A sequence approach to linear perfect hash families*

S. G. Barwick and Wen-Ai Jackson
School of Pure Mathematics, University of Adelaide,
Adelaide 5005, Australia

May 12, 2006

Abstract

A linear (q^d, q, t)-perfect hash family of size s in a vector space V of order q^d over a field F of order q consists of a set $S = \{\phi_1, \ldots, \phi_s\}$ of linear functionals from V to F with the following property: for all t subsets $X \subseteq V$ there exists $\phi_i \in S$ such that ϕ_i is injective when restricted to F. A linear (q^d, q, t)-perfect hash family of minimal size $d(t-1)$ is said to be *optimal*. In this paper we extend the theory for linear perfect hash families based on sequences developed by Blackburn and Wild. We develop techniques which we use to construct new optimal linear $(q^2, q, 5)$-perfect hash families and $(q^4, q, 3)$-perfect hash families. The sequence approach also explains a relationship between linear $(q^3, q, 3)$-perfect hash families and linear $(q^2, q, 4)$-perfect hash families.

1 Introduction to Perfect Hash Families

Perfect hash families were introduced by Mehlhorn [14] in 1984 as part of compiler design. Perfect hash families have also proved useful in a large variety of applications, in particular, there have been a number of recent applications to cryptography. For example, to threshold cryptography (see Blackburn, Burmester, Desmedt and Wild [8] and Blackburn [6]), to broadcast encryption (see Fiat and Naor [11]) and to improve explicit constructions of secure frameproof codes, key distribution patterns, group testing algorithms, cover free families and separating systems (see Stinson, van Trung and Wei [15]).

Let s, t, n, q be positive integers and let V be a set of size n and let F be a set of size q. A function $\phi: V \to F$ *separates* a subset X of V if ϕ is an injection when restricted to X. An (n, q, t)-*perfect hash family* of *size* s is a set $S = \{\phi_1, \ldots, \phi_s\}$ of s functionals from V to F with the property that for all t-subsets $X \subseteq V$, at least one of $\phi_1, \ldots, \phi_s$ separates X.

*This work was supported by the Australian Research Council

We say that S is a *linear* perfect hash family if F can be identified with a finite field $\mathrm{GF}(q)$ and V can be identified with a vector space over $\mathrm{GF}(q)$ in such a way that S is a set of linear functionals under this identification. Thus in the linear case, q is a prime power and $n = q^d$ for some non-negative integer d. This paper deals solely with linear perfect hash families and throughout we use q to denote a prime power. Linear perfect hash families also have a geometric interpretation which was used in [3, 4] to construct linear perfect hash families.

The following result from Blackburn and Wild [9] gives a bound on the size of a linear perfect hash family.

Theorem 1.1 ([9]) *Let d and t be integers such that $d \geq 2$ and $t \geq 2$ and let q be a prime power. If S is a linear (q^d, q, t)-perfect hash family, then $|S| \geq d(t-1)$.*

If $|S| = d(t-1)$ then S is called *optimal*. Blackburn and Wild give conditions for the existence of optimal linear perfect hash families.

Theorem 1.2 ([9]) *An optimal linear (q^d, q, t)-perfect hash family S exists if $q \geq \binom{t}{2}^{d(t-1)}$.*

Perfect hash families are an interesting combinatorial structure with practical applications. They are hard to construct, and there are limited known constructions of linear perfect hash families. Blackburn and Wild [9] give a general construction of an optimal linear perfect hash family which works for q much larger than the bound of Theorem 1.2 and of a certain form, namely $q = q_0^{\alpha_1\alpha_2\cdots\alpha_{d(t-2)}}$ where q_0 is any prime power and each $\alpha_i \geq d$. Blackburn [7] gives a construction of optimal linear $(p^2, p, 4)$-perfect hash families where p is a prime, $p = 11$ or $p \geq 17$. Wang and Xing [16] construct linear perfect hash families but their constructions are not optimal. In [3] the authors use geometric techniques to show that optimal linear $(q^2, q, 4)$-perfect hash families exist if and only if $q = 11$ or $q \geq 17$ (q a prime power) and constructions are given for each such q. In [4] geometric techniques are used to show that optimal linear $(q^3, q, 3)$-perfect hash families exist if and only if $q \geq 11$ (q a prime power) and constructions are given for each such q. The authors also give constructions of optimal linear $(q^2, q, 5)$-perfect hash families for restricted values of q. In [1, 2], recursive algorithms for constructing perfect hash families are given. These algorithms need as input perfect hash families with small parameters. This gives motivation for constructing small perfect hash families.

In this paper we investigate the sequence representation of linear perfect hash families developed by Blackburn and Wild [9] in more detail. We generalise their results with the aim of using a sequence approach to construct optimal linear perfect hash families. In Theorem 2.14, we give a set of conditions that can be used to check whether a given set of functionals forms a

perfect hash family. We then use the theory developed to give a general construction technique in Section 2.2. Known constructions methods of perfect hash families are ad hoc and this sequence approach gives a general method which can be used to recursively construct optimal linear perfect hash families for any q, d, t for which they exist. We illustrate how this approach can be used to construct the optimal linear $(q^2, q, 4)$- and $(q^3, q, 3)$-perfect hash families found in [3, 4]. Further, the sequence approach explains a correspondence between these two cases. We also use these techniques to construct new optimal linear $(q^2, q, 5)$- and $(q^4, q, 3)$-perfect hash families.

2 The sequence approach to linear perfect hash families

In this section we introduce the sequence approach used by Blackburn and Wild in [9]. In their paper they used sequences to prove Theorem 1.2, that is, to show that optimal linear perfect hash families exist if q is large enough. The proof of this result is a probabilistic argument and does not use the sequences to produce constructions of perfect hash families. Our approach here is to study the sequences with the aim of characterizing when they are perfect hash families in order to develop techniques to construct perfect hash families. Consequently we need to study these sequences in more detail than previously done, in particular we need to generalise the results of [9] to results that hold for smaller q; and we need to obtain some more specialised results on these sequences. Some of the preliminary results here are from [9], but we have included some short proofs to assist in understanding the theory. We use the notation of [9]. As a lot of notation and definitions are involved, we include a notation and definition index in the appendix for easy reference by the reader.

2.1 The sequence approach to the characterization of optimal linear perfect hash families

Let V be a vector space of dimension $d > 0$ over $\mathrm{GF}(q)$, where $q \geq 2$. Let V^* be the dual space of V, consisting of the set of linear functionals $\phi\colon V \to \mathrm{GF}(q)$. If $p \in V$ and $\phi \in V^*$, we will use the notation p^ϕ to represent ϕ acting on p. Given a set $S = \{\phi_1, \ldots, \phi_k\}$ of k linear functionals, we may order them in some arbitrary way to produce an *associated functional sequence* $\Phi = (\phi_1, \ldots, \phi_k) \in (V^*)^k$. In [9], conditions were found on Φ so that S (or equivalently Φ) is a perfect hash family. To explain these conditions we need to introduce the following notation and concepts from [9].

Note that the $t = 1$ case is trivial and if $q < t$ then no functional can separate t points.

Hence let $2 \leq t \leq q$ and consider $(V^*)^t$, a dt-dimensional vector space over GF(q). Let $\Psi = (\psi_1, \ldots, \psi_t) \in (V^*)^t$. Let $P = (p_1, \ldots, p_t)$ be a sequence of t elements of V. We can regard Ψ as a linear map $\Psi\colon V^t \to \mathrm{GF}(q)$ over GF(q) where $P^\Psi = p_1^{\psi_1} + \cdots + p_t^{\psi_t}$. We define a subspace U_P of $(V^*)^t$ given by

$$U_P = \left\{ \Psi = (\psi_1, \ldots, \psi_t) \in (V^*)^t : p_1^{\psi_1} + \cdots + p_t^{\psi_t} = 0 \right\}.$$

Let $T = \{1, 2, \ldots, t\}$ and define T^2 to be $\{(a, b) \in T \times T : 1 \leq a < b \leq t\}$. Note that in our notation here, T^2 is a *proper* subset of $T \times T$. An element (a, b) of T^2 is called a *pair on* T, and a sequence c of k pairs (possibly repeated) from T^2 is called a *k-pair sequence on* T. For a pair (a, b) on T and all $\phi \in V^*$, define the vector $w_{(a,b),\phi} \in (V^*)^t$ by

$$w_{(a,b),\phi} = (\psi_1, \ldots, \psi_t) \quad \text{where} \quad \psi_i = \begin{cases} \phi & \text{if } i = a, \\ -\phi & \text{if } i = b, \\ 0 & \text{if } i \neq a, b. \end{cases}$$

So for example, $w_{(1,2),\phi} = (\phi, -\phi, 0, \ldots, 0)$. We use $w_{(a,b),\phi}$ to determine whether ϕ separates a t-subset $\{p_1, \ldots, p_t\}$ of V. We let $P = (p_1, \ldots, p_t)$ be an arbitrary ordering of the t-subset and say that ϕ *separates* the sequence P if ϕ separates the set $\{p_1, \ldots, p_t\}$.

Result 2.1 ([9]) *Let $P = (p_1, \ldots, p_t)$ be a sequence of t distinct elements of V. Let $\phi \in V^*$. Then ϕ fails to separate P if and only if $w_{(a,b),\phi} \in U_P$ for some $(a, b) \in T^2$.*

Proof The linear functional ϕ fails to separate P if and only if ϕ maps two elements of P to the same element of GF(q). That is, if and only if there exists $(a, b) \in T^2$ such that $p_a^\phi = p_b^\phi$, that is, $w_{(a,b),\phi} \in U_P$. □

So, for example, if $w_{(1,2),\phi} \in U_P$, then $p_1^\phi + p_2^{-\phi} + p_3^0 + \ldots + p_t^0 = 0$, that is $p_1^\phi = p_2^\phi$, so ϕ maps p_1 and p_2 to the same element of GF(q). Thus ϕ fails to separate P.

For $(a, b) \in T^2$, we define the subspace $V_{(a,b)}$ of $(V^*)^t$ by

$$V_{(a,b)} = \{w_{(a,b),\phi} : \phi \in V^*\}.$$

Note that $V_{(a,b)}$ is a vector space of dimension d over GF(q). We are interested in t-subsets $\{p_1, \ldots, p_t\}$ of V which necessarily have distinct elements. However, when we consider an arbitrary t-sequence $P = (p_1, \ldots, p_t)$ from V^t, P may have repeated elements. The subspace $V_{(a,b)}$ gives us a way to determine whether P has repeated elements.

