij}^{\text{inner}}\mathcal{D}$.

Proposition 5.7. *Suppose there exists a $((\Gamma \cup \{\infty\}) \times [k])$-base-set for a (k^{m+1}, k) EKGDD. Then a (k^{m+1}, k) EKGDD exists.*

Proof. Let $X = (\Gamma \cup \{\infty\}) \times [k]$, let $(X, \mathcal{D})$ be the X-base-set, and

$$\mathcal{G} = \{\{\gamma\} \times [k] : \gamma \in \Gamma\} \cup \{\{\infty\} \times [k]\},$$
$$\mathcal{C} = \{D + \gamma : D \in \mathcal{D}, \gamma \in \Gamma\}.$$

Then $(X, \mathcal{G}, \mathcal{C})$ is a (k^{m+1}, k) EKGDD. □

Corollary 5.8. *A $(2^u, 2)$ EKGDD exists for $u \in \{5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 18, 19, 23\}$.*

Proof. The $(\mathbb{Z}_m \times [2])$-base-sets for $(2^m, 2)$ EKGDD for $m \in \{5, 7, 9, 11, 15, 19, 23\}$ and the $((\mathbb{Z}_m \cup \{\infty\}) \times [2])$-base-sets for $(2^{m+1}, 2)$ EKGDD for $m \in \{5, 7, 9, 11, 13, 17\}$ are given in [8]. □

Corollary 5.9. *A $(3^u, 3)$ EKGDD exists for $u \in \{7, 8, 9, 11\}$.*

Proof. The required $(\mathbb{Z}_m \times [3])$-base-sets for $(3^m, 3)$ EKGDDs for $m \in \{7, 11\}$, the $((\mathbb{Z}_7 \cup \{\infty\}) \times [3])$-base-set for a $(3^8, 3)$ EKGDD and $(\mathbb{Z}_3 \times \mathbb{Z}_3 \times [3])$-base-set for a $(3^9, 3)$ EKGDD are given in [8]. □

To end this section, we construct some EKGDDs and EK packings that are required to seed the recursion techniques. In general, the construction uses the method of differences. However, the detailed construction is *ad hoc* and we refer to the interested reader to [8].

Proposition 5.10. *The following EKGDDs exist:*

(i) *a $(1^{24}t, 2)$ EKGDD for $t \in \{3, 5, 7\}$,*
(ii) *an $(8^u t, 2)$ EKGDD for $u \in \{6, 7\}$ and $t \in \{4, 6\}$,*
(iii) *a $(3^{10}, 3)$ EKGDD,*
(iv) *an $(18^u 27, 3)$ EKGDD for $u \in \{7, 8\}$.*

Proposition 5.11. *There is a 2-$(n, 2)$ EK packing of size $\lfloor n(n-1)/8 \rfloor - 1$ for $n \in \{15, 21\}$.*

6. $C(n,k,d)$ for $(k,d) \in \{(2,2), (2,3), (3,5)\}$

We apply recursive constructions in Section 4 with the objects constructed in Section 5 to prove Theorems 1.4(ii) and 1.4(iii). For purposes of exposition, we rewrite Theorems 1.4(ii) and 1.4(iii) as the following set of equations:

$$C(n, 2, 2) = \begin{cases} n(n-1)(n-2)/8 & \text{if } n \equiv 0 \bmod 2, \quad (6.1) \\ n(n-1)(n-3)/8 & \text{otherwise}, \quad (6.2) \end{cases}$$

$$C(n,2,3)=\begin{cases} 2 & \text{if } n=6, \\ n(n-2)/8 & \text{if } n\equiv 0 \bmod 2,\ n\neq 6, \qquad (6.3)\\ \lfloor n(n-1)/8\rfloor & \text{if } n\equiv 1,3 \bmod 8, \qquad (6.4)\\ \lfloor n(n-1)/8\rfloor-1 & \text{if } n\equiv 5,7 \bmod 8, \qquad (6.5)\end{cases}$$

$$C(n,3,5)=\begin{cases} 2 & \text{if } n=10, \\ n(n-3)/18 & \text{if } n\equiv 0 \bmod 3,\ n\geqslant 1029, \qquad (6.6)\\ n(n-1)/18 & \text{if } n\equiv 1 \bmod 9,\ n\notin\{10,64,100,136\}. \qquad (6.7)\end{cases}$$

Note that we determine $C(6,2,3)=2$ and $C(10,3,5)=2$ via an exhaustive computer search.

6.1. Determining $C(n,2,2)$

Lemma 6.1. *There exists a* 3-$(n,2)$ *EK design for* $n\geqslant 4$ *and* $n\equiv 0 \bmod 2$.