Result 2.2 ([9]) *Let $P = (p_1, \ldots, p_t)$ be a sequence of t elements of V. Then P has a repeated element $p_a = p_b$ (for some a, b, $1 \leq a < b \leq t$) if and only if $V_{(a,b)} \subseteq U_P$.*

Proof We have $p_a = p_b$ if and only if $p_a^\phi = p_b^\phi$ for all $\phi \in V^*$, if and only if $w_{(a,b),\phi} \in U_P$ for all $\phi \in V^*$, if and only if $V_{(a,b)} \subseteq U_P$. □

Let $c = ((a_1, b_1), \ldots, (a_k, b_k))$ be a k-pair sequence on T and let $\Phi = (\phi_1, \ldots, \phi_k) \in (V^*)^k$ be a k-functional sequence. Define the subspace $W_{c,\Phi}$ of $(V^*)^t$ to be

$$W_{c,\Phi} = \left\langle w_{(a_i,b_i),\phi_i} : 1 \leq i \leq k \right\rangle.$$

So, for example, let $k = 2$, $c = ((1,2),(1,3))$ and $\Phi = (\phi_1, \phi_2)$, then $W_{c,\Phi} = \langle (\phi_1, -\phi_1, 0, \ldots, 0), (\phi_2, 0, -\phi_2, 0, \ldots, 0) \rangle$.

It will often to be convenient to write the subspace $W_{c,\Phi}$ in matrix form. We let $[W_{c,\Phi}]$ be the matrix with k rows $w_{(a_i,b_i),\phi_i}$, $i = 1, \ldots, k$. So in the above example, we have

$$[W_{c,\Phi}] = \begin{bmatrix} \phi_1 & -\phi_1 & 0 & 0 & \ldots & 0 \\ \phi_2 & 0 & -\phi_2 & 0 & \ldots & 0 \end{bmatrix}.$$

This matrix has t columns over V^*; and dt columns over GF(q). The row space of the matrix $[W_{c,\Phi}]$ is equal to the subspace $W_{c,\Phi}$, and $\dim W_{c,\Phi} = \text{rank } [W_{c,\Phi}]$. In what follows, we will drop the brackets and use $W_{c,\Phi}$ to refer to both the subspace and to the matrix, and we will use rank $W_{c,\Phi}$ to mean the rank of the matrix.

The next result involving $W_{c,\Phi}$ shows when the set $\{\phi_1, \ldots, \phi_k\}$ of functionals from V^* separates the t-subset $\{p_1, \ldots, p_t\}$ of V.

Result 2.3 ([9]) *Let $P = (p_1, \ldots, p_t)$ be a sequence of t distinct elements of V and let $S = \{\phi_1, \ldots, \phi_k\}$ be a set of k linear functionals with associated k-functional sequence Φ. Then S fails to separate P if and only if there exists a k-pair sequence c on T such that $W_{c,\Phi} \subseteq U_P$.*

Proof The set S fails to separate P if and only if every functional $\phi_i \in \Phi$ maps two distinct elements of P to the same element of GF(q). That is, if and only if for each $\phi_i \in \Phi$ there exists $(a_i, b_i) \in T$ such that $p_{a_i}^{\phi_i} = p_{b_i}^{\phi_i}$, that is, $w_{(a_i,b_i),\phi_i} \in U_P$. This occurs if and only if there exists a k-pair sequence c on T with $W_{c,\Phi} \subseteq U_P$. □

Note that if a set $\{\phi_1, \ldots, \phi_k\}$ fails to separate a sequence $(p_1, \ldots, p_t)$ of distinct elements, then it fails to separate any supersequence $(p_1, \ldots, p_t, p_{t+1})$.

In the next Lemma we will use the subspace Y of $(V^*)^t$ defined as:

$$Y = \left\{ (\psi_1, \ldots, \psi_t) \in (V^*)^t : \psi_1 + \cdots + \psi_t = 0 \right\}.$$

Note that Y has dimension $d(t-1)$ over GF(q) (since $\psi_1, \ldots, \psi_{t-1}$ can be chosen arbitrarily, and then $\psi_t = -(\psi_1 + \ldots + \psi_{t-1})$).

Lemma 2.4 *Suppose S is a set of $d(t-1)$ linear functionals with associated functional sequence Φ. Let c be a $d(t-1)$-pair sequence on T. Then $W_{c,\Phi} \subseteq Y$, and so* rank $W_{c,\Phi} \leq d(t-1)$. *In particular, if* rank $W_{c,\Phi} = d(t-1)$, *then $W_{c,\Phi} = Y$ and so $V_{(a,b)} \subseteq W_{c,\Phi}$ for all pairs (a,b) on T.*

Proof As $w_{(a,b),\phi} \in Y$ for all pairs (a,b) on T and $\phi \in V^*$, it follows that $W_{c,\Phi} \subseteq Y$. Hence rank $W_{c,\Phi} \leq \dim Y$, that is rank $W_{c,\Phi} \leq d(t-1)$ (this rank bound also follows directly as c, Φ have length $d(t-1)$, so $W_{c,\Phi}$ is the span of $d(t-1)$ vectors). The remaining statement follows as $V_{(a,b)} \subseteq Y$ by definition. □

The following result was stated in [9] without proof (the proof is very similar to the proof of [9, Theorem 4]).

Result 2.5 ([9]) *Suppose S is a linear (q^d, q, t)-perfect hash family of optimal size $d(t-1)$. Then every d elements of S are linearly independent over GF(q).*

We now describe the geometrical interpretation of linear perfect hash families as introduced in [9]. This allows us to give an easy proof for the next two results. For an introduction to finite projective geometry see [10]. Let π_∞ be the hyperplane at infinity in the projective space PG(d,q), so π_∞ is a subspace of projective dimension $d-1$. We identify the elements of V with the affine points of PG(d,q); that is, the points of AG$(d,q) =$ PG$(d,q)\backslash\pi_\infty$. For any linear functional $\phi \in V^*$ and $\gamma \in$ GF(q), the points $v \in V$ with $v^\phi = \gamma$ form a hyperplane of AG(d,q) and so ϕ corresponds to a parallel class of hyperplanes. These parallel hyperplanes all meet π_∞ in a hyperplane $[\phi]$ of π_∞ (so $[\phi]$ is a subspace of projective dimension $d-2$). So there is a 1-1 correspondence between parallel classes of hyperplanes in PG(d,q) and the hyperplanes of π_∞, that is, there is a 1-1 correspondence between linear functionals ϕ and hyperplanes $[\phi]$ of π_∞. Two affine points $p_1, p_2 \in$ PG$(d,q)\backslash\pi_\infty$ are separated by ϕ if and only if they belong to different hyperplanes of the parallel class determined by ϕ. That is, p_1, p_2 are separated by ϕ if and only if the line p_1p_2 meets π_∞ in a point not in $[\phi]$.

We note that perfect hash family results have a geometric interpretation in PG(d,q). For example, Result 2.5 states that every d linear functionals of a perfect hash family $S = \{\phi_1, \ldots, \phi_{d(t-1)}\}$ are independent. Hence the hyperplanes $[\phi_1], \ldots, [\phi_{d(t-1)}]$ corresponding to the linear functionals form a dual arc in π_∞. We use this geometric interpretation to classify optimal linear perfect hash families with $t = 2$.

Lemma 2.6 *The set $S = \{\phi_1, \ldots, \phi_d\}$ is an optimal linear $(q^d, q, 2)$-perfect hash family if and only if $\phi_1, \ldots, \phi_d$ are linearly independent.*

Proof The functionals $\phi_1, \ldots, \phi_d$ correspond to hyperplanes $[\phi_1], \ldots, [\phi_d]$ in π_∞. Consider two distinct affine points $p_1, p_2 \in \mathrm{PG}(d,q)\backslash\pi_\infty$. Then $p_1p_2 \cap \pi_\infty$ is a point r. If $r \in [\phi_i]$, then ϕ_i does not separate p_1, p_2. Hence we need the d hyperplanes $[\phi_1], \ldots, [\phi_d]$ of $\pi_\infty \cong \mathrm{PG}(d-1,q)$ to have no common point. This will happen if and only if $[\phi_1], \ldots, [\phi_d]$ are linearly independent. That is, $\phi_1, \ldots, \phi_d$ separate every 2-set $\{p_1, p_2\}$ of V if and only if $\phi_1, \ldots, \phi_d$ are linearly independent. □

Theorem 2.7 *Suppose $2 \leq t \leq q$. If Φ is a linear (q^d, q, t)-perfect hash family, then for each t' $(2 \leq t' < t)$, any $d(t'-1)$ subset Φ' of Φ is a (q^d, q, t')-perfect hash family.*

Proof If $t = 2$ there is nothing to prove. So suppose firstly that $t' = t - 1 \geq 2$. Let Φ' be a subset of Φ of size $d(t'-1) = d(t-2)$. Suppose that Φ' is not a $(q^d, q, t-1)$-perfect hash family. So there exists a set $\{p_1, \ldots, p_{t-1}\}$ of $t-1$ distinct elements of V that are not separated by Φ'. In the geometrical representation, these $t-1$ elements of V correspond to $t-1$ distinct points $p_1, \ldots, p_{t-1}$ of $\mathrm{PG}(d,q)\backslash\pi_\infty$. Consider the d linear functionals $\phi_1, \ldots, \phi_d$ in $\Phi\backslash\Phi'$. They are linearly independent by Result 2.5, so they correspond to linearly independent hyperplanes $[\phi_1], \ldots, [\phi_d]$ of π_∞. Hence the intersection $[\phi_1] \cap [\phi_2] \cap \cdots \cap [\phi_{d-1}]$ is a point $r \in \pi_\infty$ which is not contained in the final hyperplane $[\phi_d]$.

Now the line p_1r intersects $\langle p_2, [\phi_d]\rangle$ in a point p. Suppose firstly that $p \notin \{p_1, \ldots, p_{t-1}\}$ (in particular, $p \neq p_1, p_2$). We have $p_2p \cap \pi_\infty \in [\phi_d]$, hence ϕ_d does not separate $\{p_2, p\}$. Further, $p_1p \cap \pi_\infty = r \in [\phi_1] \cap [\phi_2] \cap \cdots \cap [\phi_{d-1}]$, so $\phi_1, \ldots, \phi_{d-1}$ do not separate $\{p_1, p\}$. As no functional in Φ' separates the set $\{p_1, \ldots, p_{t-1}\}$ we have shown that no functional in Φ separates the set $\{p_1, \ldots, p_{t-1}, p\}$ of t distinct elements. Thus Φ is not a (q^d, q, t)-perfect hash family, a contradiction.

If $p = p_1$ then let p_t be any point of $p_1r\backslash\{p_1, \ldots, p_{t-1}, r\}$, this is possible as $q \geq t$ implies every line contains $q + 1 \geq t + 1$ points. Then no functional in $\{\phi_1, \ldots, \phi_{d-1}\}$ separates $\{p_1, p_t\}$; and ϕ_d does not separate $\{p_1, p_2\}$. As no functional in Φ' separates the set $\{p_1, \ldots, p_{t-1}\}$, we have shown that no functional in Φ separates the set $\{p_1, \ldots, p_{t-1}, p_t\}$ of t distinct elements. Thus Φ is not a (q^d, q, t)-perfect hash family, a contradiction. Similarly, if $p = p_2$ then $\phi_1, \ldots, \phi_{d-1}$ do not separate $\{p_1, p_2\}$. Let $p_t \in \langle p_2, [\phi_d]\rangle\backslash\{p_1, \ldots, p_{t-1}, \pi_\infty\}$ (this is possible as $q \geq t$), then ϕ_d does not separate $\{p_2, p_t\}$. Hence Φ does not separate the distinct elements $\{p_1, \ldots, p_t\}$ and so Φ is not a perfect hash family, a contradiction. Finally, if $p \in \{p_3, \ldots, p_{t-1}\}$, then Φ does not separate $\{p_1, \ldots, p_{t-1}\}$ and so does not separate any t-set containing them, a contradiction.

Hence Φ' is a $(q^d, q, t-1)$-perfect hash family. The remaining values of t' follow by repeating the argument. □

In the following lemmas we relate pair sequence properties with the rank properties of $W_{c,\Phi}$.

Recall that if Φ is a perfect hash family, then Φ is a $d(t-1)$-functional sequence and so $W_{c,\Phi}$ is the span of $d(t-1)$ vectors, hence rank $W_{c,\Phi} \leq d(t-1)$. The next result shows that $d(t-1)$ is the maximal rank of $W_{c,\Phi}$ for any k-functional sequence Φ and k-pair sequence c. We will show later that Φ is a perfect hash family if and only if the rank of $W_{c,\Phi}$ is this maximal value $d(t-1)$ for all relevant k-pair sequences c. We will use the subspace

$$R = \{(p, \ldots, p) : p \in V\} \subseteq V^t.$$

Note that R is a vector space of dimension 1 over V, and consequently of dimension d over GF(q).

Lemma 2.8 *Let Φ be a k-functional sequence and let c be any k-pair sequence on T. Then we have* rank $W_{c,\Phi} \leq d(t-1)$. *Further,* rank $W_{c,\Phi} < d(t-1)$ *if and only if there exists $P \in V^t \setminus R$ with $W_{c,\Phi} \subseteq U_P$.*

Proof Let $\Phi = (\phi_1, \ldots, \phi_k)$, $c = (c_1, \ldots, c_k)$ and consider $W_{c,\Phi}$ in matrix form. It satisfies the dimension theorem over GF(q), that is,

$$\dim \ker W_{c,\Phi} + \text{rank } W_{c,\Phi} = \dim V^t.$$

Now dim $V^t = dt$ and $R \subseteq \ker W_{c,\Phi}$, hence as $\dim R = d$ we have $\dim \ker W_{c,\Phi} \geq d$. Thus rank $W_{c,\Phi} \leq d(t-1)$, with equality if and only if $\ker W_{c,\Phi} = R$. Now $\ker W_{c,\Phi}$ strictly contains R, if and only if there exists $P \notin R$ with $P \in \ker W_{c,\Phi}$, if and only if $w_{c_1,\Phi}, \ldots, w_{c_k,\Phi} \subseteq U_P$, if and only if $W_{c,\Phi} \subseteq U_P$. So $\ker W_{c,\Phi} = R$ if and only if for all $P \notin R$, $W_{c,\Phi} \not\subseteq U_P$. □

Let I be a subset of $T = \{1, \ldots, t\}$. If c is a pair sequence on T, define c_I to be the subsequence of c which contains the pairs of c whose entries lie entirely in I. So c_I is a pair sequence on I. Similarly, if Φ is a functional sequence in correspondence with c, let Φ_I be the subsequence of Φ consisting of those elements of Φ corresponding to the pairs in c_I. That is, if c_I is c without the pairs in positions $i_1, \ldots, i_j$, then Φ_I is Φ without the functionals in positions $i_1, \ldots, i_j$. For example, let $T = \{1,2,3,4\}$, $c = ((1,2),(1,3),(1,4),(2,3),(2,4))$, $\Phi = (\phi_1, \phi_2, \phi_3, \phi_4, \phi_5)$. Let $I = \{1,2,3\}$. Then $c_I = ((1,2),(1,3),(2,3))$ and $\Phi_I = (\phi_1, \phi_2, \phi_4)$.

We remind the reader that a k-pair sequence $c = ((a_1, b_1), \ldots, (a_k, b_k))$ on T may contain repeated pairs. We say c is (d,t)-*limited* if for every t' $(2 \leq t' < t)$, and t'-subset I of T, the number of pairs in c_I is *strictly less* than $d(t'-1)$. That is, $|c_I| < d(|I|-1)$ for all proper subsets $I \subset T$.

The concept of (d,t)-limited sequences is fundamental to our technique for constructing perfect hash families. This concept was not used in Blackburn and Wild [9]; it is important for our constructions as it considerably reduces the number of sequences to be considered in the construction of a perfect hash family.

Lemma 2.9 *Consider a (q^d, q, t)-perfect hash family with associated linear functional sequence Φ. Let c be a $d(t-1)$-pair sequence. Let c' be a subsequence of c with $d(t-1)-1$ pairs. If c' is (d,t)-limited, then* rank $W_{c,\Phi} = d(t-1)$.

Proof Suppose that c' is a (d,t)-limited sequence and let Φ' be the subsequence of Φ corresponding to c'. That is, if c' is c without the i^{th} pair, then Φ' is Φ without the i^{th} functional. We prove this result by contradiction, so suppose that rank $W_{c,\Phi} < d(t-1)$. By Lemma 2.8 there exists $P = (p_1, \ldots, p_t) \in V^t \setminus R$ with $W_{c,\Phi} \subseteq U_P$. Note that $W_{c',\Phi'} \subseteq W_{c,\Phi}$, so we have $W_{c',\Phi'} \subseteq U_P$. As Φ is a perfect hash family, by Result 2.3, P can not have distinct elements. So P has some repeated elements, but at least two different elements, as $P \notin R$. Suppose P has r $(1 < r < t)$ distinct elements $q_1, \ldots, q_r$. For each $i = 1, \ldots, r$, let $I_i \subseteq T$ be the set $\{j : p_j = q_i\}$. Let $x_i = |I_i|$, the number of times q_i occurs in P, so $1 \leq x_i < t$. If no $x_i = 1$, set s to 0. Otherwise suppose without loss of generality that $1 = x_1 = x_2 = \cdots = x_s$, and $x_i > 1$ for $i = s+1, \ldots, r$. Note that $s < r$ and

$$s + \sum_{i=s+1}^{r} x_i = t. \qquad (*)$$

Our aim is to show that the number of pairs in c (denoted $|c|$) is strictly bounded above by $d(t-1)$, contradicting c being of length $d(t-1)$. We partition the pairs of c' into $r+1$ categories $(0), (1), \ldots, (r)$. Category (i) consists of the pairs (a,b) of c' where $a, b \in I_i$, $(1 \leq i \leq r)$. If $x_i = 1$, then there are no pairs in this category (as a pair has two distinct entries), so $|c'_{I_i}| = 0$ for $i = 1, \ldots, s$. Otherwise, as c' is (d,t)-limited, we have a bound on the number of pairs in category (i):

$$|c'_{I_i}| \leq d(x_i - 1) - 1 \quad \text{for } i = s+1, \ldots, r.$$

Category (0) consists of the remaining pairs (a,b) of c', so $a \in I_i$, $b \in I_j$ and $i \neq j$. We use these pairs to define a new pair sequence c'_0. For each pair (a,b) of c' of category (0), define a corresponding pair (i,j) of c'_0, where $a \in I_i$, $b \in I_j$. By definition of category (0), $i \neq j$, so without loss of generality, write these pairs as (i,j) with $i < j$. Let Φ'_0 be those elements of Φ' corresponding to the pairs in category (0). Let $Q = (q_1, \ldots, q_r)$. As $W_{c',\Phi'} \subseteq U_P$, we have $W_{c'_0,\Phi'_0} \subseteq U_Q$. Now Q has distinct elements, so by Result 2.3, Φ'_0 does not separate Q, and so Φ'_0 is not a (q^d, q, r)-perfect hash family. Hence by Theorem 2.7, $|\Phi'_0| < d(r-1)$. Thus the number of pairs in category (0) is $|c'_0| \leq d(r-1) - 1$.

We are now ready to bound the length of c', using the bounds on the pairs in categories $(0), (1), \ldots, (r)$. We have

$$|c'| \;\leq\; d(r-1) - 1 + \sum_{i=s+1}^{r} (d(x_i - 1) - 1)$$

$$
\begin{aligned}
&= dr - d - 1 + \left(\sum_{i=s+1}^{r} dx_i \right) - (r-s)d - (r-s) \\
&= d\left(s + \sum_{i=s+1}^{r} x_i \right) - d - 1 - (r-s) \\
&= d(t-1) - 1 - (r-s) \qquad \text{by (*)} \\
&\leq d(t-1) - 2
\end{aligned}
$$

as $s < r$. As $|c| = |c'| + 1$, this contradicts c being of length $d(t-1)$. Hence we must have rank $W_{c,\Phi} = d(t-1)$ as required. □

Lemma 2.10 *Let Φ be a $d(t-1)$-functional sequence with the property that every $d(t-2)$-subset of Φ is a $(q^d, q, t-1)$-perfect hash family. Let c be a $d(t-1)$-pair sequence on T which is not (d,t)-limited. Then $V_{(a,b)} \subseteq W_{c,\Phi}$ for some pair (a,b) on T.*

Proof As c is not (d,t)-limited, let t' be the smallest value ($2 \leq t' < t$) with a t'-subset I of T satisfying $|c_I| \geq d(t'-1)$. Choose any subsequence c' of c_I of length $d(t'-1)$, and let Φ' be the corresponding subsequence of Φ_I. By choice of t', c' is (d,t')-limited. As $t' < t$, by Theorem 2.7, Φ' is a (q^d, q, t')-perfect hash family. Applying Lemma 2.9 with Φ' and c', we have rank $W_{c',\Phi'} = d(t'-1)$. By Lemma 2.4, there exists $a, b \in I$ with $V_{(a,b)} \subseteq W_{c',\Phi'} \subseteq W_{c,\Phi}$, as required. □

The following is our first theorem that characterises perfect hash families in terms of sequences. In [9] the forward implication was indicated, but we include the proof for completeness.

Theorem 2.11 *Suppose $2 \leq t \leq q$. Suppose S is a set of $d(t-1)$ linear functionals with associated functional sequence Φ. Then S is an optimal linear (q^d, q, t)-perfect hash family if and only if, for all $|S|$-pair sequences c we have either: (a) there exists a pair (a,b) on T with $V_{(a,b)} \subseteq W_{c,\Phi}$, or (b)* rank $W_{c,\Phi} = d(t-1)$.