Proof. By Proposition 5.1(i), a 3-$(n,2)$ EK design exists for $n\in\{4,6\}$. For $n\geqslant 8$, there exists a 3-BD$(n,\{4,6\})$ by Theorem 4.4. The lemma follows from Proposition 4.11. □

Equation (6.1) follows from Theorem 1.2, Proposition 2.1 and Lemma 6.1.

Lemma 6.2. *There exists a* 3-$(n,2)$ *EK packing of size* $n(n-1)(n-3)/8$ *for* $n\geqslant 5$ *and* $n\equiv 1 \bmod 2$.

Proof. By Lemma 6.1, there exists a 3-$(n+1,2)$ EK design $(X,\mathcal{C})$. Fix any point $x\in X$. Define $X'=X\setminus\{x\}$ and $\mathcal{C}'=\{\{A,B\}\in\mathcal{C} : x\notin A\cup B\}$.

Since x is contained in exactly $n(n-1)/2$ set-pairs in $\mathcal{C}$, then

$$|\mathcal{C}'|=(n+1)n(n-1)/8-n(n-1)/2=n(n-1)(n-3)/8.$$

Hence, $(X',\mathcal{C}')$ is the required EK packing. □

Equation (6.2) follows from Theorem 1.2, Proposition 2.1 and Lemma 6.2.

6.2. Determining $C(n,2,3)$

Lemma 6.3. *There exists a* $(2^m,2)$ *EKGDD and an* $(8^m,2)$ *EKGDD for* $m\geqslant 4$.

Proof. By Proposition 5.1(iii), there exists a $(2^m,2)$ EKGDD for $m\in\{4,5,6,7,8,9,10,11,12,14,15,18,19,23\}$. For other values of m, we apply Proposition 4.9 with a 2-BD $(m,\{4,5,6\})$ to obtain a $(2^m,2)$ EKGDD. For $m\geqslant 4$, apply Proposition 4.7 with a $(2^m,2)$ EKGDD and a TD$(4,4)$ to obtain an $(8^m,2)$ EKGDD. □

Equation (6.3) follows from Theorem 1.2, Propositions 2.2 and 5.1(iv), and Lemma 6.3.

Proposition 6.4. *Let $n \equiv 3, 5, 7 \bmod 8$, X be an n-set and $\mathcal{C} \subseteq \mathcal{E}(X, 2)$. Then $\mathcal{C}$ is an optimal $(n, 2, 3)$ EK code if*

(i) $n \equiv 3 \bmod 8$ *and $(X, \mathcal{G}, \mathcal{C})$ is a $(1^{n-3}3, 2)$ EKGDD, where $\mathcal{G}$ is a partition of X into $n-3$ groups of single points and one group of three points, or*

(ii) $n \equiv 5, 7 \bmod 8$ *and $(X, \mathcal{C})$ is a 2-$(n, 2)$ EK packing of size exactly $\lfloor n(n-1)/8 \rfloor - 1$.*

Proof. If $n \equiv 3 \bmod 8$ and $(X, \mathcal{G}, \mathcal{C})$ is a $(1^{n-3}3, 2)$ EKGDD, then $(X, \mathcal{C})$ is a 2-$(n, 2)$ EK packing of size $\lfloor n(n-1)/8 \rfloor$. Optimality of $\mathcal{C}$ follows from Theorem 1.2.

If $n \equiv 5, 7 \bmod 8$ and $(X, \mathcal{C})$ is a 2-$(n, 2)$ EK packing of size $\lfloor n(n-1)/8 \rfloor - 1$, then $\mathcal{C}$ is an $(n, 2, 3)$ EK code of size $\lfloor n(n-1)/8 \rfloor - 1$ by Proposition 2.1. From Theorem 1.2, it suffices to show that there exists no $(n, 2, 3)$ EK code of size $\lfloor n(n-1)/8 \rfloor$. Suppose otherwise. Then equivalently, there exists a 2-$(n, 2)$ EK packing $(X, \mathcal{D})$ of size $\lfloor n(n-1)/8 \rfloor$. Consider the following subset $\mathcal{P} \subseteq \binom{X}{2}$:

$$\mathcal{P} := \{\{x, y\} : x \in A, y \in B \quad \text{for some } \{A, B\} \in \mathcal{D}\}.$$

Since $(X, \mathcal{D})$ is an EK-packing, we obtain

$$\left| \binom{X}{2} \setminus \mathcal{P} \right| = \binom{n}{2} - 4 \left\lfloor \frac{n(n-1)}{8} \right\rfloor = \begin{cases} 2 & \text{if } n \equiv 5 \bmod 8, \\ 1 & \text{if } n \equiv 7 \bmod 8. \end{cases}$$