Proof ($\Rightarrow$) Suppose that S is a perfect hash family and let c be a $|S|$-pair sequence. If c is not (d,t)-limited, then (a) holds by Theorem 2.7 and Lemma 2.10. If c is (d,t)-limited, then every $(d(t-1)-1)$-pair subsequence of c is (d,t)-limited and so (b) holds by Lemma 2.9.

($\Leftarrow$) Suppose that S is not a perfect hash family. Then there exists a set P of t distinct elements of V that S does not separate. Thus by Result 2.3, there exists a $|S|$-pair sequence c with $W_{c,\Phi} \subseteq U_P$. If c satisfies (a), then there exists $(a,b) \in T^2$ such that $V_{(a,b)} \subseteq W_{c,\Phi} \subseteq U_P$ and so by Result 2.2, P does not have distinct elements, a contradiction. If c satisfies (b), then by Lemma 2.4 we have $V_{(a,b)} \subseteq W_{c,\Phi} \subseteq U_P$ for all $(a,b) \in T^2$, and so by Result 2.2, P does not have distinct elements, a contradiction. Thus S is a perfect hash family. □

We are interested in the following property. Let Φ be a sequence $d(t-1)$ linear functionals.

Property (*) for (Φ, d, t): *For every $d(t-1)$-pair sequence c of the form $c = (c', (a, b))$, where c' is a (d, t)-limited pair sequence of length $d(t-1)-1$, we have* rank $W_{c,\Phi} = d(t-1)$.

Lemma 2.12 *Suppose $2 \leq t \leq q$. Let Φ be a $d(t-1)$-functional sequence. Suppose Property $(*)$ for (Φ, d, t) holds. Then Φ is a (q^d, q, t)-perfect hash family.*

Proof We begin by summarising the proof, noting which parts need further argument. We prove the lemma by induction on t. The first step is to prove that it holds in the case $t = 2$ (this is Assertion 1 and is proved below).

We now assume that the lemma holds for the value $t-1$ (≥ 2), and prove the lemma for the value t. Thus, we will assume the following:

1. If Property $(*)$ holds for $(\Psi, d, t-1)$ (where Ψ is any $d(t-2)$ sequence of linear functionals), then Ψ is a $(q^d, q, t-1)$-perfect hash family.
2. Property $(*)$ holds for (Φ, d, t).

We need to show that Φ is a (q^d, q, t)-perfect hash family. To this end, we first prove that any $d(t-2)$ subset Ψ of Φ is a $(q^d, q, t-1)$-perfect hash family (this is Assertion 2 and is proved below). Assuming this, consider a $d(t-1)$-pair sequence c. If c is not (d, t)-limited, by Lemma 2.10, $V(a, b) \subseteq W_{c,\Phi}$. If c is (d, t)-limited then any subsequence is (d, t)-limited, so by Property $(*)$ for (Φ, d, t) we have rank $W_{c,\Phi} = d(t-1)$. So by Theorem 2.11, Φ is a (q^d, q, t)-perfect hash family. This completes the proof.

It now remains to prove Assertions 1 and 2.

Proof of Assertion 1: We prove that the lemma holds for $t = 2$. In this case $T = \{1, 2\}$ and c is the sequence with the pair $(1, 2)$ repeated d times. The condition that rank $W_{c,\Phi} = d$ is equivalent to every d functionals in Φ being linearly independent. By Lemma 2.6, this is equivalent to Φ being a $(q^d, q, 2)$-perfect hash family.

Proof of Assertion 2: We prove that any $d(t-2)$ subset Ψ of Φ is a $(q^d, q, t-1)$-perfect hash family.

Let $T' = \{1, \ldots, t-1\}$ and let $e = (e', (a, b))$ be a $d(t-2)$-pair sequence on T' where e' is $(d, t-1)$-limited. Extend e' to a $(d(t-1)-1)$-sequence $c' = (e', (t-2, t), (t-1, t), \ldots, (t-1, t))$ on T (that is, the pair $(t-1, t)$ occurs as the last $d-1$ pairs of c'). We will show that c' is (d, t)-limited. Note that as e' is $(d, t-1)$-limited, for any subset I' of T', we have

$$|c'_{I'}| \leq d(|I'| - 1) - 1.$$

Consider $I \subset T$, we need to show that $|c'_I| \leq d(|I|-1)-1$. There are four cases to consider:

$$(1)\ t-1 \in I, t \in I,\ (2)\ t-1 \in I, t \notin I,\ (3)\ t-1 \notin I, t \in I,\ (4)\ t-1 \notin I, t \notin I.$$

In Case (1), write $I = I' \cup \{t\}$, so $|I'| \geq 1$. If $|I'| > 1$, then $|c'_I| \leq d(|I'|-1)-1+d = d(|I|-1)-1$ as required. Otherwise $|I'| = 1$ and $I = \{t-1, t\}$ and so $|c'_I| = d-1 = d(|I|-1)-1$ as required. In Case (2), $|c'_I| = |e'_I| \leq d(|I|-1)-1$ as required. In Case (3), write $I = I' \cup \{t\}$. If $|I'| > 1$, then $|c'_I| \leq d(|I'|-1)-1+1 \leq d(|I|-1)-1$. If $|I'| = 1$ then $|c'_I| \leq 1 \leq d(|I|-1)-1$, as required. In Case (4), $|c'_I| = |e'_I| \leq d(|I|-1)-1$, as required.

Hence, c' is (d,t)-limited. Thus by Property $(*)$ for (Φ, d, t), if $c = (c', (a,b))$, then rank $W_{c,\Phi} = d(t-1)$. So if Ψ is any $d(t-2)$ subset of Φ, then as e is a subsequence of c and $W_{c,\Phi}$ has maximal rank, we deduce that $W_{e,\Psi}$ has maximal rank. That is, rank $W_{e,\Psi} = d(t-2)$, that is, Property $(*)$ holds for $(\Psi, d, t-1)$. So by the inductive hypothesis, Ψ is a $(q^d, q, t-1)$-perfect hash family. □

Theorem 2.13 *Suppose $2 \leq t \leq q$. Suppose S is a set of $d(t-1)$ linear functionals with associated functional sequence Φ. Then S is a (q^d, q, t)-perfect hash family if and only if both*

1. *Every $d(t-2)$-subset of S is a $(q^d, q, t-1)$-perfect hash family.*
2. *For all (d,t)-limited $|S|$-pair sequences c we have* rank $W_{c,\Phi} = d(t-1)$.

Proof The forward direction follows immediately from Theorem 2.7 and Lemma 2.9. The reverse direction follows from Lemma 2.10 and Theorem 2.11. □

We are now able to prove the main theorem which characterizes when a functional sequence Φ is a perfect hash family in terms of the rank of $W_{c,\Phi}$ and (d,t)-limited pair sequences.

Theorem 2.14 *Suppose $2 \leq t \leq q$. Suppose S is a set of $d(t-1)$ linear functionals with associated functional sequence Φ. Then S is an optimal linear (q^d, q, t)-perfect hash family if and only if, for all $|S|$-pair sequences of the form $c = (c', (a,b))$ where c' is (d,t)-limited, and all pairs (a,b) on T, we have*

$$\text{rank } W_{c,\Phi} = d(t-1).$$

Proof The forward direction follows immediately from Lemma 2.9. The reverse direction is Lemma 2.12. □

2.2 Methods to construct linear perfect hash families

These results give us two approaches to constructing perfect hash families. Both these methods involve finding a set of $d(t-1)$ linear functionals that satisfy Theorem 2.14.

Before describing these methods, we first note that if c_1, c_2 are two k-pair sequences, then we can determine whether they are isomorphic under the induced action of the symmetric group S_t acting on the collection of k-pair sequences. If c_1 and c_2 are isomorphic, then for any k-functional sequence Φ, $W_{c_1,\Phi}$ is a column permutation of $W_{c_2,\Phi}$ and so rank $W_{c_1,\Phi}$ = rank $W_{c_2,\Phi}$. Thus, to show that a $d(t-1)$-functional sequence Φ satisfies Theorem 2.14, it is sufficient to show that it satisfies the rank condition for one representative of each isomorphism group under the induced action of S_t of the sequences $c = (c', (a, b))$ with c' a (d, t)-limited $(d(t-1)-1)$-pair sequence. We will call *representative* k-pair sequences the collection of one sequence from each isomorphism group. In the next sections, we present some examples and show how to find the representative pair sequences.

Method 1. The first method we could use to construct a perfect hash family involves guessing a $d(t-1)$-functional sequence Φ, and then checking if it is a linear (q^d, q, t)-perfect hash family. To do this we check that for each representative $d(t-1)$-pair sequence c with a subsequence c' of length $d(t-1)-1$ which is (d, t)-limited, we have rank $W_{c,\Phi} = d(t-1)$. Then by Theorem 2.14, Φ is an optimal linear (q^d, q, t)-perfect hash family. This method is not very practical for small q as it will result in a high failure rate. However, for large q (relative to d, t) there are many perfect hash families and this guessing method has a better chance of succeeding.

Method 2. We now present a method that gives us a technique for recursively building a perfect hash family for any q, d, t where they exist. This method has two main steps. The first step is to find an appropriate $d(t-1)-1$ functional sequence Φ'. We say a $(d(t-1)-1)$-functional sequence Φ' is (d, t)-*suitable* if

(i) every $d(t-2)$ subset of Φ' is a $(q^d, q, t-1)$-perfect hash family and

(ii) rank $W_{c',\Phi'} = d(t-1)-1$ for all (representative) (d, t)-limited sequences c' of length $d(t-1)-1$.

The second step is to consider all (d, t)-limited sequences c' of length $d(t-1)-1$, and all pairs (a, b) and find the functionals ϕ for which

$$\text{rank } W_{c',\Phi'} = \text{rank } W_{(c',(a,b)),(\Phi',\phi)}.$$

We call such a functional ϕ *excluded*. These functionals ϕ cannot be added to Φ' to form a perfect hash family. We show that any other functional in V^* will complete Φ' to a perfect hash family.

Theorem 2.15 *Let $2 \leq t \leq q$. Let Φ' be a (d,t)-suitable $(d(t-1)-1)$-functional sequence. For any linear functional ϕ which is not excluded, (Φ', ϕ) forms a (q^d, q, t)-perfect hash family. Further, every (q^d, q, t)-perfect hash family can be constructed in this way.*

Proof As Φ' is (d,t)-suitable, rank $W_{c',\Phi'} = d(t-1)-1$ for all (d,t)-limited $(d(t-1)-1)$-pair sequences c'. If ϕ is not excluded, then rank $W_{(c',(a,b)),(\Phi',\phi)} = \text{rank } W_{c',\Phi'} + 1 = d(t-1)$. Hence, by Theorem 2.14, (Φ', ϕ) is a (q^d, q, t)-perfect hash family.