Then there exists $x \in X$ that occurs exactly once in the pairs of $\binom{X}{2} \setminus \mathcal{P}$. Consider all pairs in $\mathcal{P}$ that contain x. On one hand, since there is exactly one pair in $\binom{X}{2} \setminus \mathcal{P}$ that contains x, the number $|\{\{x, y\} \in \mathcal{P} : y \in X\}| = (n-1) - 1 = n - 2$ is odd. On the other hand, if we consider all set-pairs in $\mathcal{D}$ containing x, then

$$|\{\{x, y\} \in \mathcal{P} : y \in X\}| = 2|\{x \in A \cup B : \{A, B\} \in \mathcal{D}\}|$$

is even, a contradiction. Therefore, (ii) follows. □

Lemma 6.5. *There exists a 2-$(n, 2)$ EK design for $n \geqslant 9$ and $n \equiv 1 \bmod 8$.*

Proof. A 2-$(n, 2)$ EK design exists for $n \in \{9, 17, 25\}$ by Proposition 5.1(ii). For $n \geqslant 33$, write $n = 8m + 1$ for $m \geqslant 4$. Then an $(8^m, 2)$ EKGDD exists by Lemma 6.3. Apply Corollary 4.2 to obtain the desired design. □

Lemma 6.6. *There exists a $(1^{n-3}3, 2)$ EKGDD for $n \geqslant 11$ and $n \equiv 3 \bmod 8$.*

Proof. A $(1^{n-3}3, 2)$ EKGDD exists for $n \in \{11, 19, 27\}$ by Proposition 5.1(iii). For $n \geqslant 35$, write $n = 8m + 3$ for $m \geqslant 4$. Then an $(8^m, 2)$ EKGDD exists by Lemma 6.3. Apply Proposition 4.1 to obtain the desired EKGDD. □

Equation (6.4) follows from Propositions 2.1 and 6.4, and Lemmas 6.5 and 6.6.

Lemma 6.7. *Let $r \in \{0, 1\}$ and $r + s > 0$ and suppose there exists a* TD$(4 + r + s, 4m + r)$. *Let $0 \leqslant g_i \leqslant 4m$, $i \in [s]$ and suppose there exists a 2-$(n, 2)$ EK design for all*

Table 3. Existence of 2-$(n,2)$ EK packings of size $\lfloor n(n-1)/8 \rfloor - 1$ for $n \equiv 5,7 \bmod 8$.

Authority	n
Proposition 5.1(iv)	5, 7, 13, 15, 21
Lemma 6.8	23, 29, 31, 53, 55, 61, 63
Lemma 6.7 with $m=1$, $r=0$, $g_1 \in \{2,3\}$	37, 39
Lemma 6.7 with $m=1$, $r=1$, $g_1 \in \{2,3\}$	45, 47
Lemma 6.7 with $m=2$, $r=0$, $g_1,g_2,g_3 \in \{0,4,8\}$, $g_4 \in \{2,3,6,7\}$	69–127

$n \in \{8m+1\} \cup \{2g_i+1 : i \in [s-1]\}$ *and a* 2-$(2g_s+1,2)$ *EK packing of size* $\lfloor g_s(2g_s+1)/4 \rfloor - 1$. *Then there exists a* 2-$(N,2)$ *EK packing of size* $\lfloor N(N-1)/8 \rfloor - 1$, *where*

$$N = 1 + 8m(4+r) + 2\sum_{i=1}^{s} g_i.$$

Proof. By Lemma 6.3, a $(2^u,2)$ EKGDD exists for $u \geqslant 4$. Apply Proposition 4.10 to obtain an $((8m)^{4+r}(2g_1)\cdots(2g_s),2)$ EKGDD. Now apply Proposition 4.3 to obtain a 2-$(N,2)$ EK packing of size $\lfloor N(N-1)/8 \rfloor - 1$. □

Lemma 6.8. *A* 2-$(n,2)$ *EK packing of size* $\lfloor n(n-1)/8 \rfloor - 1$ *exists for* $n \in \{23,29,31,53,55,61,63\}$.