Conversely, let Φ be a (q^d, q, t)-perfect hash family. Let $\Phi = (\Phi', \phi)$, where Φ' is the sequence of the first $d(t-1)-1$ elements of Φ and ϕ is the remaining linear functional. Then Φ' is (d,t)-suitable by Theorems 2.7 and 2.14, and the functional ϕ is not excluded by Theorem 2.14. □

It is important to note that by definition, $W_{c,\Phi}$ and its rank is dependent on the value of q. We also note that there are two types of excluded functionals. In the case that $(c',(a,b))$ is (d,t)-limited, we say that ϕ is *limit-excluded*. Now suppose $(c',(a,b))$ is not (d,t)-limited. As c' is (d,t)-limited, this means there exists $I \subset T$ with $|(c',(a,b))_I| \geq d(|I|-1)$. If ϕ is a linear functional with rank $W_{c',\Phi'} = \text{rank } W_{(c',(a,b)),(\Phi',\phi)}$, then (Φ'_I, ϕ) is not a $(q^d, q, |I|-1)$-perfect hash family, for I with $|I| < t$. We call ϕ *sub-excluded* (reflecting a *sub*set of a perfect hash family also being a perfect hash family). Thus the sub-excluded functionals are the functionals which are excluded because adding them would result in a functional sequence contradicting Theorem 2.7.

In previous sections, we have used c, Φ and c', Φ' to represent sequences of length $d(t-1)$ and $d(t-1)-1$ respectively. In the construction of perfect hash families, we are only concerned with sequences of length $d(t-1)-1$. Hence from now on we will omit the $'$, and use c and Φ to denote sequences of length $d(t-1)-1$. We summarise our method using this new notation.

Algorithm to find a (q^d, q, t)-perfect hash family

1. *Suitability Phase:* Find a (d,t) suitable $d(t-1)-1$ functional sequence Φ, that is, so that
   (i) every $d(t-2)$ subsequence of Φ is a $(q^d, q, t-1)$-perfect hash family;
   (ii) for every (d,t)-limited $(d(t-1)-1)$-pair sequence c, we have rank $W_{c,\Phi} = d(t-1)-1$.
2. *Exclusion Phase:* Find the (representative) (d,t)-limited $(d(t-1)-1)$-pair sequences c. For every (representative) $d(t-1)$-pair sequence $(c,(a,b))$, calculate the excluded functionals of Φ. That is, find the linear functionals ϕ with rank $W_{c,\Phi} = \text{rank } W_{(c,(a,b)),(\Phi,\phi)}$
3. *Construction Phase:* We can now add any functional of V^* that is not excluded to Φ and we will have an optimal linear (q^d, q, t)-perfect hash family.

In the following sections we give examples which demonstrate how to find suitable functional sequences and representative pair sequences.

2.3 Sequence condition for every d functionals linearly independent

We can simplify the search for perfect hash families by using the fact that every d linear functionals in a linear (q^d, q, t)-perfect hash family are linearly independent (by Result 2.5). We now show that this condition relates to the sub-excluded linear functionals.

We consider the case $d = 2$. An optimal linear (q^2, q, t)-perfect hash family S has size $2(t-1)$. A linear functional $\phi \in V^*$ can be represented by $\phi = [a, b]$ where $a, b \in \mathrm{GF}(q)$ (we will use square brackets to denote linear functionals). Suppose that c is a $(2, t)$-limited sequence of length $2(t-1)-1$, and suppose the pair (a, b) occurs $d-1 = 2-1 = 1$ times in c. Then $(c, (a, b))$ is a $2(t-1)$-pair sequence that is not $(2, t)$-limited (since if we pick $I = \{a, b\} \subset T$, then $|c_I| = 2 = d(|I| - 1)$). Hence this sequence leads to sub-excluded linear functionals. Suppose the pair (a, b) occurs in position i in c. Let $\Phi = (\phi_1, \ldots, \phi_{t-1})$ be a $(2, t)$-suitable $(2(t-1)-1)$-functional sequence and consider the matrix $W_{(c,(a,b)),(\Phi,\phi_t)}$. Row i has functional ϕ_i in column a, and functional $-\phi_i$ in column b. Row t has functional ϕ_t in column a, and functional $-\phi_t$ in column b. Hence in order for this matrix to have maximal rank $2(t-1)$, we must have all rows linearly independent, that is, we must have ϕ_i and ϕ_t linearly independent. That is, the functionals which are linearly dependent on a functional in Φ are sub-excluded functionals. Hence, if $\phi = [a, b] \in \Phi$ then $\alpha\phi = [\alpha a, \alpha b]$ is a sub-excluded linear functional. Thus, without loss of generality, in the case $t = 2$ we only consider linear functionals of the form $\phi = [1, \alpha]$ where $\alpha \in \mathrm{GF}(q) \cup \{\infty\}$ (where $\phi = [1, \infty]$ is used to represent the linear functional $[0, 1]$).

For general d, let Φ be a (d, t)-suitable $(d(t-1)-1)$-functional sequence. If c is a (d, t)-limited sequence of length $d(t-1)-1$ with the pair (a, b) occurring at least $d-1$ times in c (and hence exactly $d-1$ times), then $(c, (a, b))$ is not (d, t)-limited. Thus those functionals which are linearly dependent on $d-1$ functionals of Φ are sub-excluded linear functionals.

3 Case $d = 2,\ t = 4$

We will use the sequence theory developed above to show how to construct optimal linear $(q^2, q, 4)$-perfect hash families. We take a brief look at this case for three reasons. Firstly, this case was considered using a geometrical approach in [3] and we want to illustrate the difference between the geometric approach and the sequence approach. Secondly, we want to explain

the correspondence between this case and a case of $d = 3$, $t = 3$ in Section 5. Thirdly, when considering the $d = 2$, $t = 5$ case in Section 4, we will see that (as implied by Theorem 2.7) the results for $t = 4$ appear as part of the conditions for the $t = 5$ case.

3.1 $(q^2, q, 3)$-perfect hash families

In order to construct a $(q^2, q, 4)$-perfect hash family S, we need to find a $(2, 4)$-suitable 5-functional sequence Φ, so we need every $d(t-2) = 2(4-2) = 4$ subset of Φ to be a $(q^2, q, 3)$-perfect hash family. We show here that a set of 4 functionals is an optimal linear $(q^2, q, 3)$-perfect hash family if and only if they are pairwise linearly independent.

By the argument in Section 2.3, we only need consider functionals of the form $[1, \alpha]$, $\alpha \in \mathrm{GF}(q) \cup \{\infty\}$. To construct a $(q^2, q, 3)$-perfect hash family, we begin with $T = \{1, 2, 3\}$ and $T^2 = \{(1, 2), (1, 3), (2, 3)\}$. We need to construct a $(2, 3)$-limited sequence c of length $d(t-1)-1 = 3$; c will be $(2, 3)$-limited if each pair from T^2 occurs exactly once in c. Consider the 3-functional sequence $\Phi = ([1, u], [1, v], [1, w])$. We need Φ to be $(2, 3)$-suitable. Firstly, by Lemma 2.6, every 2-subsequence of Φ is a $(q^2, q, 2)$-perfect hash family if every 2 functionals of Φ are linearly independent, that is, we need u, v, w distinct. Secondly, we need to check that rank $W_{c,\Phi} = 3$ for every $(2, 3)$-limited sequence c. For $c = ((1, 2), (1, 3), (2, 3))$, we have

$$W_{c,\Phi} = \left[\begin{array}{cc|cc|cc} 1 & u & -1 & -u & 0 & 0 \\ 1 & v & 0 & 0 & -1 & -v \\ 0 & 0 & 1 & w & -1 & -w \end{array}\right].$$

This has rank 3 provided u, v, w are distinct. Further, permuting pairs in c is equivalent to permuting u, v, w in the matrix, so $W_{c,\Phi}$ has rank 3 for any $(2, 3)$-limited sequence c. Thus $\Phi = ([1, u], [1, v], [1, w])$ is $(2, 3)$-suitable if u, v, w are distinct.

We now calculate the excluded functionals. We can complete c to a 4-pair sequence in 3 ways. Note that none of these gives a $(2, 3)$-limited sequence, so there are no limited-excluded functionals. If we add the pair $(1, 2)$, then we have the matrix

$$W_{(c,(1,2)),(\Phi,[1,x])} = \left[\begin{array}{cc|cc|cc} 1 & u & -1 & -u & 0 & 0 \\ 1 & v & 0 & 0 & -1 & -v \\ 0 & 0 & 1 & w & -1 & -w \\ 1 & x & -1 & -x & 0 & 0 \end{array}\right].$$

Thus rank $W_{(c,(1,2)),(\Phi,[1,x])}$ = rank $W_{c,\Phi}$ if and only if $x = u$. That is, the functional $[1, u]$ is excluded. Similarly, adding the pairs $(1, 3)$ or $(2, 3)$ to c result in the excluded linear functionals $[1, v], [1, w]$ respectively. Hence we can add any remaining functional in V^* to Φ and get a perfect hash family. That is, $\{[1, u], [1, v], [1, w], [1, x]\}$ is an optimal linear $(q^2, q, 3)$-perfect hash family if and only if u, v, w, x are distinct.

3.2 Constructing $(q^2, q, 4)$-perfect hash families

We will use the algorithm described in Section 2.2 to construct a perfect hash family. As explained in Section 2.3, we will assume the linear functionals are of the form $\phi = [1, \alpha]$ where $\alpha \in \mathrm{GF}(q) \cup \{\infty\}$. We have $T^2 = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} : 1 \leq a < b \leq 4\} = \{(1,2), (1,3), (1,4), (2,3), (2,4), (3,4)\}$. Note that a $(2,4)$-limited k-pair sequence on T uses each pair on T at most once.

We briefly outline the steps we need to take following the algorithm:

1. We find a $(2,4)$-suitable 5-functional sequence Φ.
2. We find all the representative $(2,4)$-limited 5-pair sequences c on T, these consist of 5 distinct pairs from T^2. We can complete such a sequence to a 6-pair sequence in 6 ways. Five of these ways involve repeating a pair, and so lead to a sequence that is not $(2,4)$-limited. Since we have chosen our functionals to have form $[1, \alpha]$, the resulting sub-excluded functionals in these cases are just the functionals in Φ. There is a unique way to complete a $(2,4)$-limited 5-pair sequence c to a $(2,4)$-limited 6-pair sequence $(c, (a_6, b_6))$. For each representative $(2,4)$-limited 5-pair sequence c, we form $W_{c,\Phi}$ and row reduce to calculate the limit-excluded linear functional ϕ corresponding to the remaining pair (a_6, b_6) on T.
3. Once we have calculated all the excluded linear functionals, any of the remaining linear functionals in $V^* \setminus \Phi$ of the form $[1, \alpha]$, $\alpha \in \mathrm{GF}(q) \cup \{\infty\}$ can be added to Φ to form a perfect hash family.

We first count the number of representative $(2,4)$-limited 5-pair sequences on T. The action of the symmetric group S_4 on $1, 2, 3, 4$ induces an action on T. For example, the mapping $\sigma = (123)$ (that is, $1 \mapsto 2 \mapsto 3 \mapsto 1$) maps the pair $(2,3)$ on T onto $(3,1) \equiv (1,3)$. Without loss of generality, we can assume that a representative 5-pair sequence contains the 5 pairs $\mathcal{Q} = \{(1,2), (1,3), (1,4), (2,3), (2,4)\}$. So the subgroup $\langle (1,2), (3,4) \rangle$ of S_4 of size 4 fixes the set $\mathcal{Q}$. The number of representative 5-pair sequences consisting of the pairs $\{(1,2), (1,3), (1,4), (2,3), (2,4)\}$ under S_4 is therefore $5!/4 = 30$.

We choose a general 5-functional sequence Φ and consider the following 5-pair sequence c:

$$\begin{array}{rcccccc} \Phi & = & ([1,v], & [1,w], & [1,x], & [1,y], & [1,z]) \\ c & = & ((1,2), & (1,3), & (1,4), & (2,3), & (2,4)). \end{array}$$

We need to construct $W_{c,\Phi}$. Recall that $w_{(1,2),\phi} = (\phi, -\phi, 0, 0)$, thus the first row of $W_{c,\Phi}$ is

$(1, v, -1, -v, 0, 0, 0, 0)$. Similarly we obtain all the $w_{c,\phi}$ and so $W_{c,\Phi}$ written as a matrix is:

$$W_{c,\Phi} = \left[\begin{array}{cc|cc|cc|cc} 1 & v & -1 & -v & 0 & 0 & 0 & 0 \\ 1 & w & 0 & 0 & -1 & -w & 0 & 0 \\ 1 & x & 0 & 0 & 0 & 0 & -1 & -x \\ 0 & 0 & 1 & y & -1 & -y & 0 & 0 \\ 0 & 0 & 1 & z & 0 & 0 & -1 & -z \end{array}\right].$$

We need Φ to be $(2,4)$-suitable. Firstly, we need every 4-subset of Φ to be a $(q^2, q, 3)$-perfect hash family. That is, by Section 3.1, we need every 4 elements of v, w, x, y, z to be distinct. Secondly, we need the rank of $W_{c,\Phi}$ to be 5 for all $(2,4)$-limited 5-pair sequences c. By inspection the matrix above has rank 5 if v, w, x, y, z are distinct. Now, different representative 5-pair sequences correspond to permuting the pairs in c and so correspond to permuting v, w, x, y, z in this matrix. Hence for any representative 5-pair sequence c, rank $W_{c,\Phi} = 5$. Thus any set of 5 distinct values of v, w, x, y, z results in a $(2,4)$-suitable 5-functional sequence Φ.

We can now easily see that completing c to a 6-pair sequence by adding a pair already in c results in the sub-excluded functionals of form $[1, \alpha]$ being the functionals already in Φ (that is, the functionals $[1, v], [1, w], [1, x], [1, y], [1, z]$). There is a unique way to complete c to a $(2,4)$-limited 6-pair sequence, namely by adding the pair $(3,4)$. We now perform row operations on $W_{c,\Phi}$ to find the functional $[1, \theta]$ such that $w_{(3,4),[1,\theta]}$ is in the row space of $W_{c,\Phi}$. That is, we perform row operations to construct a row $[0, 0, 0, 0, 1, \theta, -1, -\theta]$. For the above matrix, the linear functional corresponding to the final pair $(3,4)$ is $[1, \theta]$ where

$$\theta \;=\; -\frac{xy(v-w-z) + (x+y)wz - vwz}{xy - (x+y)v + zv - zw + wv}.$$

Thus we have rank $W_{c,\Phi} = \text{rank } W_{(c,(3,4)),(\Phi,[1,\theta])} = 5$ and so we cannot add the linear functional $[1, \theta]$ to our set $\{[1, v], [1, w], [1, x], [1, y], [1, z]\}$ to obtain a perfect hash family. That is, $[1, \theta]$ is a limit-excluded linear functional.

There are 30 representative 5-pair sequences, so there are 30 limit-excluded functionals. However, we see that θ is symmetrical in x and y (with corresponding pairs $(1,4)$ and $(2,3)$) and so two sequences with the pairs $(1,4)$ and $(2,3)$ swapped yield the same limit-excluded linear functional $[1, \theta]$. This gives us a total of 15 representative 5-pair sequences to consider. These can be calculated directly by hand; Table 1 lists 15 representative 5-pair sequences. (Note also that θ is symmetrical in w and z but this does not further reduce the number of sequences to be considered.)

To construct a perfect hash family, we choose the 5-functional sequence to be

$$\Phi = ([0, 1], [1, 0], [1, 1], [1, a], [1, b]).$$

Table 1: The 15 representative sequences $c_1, \ldots, c_{15}$ for $d = 2$, $t = 4$

c_1	$((1,2),(1,3),(1,4),(2,3),(2,4))$	c_9	$((1,3),(1,4),(2,3),(1,2),(2,4))$
c_2	$((1,2),(1,3),(1,4),(2,4),(2,3))$	c_{10}	$((1,3),(1,4),(2,3),(2,4),(1,2))$
c_3	$((1,2),(1,3),(2,4),(1,4),(2,3))$	c_{11}	$((1,3),(1,4),(2,4),(1,2),(2,3))$
c_4	$((1,3),(1,2),(1,4),(2,3),(2,4))$	c_{12}	$((1,3),(1,4),(2,4),(2,3),(1,2))$
c_5	$((1,3),(1,2),(1,4),(2,4),(2,3))$	c_{13}	$((1,3),(2,4),(1,2),(1,4),(2,3))$
c_6	$((1,3),(1,2),(2,4),(1,4),(2,3))$	c_{14}	$((1,3),(2,4),(1,4),(1,2),(2,3))$
c_7	$((1,3),(1,4),(1,2),(2,3),(2,4))$	c_{15}	$((1,3),(2,4),(1,4),(2,3),(1,2))$
c_8	$((1,3),(1,4),(1,2),(2,4),(2,3))$		

We can do this without loss of generality as the group $\mathrm{GL}(2,q)$ of linear functionals of $\mathrm{GF}(q)^2$ is 3-transitive on a one dimensional subspace (see for example [13]). If we have $a \neq b$ and $a, b \neq 0, 1, \infty$ then Φ is $(2,4)$-suitable.

We want to find the limit-excluded functional for each representative 5-pair sequence c in Table 1. Rather than performing row operations on a new $W_{c,\Phi}$ for each sequence c, we use the $W_{c,\Phi}$ matrix above, and note that changing the sequence c by permuting the pairs in c is equivalent to permuting v, w, x, y, z in the matrix. Hence it is equivalent to permuting v, w, x, y, z in the excluded functional $[1, \theta]$. So for the first sequence $c_1 = ((1,2),(1,3),(1,4),(2,3),(2,4))$, we have $v = \infty$, $w = 0$, $x = 1$, $y = a$, $z = b$, hence $\theta = -a/(b-a-1)$. Thus we obtain the limit-excluded linear functional $\phi_1 = [1, -a/(b-a-1)]$ which appears against c_1 in Table 2. For the second sequence $c_2 = ((1,2),(1,3),(1,4),(2,4),(2,3))$, we have $v = \infty$, $w = 0$, $x = 1$, $y = b$, $z = a$ and so $\theta = -b/(a-b-1)$, giving the limit-excluded linear functional $\phi_2 = [1, -b/(a-b-1)]$. Similarly, we obtain ϕ_3. For the remaining 12 sequences c_4 to c_{15}, the pair $(1,3)$ corresponds to ∞, that is $w = \infty$. Thus in these cases θ becomes $-(-xy + zx - zv + yz)/(-z + v)$. So, for example, for the sequence $c_4 = ((1,3),(1,2),(1,4),(2,3),(2,4))$, we have $w = \infty$, $v = 0$, $x = 1$, $y = a$, $z = b$ and hence $\theta = -(-a + b + ab)/(-b)$. Table 2 lists the 15 limit-excluded linear functionals obtained.

Table 2: The 15 excluded linear functionals for $d = 2$, $t = 4$

c_1	$\phi_1 = [1, -a/(-a+b-1)]$	c_9	$\phi_9 = [1, b(a-1)/(a-b)]$
c_2	$\phi_2 = [1, -b/(-b+a-1)]$	c_{10}	$\phi_{10} = [1, a(b-1)/(b-a)]$
c_3	$\phi_3 = [1, ab/(a+b-1)]$	c_{11}	$\phi_{11} = [1, (a-b)/(a-1)]$
c_4	$\phi_4 = [1, (ab-a+b)/b]$	c_{12}	$\phi_{12} = [1, (b-a)/(b-1)]$
c_5	$\phi_5 = [1, (ab+a-b)/a]$	c_{13}	$\phi_{13} = [1, ab]$
c_6	$\phi_6 = [1, -ab+a+b]$	c_{14}	$\phi_{14} = [1, b/a]$
c_7	$\phi_7 = [1, b(a-1)/(b-1)]$	c_{15}	$\phi_{15} = [1, a/b]$
c_8	$\phi_8 = [1, a(b-1)/(a-1)]$		

We note that these 15 limit-excluded linear functionals correspond to the 15 linear functionals calculated via geometrical means in [3]. Constructions there used these 15 excluded linear functionals to show the following existence result.

Theorem 3.1 *Optimal linear $(q^2, q, 4)$-perfect hash families exist for all prime powers q except when $q = 2, 3, 4, 5, 7, 8, 9, 13$.*

4 Case $d = 2$, $t = 5$

An optimal linear $(q^2, q, 5)$-perfect hash family has size $d(t-1) = 8$. Theorem 1.2 shows that they exist for $q \geq 10^8$. In this section we prove an existence bound that is much smaller. Further, we construct optimal linear $(q^2, q, 5)$-perfect hash families for much smaller q than the known examples.

We first list the known constructions of optimal linear $(q^2, q, 5)$-perfect hash families. They are all based on chains of subfields. The Blackburn-Wild [9] construction gives optimal linear $(q^2, q, 5)$-perfect hash families for q of the form $q = q_0^{\alpha_1\alpha_2\cdots\alpha_6}$ where q_0 is any prime power and each $\alpha_i \geq 2$. In [4], optimal linear $(q^2, q, 5)$-perfect hash families are constructed for $q = r^2$, r any prime with $\mathrm{char}(r) \geq 31$; and for $q = r^4$, r any prime such that $r \geq 11$ $(r \neq 13)$.