Proof. For $(u,t) \in \{(16,7),(24,5),(24,7)\}$, apply Proposition 4.3 to $(1^u t,2)$ EKGDDs constructed in Proposition 5.1(iii) to obtain 2-$(n,2)$ EK packings for $n \in \{23,29,31\}$. For $u \in \{6,7\}$ and $t \in \{4,6\}$, apply Proposition 4.3 to the $(8^u t,2)$ EKGDDs constructed in Proposition 5.1(iii) to obtain 2-$(n,2)$ EK packings for $n \in \{53,55,61,63\}$. □

Lemma 6.9. *There exists a* 2-$(n,2)$ *EK packing of size* $\lfloor n(n-1)/8 \rfloor - 1$ *for* $n \geqslant 5$ *and* $n \equiv 5,7 \bmod 8$.

Proof. A 2-$(n,2)$ EK packing of size $\lfloor n(n-1)/8 \rfloor - 1$ can be constructed for $n \equiv 5,7 \bmod 8$, $5 \leqslant n \leqslant 127$. Details are provided in Table 3.

For $n \geqslant 133$, write $n = 32m + 2t + 1$, such that $m \geqslant 4$ and $t \in \{2,3,6,7,10,11,14,15\}$. Then a TD$(5,4m)$ exists by Theorem 4.5 and a 2-$(8m+1,2)$ EK design exists by Lemma 6.5. Apply Lemma 6.7 with $r=0$ and $g_1 = t$ to obtain the result. □

Equation (6.5) follows from Proposition 6.4 and Lemma 6.9.

6.3. Determining $C(n,3,5)$

Lemma 6.10. *There exists a* $(3^m,3)$ *EKGDD for* $m \geqslant 343$.

Proof. By Proposition 5.1(iii), there exists a $(3^m,3)$ EKGDD for $m \in \{7,8,9\}$. For $m \geqslant 343$, there exists a 2-BD$(m,\{7,8,9\})$ by Theorem 4.4. Apply Proposition 4.9 to obtain a $(3^m,3)$ EKGDD. □

Table 4. Existence of $(n, 3)$ EK designs

Authority	n
Proposition 5.1(ii)	19–55, 73–91, 109–127, 145, 163, 181, 199
Lemma 6.12	154, 172
Lemma 6.11 with $m = 3$, $r = 0$, $g_1, g_2, g_3 \in \{0, 2, 3\}$	190, 208–271
Lemma 6.11 with $m = 3$, $r = 2$, $g_1, g_2 \in \{2, 3\}$	280–298
Lemma 6.11 with $m \in \{4, 5, 6, 8, 10, 12, 14, 16\}$, $r = 1$, $0 \leqslant g_1, g_2, g_3 \leqslant m$, $g_1, g_2, g_3 \notin \{1, 7, 11, 15\}$	307–1513

Hence, equation (6.6) follows from Proposition 2.2 and Lemma 6.10.

Lemma 6.11. *Let* $r \in \{0, 1, 2\}$, $r + s \in [4]$ *and suppose there exists a* $\mathrm{TD}(7 + r + s, 3m + r)$ *with* r *disjoint blocks. If* $0 \leqslant g_i \leqslant m$, $i \in [s]$ *and there exists a* 2-$(n, 3)$ *EK design for all*

$$n \in \{9m + 1\} \cup \{9g_i + 1 : i \in [s]\},$$

then there exists a 2-$(1 + 9m(7 + r) + 9\sum_{i=1}^{s} g_i, 3)$ *EK design.*

Proof. By Proposition 5.1(iii), a $(3^u, 3)$ EKGDD exists for $u \in \{7, 8, 9, 10, 11\}$. By Proposition 4.10, there exists a $((9m)^{7+r}(9g_1) \cdots (9g_s), 3)$ EKGDD. Now apply Corollary 4.2 to obtain the desired design. □

Lemma 6.12. *A* 2-$(n, 3)$ *EK design exists for* $n \in \{154, 172\}$.

Proof. Apply Corollary 4.2 to 2-(n, k) EK designs for $n \in \{19, 28\}$ and $(18^u 27, 3)$ EKGDD for $u \in \{7, 8\}$ from Propositions 5.1(ii) and 5.1(iii), respectively, to obtain the required EK designs. □

Lemma 6.13. *A* 2-$(n, 3)$ *EK design exists for all* $n \geqslant 19$ *and* $n \equiv 1 \bmod 9$, *except possibly for* $n \in \{64, 100, 136\}$.

Proof. A 2-$(n, 3)$ EK design can be constructed for $n \equiv 1 \bmod 9$, $19 \leqslant n \leqslant 1513$, and $n \notin \{64, 100, 136\}$. Details are provided in Table 4.