4.1 An Existence Bound

We use the algorithm given in Section 2.2. As noted in Section 2.3, we only need to consider linear functionals of the form $[1, \alpha]$, $\alpha \in \mathrm{GF}(q) \cup \{\infty\}$ when constructing our perfect hash family. We are interested in $(2,5)$-limited 7-pair sequences from $T^2 = \{(1,2), (1,3), (1,4), (1,5), (2,3), (2,4), (2,5), (3,4), (3,5), (4,5)\}$. So we only need to consider 7-pair sequences on T where each pair from T^2 occurs at most once. Since $|T^2| = 10$, we can complete such a sequence to an 8-pair sequence in 10 ways, but in only 3 ways without repeating a pair.

We wish to count the number of representative $(2,5)$-limited 7-pair sequences so that we can get a bound on the number of excluded functionals. Note that different sequences may result in the same excluded functional, so this count gives us an upper bound. Let c be a $(2,5)$-limited 7-pair sequence on T. Then, under the action of the symmetry group S_5, the remaining three pairs in T^2 are one of four types: A, B, C or D which are given in Table 3.

We now provide an upper bound on the total number of sequences of each type.

A. This sequence c contains the pairs $\{(1,5), (2,3), (2,4), (2,5), (3,4), (3,5), (4,5)\}$. It is not

Table 3: Sequence types for $d = 2$ and $t = 5$

Case	remaining 3 pairs
A	$\{(1,2),(1,3),(1,4)\}$
B	$\{(1,2),(1,3),(4,5)\}$
C	$\{(1,2),(1,3),(2,3)\}$
D	$\{(1,2),(1,3),(2,4)\}$

$(2,5)$-limited as taking $I = \{2,3,4,5\}$ results in $|c_I| = 2(|I| - 1) = 6$. So we can ignore this case.

B. To preserve a sequence of type B, we can permute $2,3$ and $4,5$. Hence sequences of type B are still type B under the action of the permutation group $\langle (23),(45) \rangle$ of size 4. This group has 4 elements, hence there are $7!/4 = 1260$ representative sequences of type B.

C. Sequences of type C remain type C under the action of the group $S_3 \times S_2$ of size 12 permuting $1,2,3$ and $4,5$. Hence there are $7!/12 = 420$ representative sequences of type C.

D. A sequence of type D remains a sequence of type D under the action of the permutation group $\langle (12)(34) \rangle$ of size 2. Hence there are $7!/2 = 2520$ representative sequences of type D.

This gives a total of $1260 + 420 + 2520 = 4200$ representative $(2,5)$-limited 7-pair sequences.

For each $(2,5)$-limited sequence c above, consider repeating a pair (a,b) already occurring in the sequence c, so the sequence $(c,(a,b))$ is not $(2,5)$-limited. By the discussion in Section 2.3, the condition generated by this sequence is that every two functionals in our perfect hash family are linearly independent. Since we only consider functionals of the form $[1,\alpha]$, the resulting sub-excluded functionals are the functionals already in our $(2,5)$-suitable 7-functional sequence Φ.

So we only need consider the case where we extend c by adding a pair not in c. Some of the resulting sequences will be $(2,5)$-limited, and some will not. So each of the 4200 representative $(2,5)$-limited 7-pair sequences can be completed to an 8-pair sequence containing distinct pairs in three ways. Hence there are 4200×3 representative 8-pair sequences. Each 8-pair sequence results in one excluded linear functional of the form $[1,\alpha]$. Hence there at most 4200×3 excluded linear functionals of the form $[1,\alpha]$. However, to prove existence of a perfect hash family, we first need to verify that there exists a $(2,5)$-suitable 7-functional sequence Φ. So we need to show that there is a 7-functional sequence Φ such that firstly, every 6 elements of Φ form a $(q^2,q,4)$-perfect hash family, and secondly, $W_{c,\Phi}$ has maximal rank 7 for any $(2,5)$-

limited 7-pair sequence c. We study the form of $W_{c,\Phi}$ in the next section and will prove in Lemma 4.2 that a $(2,5)$-suitable sequence exists if $q > 95$. The next theorem gives a bound on the existence of optimal linear $(q^2, q, 5)$-perfect hash families assuming a $(2,5)$-suitable sequence Φ exists. This extra condition is removed in Theorem 4.3.

Theorem 4.1 *If $q+1 > 7+4200\times 3 = 12607$, and there exists a $(2,5)$-suitable sequence, then there exists an optimal linear $(q^2, q, 5)$-perfect hash family.*

Proof We assume there exists a $(2,5)$-suitable 7-functional sequence Φ. In the discussion above, we showed that the 7 functionals in Φ are excluded functionals. Further, each of the 4200 representative $(2,5)$-limited 7-pair sequences results in 3 excluded functionals of the form $[1,\alpha]$. There are $q+1$ linear functionals of the form $[1,\alpha]$, $\alpha \in \mathrm{GF}(q) \cup \{\infty\}$, so provided $q+1 > 7+4200\times 3 = 12607$, there are functionals in V^* that we can add to Φ to form a $(q^2, q, 5)$-perfect hash family. □

In fact we will show that $(q^2, q, 5)$-perfect hash families exist for much smaller q than given in Theorem 4.1, since the excluded linear functionals are not necessarily distinct.

4.2 Sequence Analysis for $d = 2$ and $t = 5$

In order to construct linear $(q^2, q, 5)$-perfect hash families, we use a similar technique to that used in Section 3.2. We now study in detail the representative $(2,5)$-limited 7-pair sequences from each of the cases B, C and D given in Table 3.

Case B. Consider the 7-functional sequence Φ and 7-pair sequence c:

$$\begin{aligned} \Phi &= ([1,s], \quad [1,u], \quad [1,v], \quad [1,w], \quad [1,x], \quad [1,y], \quad [1,z]) \\ c &= ((1,4), \quad (1,5), \quad (2,3), \quad (2,4), \quad (2,5), \quad (3,4), \quad (3,5)). \end{aligned}$$

Row reducing $W_{c,\Phi}$ enables us to calculate the excluded linear functionals for each of the 3 missing pairs $(4,5), (1,2), (1,3)$. They are:

$$B_{(4,5)} = \left[1, -\tfrac{-vwz-wxy+wzx+wzy-yzx+yvx}{-vy+zv-wz+vw-vx+xy}\right],$$

$$B_{(1,2)} = \left[1, \tfrac{-wuzv+zuvs+zuwy-vwsx+xwzs+xysv-xzys+wuvx-uvsx-wuxy+yusx-uysv-swuz+swuv}{zsv+wzy-zys-vsx-wxy+ysx-yzx+yvx+wzx+yuz-uvy-wuz+wuv-vwz)}\right],$$

$$B_{(1,3)} = \left[1, \tfrac{-uysv+yusx+swuv+yzsv-xzys-wzsv+xwzs+zuvs-uvsx-yuzv+yuvx-swuz+zuwy-wuxy}{-uvy+uxy+wuv-wux-yzx-vwz+wzx+zsv-vsx+yvx-wzs+wsx+wzy-wxy}\right].$$

Case C: Consider the 7-functional sequence Φ and 7-pair sequence c:

$$\begin{aligned} \Phi &= ([1,s], \quad [1,u], \quad [1,v], \quad [1,w], \quad [1,x], \quad [1,y], \quad [1,z]) \\ c &= ((1,4), \quad (1,5), \quad (2,4), \quad (2,5), \quad (3,4), \quad (3,5), \quad (4,5)). \end{aligned}$$

The excluded linear functionals for the 3 missing pairs $(1,2),(1,3),(2,3)$ are:

$$C_{(1,2)} = \left[1, \tfrac{-(-uzv-wus+wuv+uvs-svw+wzs)}{-uv-zs+zv+ws+uz-wz}\right],$$
$$C_{(1,3)} = \left[1, \tfrac{-usx+uzx+uys-uxy+ysx-zys}{-(-ys+zy-uz+ux+zs-zx)}\right],$$
$$C_{(2,3)} = \left[1, \tfrac{-(-yzv+yvx-vwx+yvw-wxy+wzx)}{zy-wz+wx-vy+zv-zx}\right].$$

Case D: Consider the 7-functional sequence Φ and 7-pair sequence c:

$$\begin{array}{rcl} \Phi & = & ([1,s], \quad [1,u], \quad [1,v], \quad [1,w], \quad [1,x], \quad [1,y], \quad [1,z]) \\ c & = & ((1,4), \quad (1,5), \quad (2,3), \quad (2,5), \quad (3,4), \quad (3,5), \quad (4,5)). \end{array}$$

The excluded linear functionals for the 3 missing pairs $(1,2),(1,3),(2,4)$ are:

$$D_{(1,2)} = \left[1, \tfrac{-uysv+uvsx+yuzv-uzvx-vwsx-wusx-wuzv+wuzx+wuvx+wuys-wuxy+wysx+wzsv-wzys}{yuz+wys-wsx-wzy+wzx-uzv+uvx-uxy-ysv+ysx+yzv+zsv-zvx-zys}\right],$$
$$D_{(1,3)} = \left[1, \tfrac{-usx+uzx+uys-uxy+ysx-zys}{-(-ys+zy-uz+ux+zs-zx)}\right],$$
$$D_{(2,4)} = \left[1, \tfrac{-vwx-yzv+zvx+vwz+wxy-wzx}{xy+wy-wx-vy+zv-zy}\right].$$

We note that the linear functional $B_{(4,5)}$ is a sub-excluded linear functional. In this case the 8-pair sequence is $c = ((1,4),(1,5),(2,3),(2,4),(2,5),(3,4),(3,5),(4,5))$. This is not $(2,4)$-limited since if we let $I = \{2,3,4,5\}$, then $|c_I| = 6$. This corresponds to the excluded linear functionals arising from the case $d = 2, t = 4$ with the 5-functional sequence and 5-pair sequence as follows:

$$\begin{array}{rcl} \Phi'' & = & ([1,v], \quad [1,w], \quad [1,x], \quad [1,y], \quad [1,z]) \\ c'' & = & ((2,3), \quad (2,4), \quad (2,5), \quad (3,4), \quad (3,5)). \end{array}$$

Similarly, $C_{(1,2)}$, $C_{(1,3)}$, $C_{(2,3)}$, $D_{(1,3)}(= C_{(1,3)})$ and $D_{(2,4)}$ are sub-excluded linear functionals. These cases correspond to the fact that given any 5-functional subsequence of Φ, the 15 excluded linear functionals given in Table 2 for the case $d = 2, t = 4$ are sub-excluded functionals here. Thus the total number of sub-excluded functionals resulting from these cases is (at most) $\binom{7}{5} \times 15 = 350$. The remaining excluded functionals $B_{(1,2)}, B_{(1,3)}$ and $D_{(1,2)}$ are limit-excluded functionals.