For $n \geqslant 1522$, we prove by induction. Write $n = 63m + 9t + 1$, where $t \in \{0, 2, 3, 4, 5, 6, 8\}$ and $23 \leqslant m < n$. Then a $\mathrm{TD}(8, 3m)$ exists by Theorem 4.5 and a 2-$(9m + 1, 3)$ EK design exists by induction hypothesis. Apply Lemma 6.11 with $r = 0$ and $g_1 = t$ to obtain the result. □

Therefore, equation (6.7) follows from Proposition 2.1 and Lemma 6.13.

7. Conclusion

New infinite families of optimal codes in the Enomoto–Katona space are obtained in this paper. In particular, we show that $C(n, k, 2k - 1)$ attains the Quistorff bound for infinitely many n. In addition, the value of $C(n,2,d)$ is also completely determined, while the value of $C(n,3,5)$ is determined for certain congruence classes with finite exceptions.

Acknowledgement

The authors are grateful to the anonymous referee and managing editor Professor Paul Balister for their helpful comments.

References

[1] Abel, R. J. R., Bennett, F. B. and Greig, M. (2007) PBD-closure. In *The CRC Handbook of Combinatorial Designs* (C. J. Colbourn and J. H. Dinitz, eds), second edition, CRC Press, pp. 247–255.

[2] Abel, R. J. R., Colbourn, C. J. and Dinitz, J. H. (2007) Mutually orthogonal Latin squares (MOLS). In *The CRC Handbook of Combinatorial Designs* (C. J. Colbourn and J. H. Dinitz, eds), second edition, CRC Press, pp. 160–193.

[3] Armstrong, W. W. (1974) Dependency structures of data base relationships. In *Proc. IFIP Congress 1974*, pp. 580–583.

[4] Beeri, C., Fagin, R. and Howard, J. H. (1977) A complete axiomatization for fuzzy functional and multivalued dependencies in fuzzy database relations. In *Proc. ACM SIGMOD International Conference on Management of Data*, pp. 47–61.

[5] Bernstein, P. A. (1976) Synthesizing third normal form relations from functional dependencies. *ACM Trans. Database Syst.* **1** 277–298.

[6] Bollobás, B., Füredi, Z., Kantor, I., Katona, G. O. H. and Leader, I. A coding problem for pairs of subsets. Preprint.

[7] Brightwell, G. and Katona, G. (2001) A new type of coding problem. *Studia Sci. Math. Hungar.* **38** 139–147.

[8] Chee, Y. M., Kiah, H. M., Zhang, H. and Zhang, X. Addendum to optimal codes in the Enomoto–Katona space. `http://sites.google.com/kiahhanmao/optimalek`

[9] Chee, Y. M., Kiah, H. M., Zhang, H. and Zhang, X. (2013) Optimal codes in the Enomoto–Katona space. In *Proc. IEEE International Symposium on Information Theory*, pp. 321–325.

[10] Codd, E. F. (1970) A relational model of data for large shared data banks. *Commun. Assoc. Comput. Mach.* **13** 377–387.

[11] Demetrovics, J., Katona, G. O. and Sali, A. (1995) Minimal representations of branching dependencies. *Acta Sci. Math.* (Szeged) **60** 213–224.

[12] Demetrovics, J., Katona, G. O. H. and Sali, A. (1992) The characterization of branching dependencies. *Discrete Appl. Math.* **40** 139–153.

[13] Demetrovics, J., Katona, G. O. H. and Sali, A. (1998) Design type problems motivated by database theory. *J. Statist. Plann. Inference* **72** 149–164.

[14] Enomoto, H. and Katona, G. O. H. (2001) Pairs of disjoint q-element subsets far from each other. *Electron. J. Combin.* **8** R7.

[15] Hanani, H. (1971) Truncated finite planes. In *Proc. Symposium Pure Mathematics* (AMS), Vol. 19, pp. 115–120.

[16] Katona, G. O. H. and Sali, A. (2004) New type of coding problem motivated by database theory. *Discrete Appl. Math.* **144** 140–148.

[17] Lamken, E. R. and Wilson, R. M. (2000) Decompositions of edge-colored complete graphs. *J. Combin. Theory Ser. A* **89** 149–200.

[18] Quistorff, J. (2009) Combinatorial problems in the Enomoto–Katona space. *Studia Sci. Math. Hungar.* **46** 121–139.

[19] Rissanen, J. (1977) Independent components of relations. *ACM Trans. Database Syst.* **2** 317–325.

[20] Sali, A. (2011) Coding theory motivated by relational databases. In *Proc. International Conference on Semantics in Data and Knowledge Bases: SDKB'10*, Springer, pp. 96–113.

[21] Wilson, R. M. (1972) An existence theory for pairwise balanced designs I: Composition theorems and morphisms. *J. Combin. Theory Ser. A* **13** 220–245.