This means we can improve our bound on the number of excluded linear functionals. There are 1260 excluded functionals of the form $B_{(1,2)}$ and 1260 of the form $B_{(1,3)}$. There are 2520 excluded functionals of the form $D_{(1,2)}$. There are in total at most 350 excluded functionals generated by $B_{(4,5)}$, $C_{(1,2)}$, $C_{(1,3)}$, $C_{(2,3)}$, $D_{(1,3)}$ and $D_{(2,4)}$. Finally, there are 7 functionals in Φ that are also excluded. Hence the number of excluded functionals is $1260 \times 2 + 2520 + 350 + 7 = 5387$. We can use this to improve the bound given in Theorem 4.1; the improved bound is given in Theorem 4.3.

Note also that the symmetry group $\langle(23)\rangle$ maps the sequence $((1,4),(1,5),(2,3),(2,4),(2,5),(3,4),(3,5))$ to the sequence $((1,4),(1,5),(2,3),(3,4),(3,5),(2,4),(2,5))$, hence $B_{(1,3)}$ is $B_{(1,2)}$

with the variables w and y swapped and x and z swapped. This reduces the search space when using a computer to construct a perfect hash family.

4.3 The existence of $(2, 5)$-suitable sequences

In this section we prove that there exists a $(2, 5)$-suitable sequence if $q > 95$. Moreover, the first paragraph of the proof gives a technique which we will use to construct suitable sequences.

Lemma 4.2 *If $q > 95$, there exists a $(2, 5)$-suitable sequence.*

Proof We first need to check that we can find a 7-functional sequence Φ such that every 6 linear functionals of Φ form a $(q^2, q, 4)$-perfect hash family. Secondly, we need to show that in each of the Cases B, C and D from Section 4.2, the functional Φ satisfies rank $W_{c,\Phi} = 7$. Recall by Section 2.3 we can consider the linear functionals to be of the form $[1, \alpha]$, $\alpha \in \mathrm{GF}(q) \cup \{\infty\}$. Suppose q is a prime power greater than 95, then by Theorem 3.1 there exists a linear $(q^2, q, 4)$ perfect hash family $S = \{\phi_1, \ldots, \phi_6\}$. For each 5-subset of S, the associated 5-functional sequence Φ' leads to 15 excluded linear functionals (given by Table 2). Including the original six linear functionals $\phi_1, \ldots, \phi_6$, this accounts for $6+6\times 15 = 96$ functionals. Hence if $q+1 > 96$, there is a functional $\phi \in V^*$ which is not excluded. Thus $S \cup \{\phi\}$ is a set of 7 functionals, such that every 6 form a $(q^2, q, 4)$-perfect hash family.

Let $\Phi = ([1, s], [1, u], [1, v], [1, w], [1, x], [1, y], [1, z])$ be an associated functional sequence for $S \cup \{\phi\}$. It is easy to see that in each of Cases B, C, D from Section 4.2, the matrix $W_{c,\Phi}$ always has rank 7 if s, u, v, w, x, y, z are distinct. For example, for Case B, $W_{c,\Phi}$ has form (writing $*$ instead of the linear functionals):

$$W_{c,\Phi} = \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \left[\begin{array}{c|cccc} * & & & * & \\ * & & & & * \\ \hline & * & * & & \\ & * & & * & \\ 0 & * & & & * \\ & & * & * & \\ & & * & & * \end{array}\right] \end{array}$$

The bottom five rows have rank 5 because every 6 functionals of Φ form a $(q^2, q, 4)$-perfect hash family. Thus the whole matrix has rank 7 as any two functionals in Φ are linearly independent. A similar argument holds for cases C and D. Hence the associated 7-functional sequence Φ is $(2, 5)$-suitable. □

We can now state a more general bound for the existence of an optimal linear $(q^2, q, 5)$-perfect hash family. By Lemma 4.2, a $(2, 5)$-suitable sequence exists if $q > 95$. Hence by Theorem 4.1, an optimal linear $(q^2, q, 5)$-perfect hash family exists if $q > 12606$. However, in the discussion in Section 4.2, we showed that we could reduce the number of excluded functionals to 5387. Hence we have that if $q + 1 > 5387$, then there exists an optimal linear $(q^2, q, 5)$-perfect hash family.

Theorem 4.3 *If $q > 5386$, then there exists an optimal linear $(q^2, q, 5)$-perfect hash family.*

4.4 Strategy for constructing perfect hash families for $d = 2$, $t = 5$

We can now use these calculations to construct $(q^2, q, 5)$-perfect hash families. As discussed, we only use functionals of the form $[1, \alpha]$, $\alpha \in \mathrm{GF}(q) \cup \{\infty\}$. We need to start with a $(2, 5)$-suitable 7-functional sequence Φ. To build Φ, we use the technique described in the first paragraph of the proof of Lemma 4.2. We begin with S'', an optimal linear $(q^2, q, 4)$-perfect hash family (constructions for these are given in [3]). Note that S'' consists of 6 linear functionals. For each 5-functional subsequence of S'' we calculate the 15 excluded linear functionals given in Table 2. We then choose any non-excluded functional from V^* to add to S'' to give us a $(2, 5)$-suitable sequence Φ. $(2, 5)$-suitable sequences were investigated in [4] and examples are found for all primes $q \geq 31$. For example, $\Phi = ([1, \infty], [1, 0], [1, 1], [1, 2], [1, 3], [1, 5], [1, 12])$ is a $(2, 5)$-suitable sequence for all primes $q > 79$.

Given a $(2, 5)$-suitable 7-functional sequence Φ, we want to calculate the excluded linear functionals for each representative $(2, 5)$-limited 7-pair sequence. In this case, there are 4200 representative $(2, 5)$-limited 7-pair sequences, which is too many to calculate by hand as we did in the $d = 2, t = 4$ case. However, using a computer, we can easily find the excluded functionals for *all* 7! of the $(2, 5)$-limited 7-pair sequences. We choose a $(2, 5)$-suitable 7-functional sequence, for example, let $\Phi = ([1, \infty], [1, 0], [1, 1], [1, 2], [1, 3], [1, 5], [1, 12])$ (for q a prime, $q > 79$). There are 7! ways of assigning s, u, v, w, x, y, z to be one of $\infty, 0, 1, 2, 3, 5, 12$. These different permutations correspond to choosing different 7-pair sequences c in the matrix $W_{c,\Phi}$. For each of these permutations, we calculate the nine excluded linear functionals $B_{(4,5)}, \ldots, D_{(2,4)}$ given in Section 4.2. This method will generate all the excluded linear functionals, and we can then choose a non-excluded linear functional (distinct from the original seven) from V^* to add to Φ to form a perfect hash family.

For simplicity we looked for constructions in the case q a prime, $q < 50,000$. However, the excluded functionals in Section 4.2 can be used to construct $(q^2, q, 5)$-perfect hash families for any (large enough) value of q. We found many constructions of optimal linear $(q^2, q, 5)$-perfect

hash families. In summary, using the $(2, 5)$-suitable sequence

$$\Phi = \{[0, 1], [1, 0], [1, 1], [1, 2], [1, 3], [1, 5], [1, 12]\},$$

we found constructions of $(q^2, q, 5)$-perfect hash families for primes $q = 359$, 397, 401, 433, 439, 443, 449, 461, and for all primes $467 \leq q \leq 50,000$, $q \neq 541$. The larger q is, the more $(q^2, q, 5)$-perfect hash families containing Φ there are. We conjecture that optimal linear $(q^2, q, 5)$-perfect hash families containing Φ will exist for all primes $q \geq 467$, $q \neq 541$.

The smallest example we found for $d = 2$ and $t = 5$ is a $(311^2, 311, 5)$-perfect hash family $S = \{[0, 1],\ [1, 0],\ [1, 1],\ [1, 2],\ [1, 3],\ [1, 5],\ [1, 16], [1, 87]\}$. In fact, starting with the $(q^2, q, 4)$-perfect hash family $S'' = \{[0, 1], [1, 0], [1, 1], [1, 2], [1, 3], [1, 5]\}$, we found optimal linear $(q^2, q, 5)$-perfect hash families for all primes $347 \leq q \leq 50,000$. Our search was by no means complete, as we only used a few suitable sequences to base our search on. It seems likely that using a different starting suitable sequence may lead to examples of optimal linear $(q^2, q, 5)$-perfect hash families for $q < 311$. For the interested reader, Table 4 lists constructions for the smallest values of q for which we found examples.

Table 4: Small values of q for which optimal linear $(q^2, q, 5)$-perfect hash families $S = \{[1, s],\ [1, u],\ [1, v],\ [1, w],\ [1, x],\ [1, y],\ [1, z], [1, i]\}$ exist.

q	s, u, v, w, x, y, z	i
311	$\infty, 0, 1, 2, 3, 5, 16$	$87, 264$
317	$\infty, 0, 1, 2, 3, 5, 19$	179
331	$\infty, 0, 1, 2, 3, 5, 14$	$75, 295$
337	$\infty, 0, 1, 2, 3, 5, 20$	268
347	$\infty, 0, 1, 2, 3, 5, 20$	227
359	$\infty, 0, 1, 2, 3, 5, 12$	329
367	$\infty, 0, 1, 2, 3, 5, 26$	334

We also wanted to find a general construction of an optimal linear $(q^2, q, 5)$-perfect hash family. The set $S = \{[0, 1], [1, 0], [1, 1], [1, 2], [1, 3], [1, 5], [1, 12], [1, 44]\}$ is an optimal linear $(q^2, q, 5)$-perfect hash family for primes q, $2400 \leq q \leq 50,000$, except for $q = 2543$, 2579 and 2843. (In these cases $\{[0, 1], [1, 0], [1, 1], [1, 2], [1, 3], [1, 5], [1, 12], [1, 46]\}$ is a perfect hash family.) This set S is also a perfect hash family for many smaller values of q. We conjecture that S is a perfect hash family for all primes $q > 2843$.

5 Case $d = 3,\ t = 3$

An optimal linear $(q^3, q, 3)$-perfect hash family has size $d(t-1) = 6$. We start by considering the representative $(3,3)$-limited 5-pair sequences from $T^2 = \{(1,2),(1,3),(2,3)\}$. To be $(3,3)$-limited, each pair on T occurs at most twice in the sequence. Using the action of the symmetric group S_3 on 1,2,3 we may assume that the sequence contains the pairs $\{(1,2),(1,2),(1,3),(1,3),(2,3)\}$ (that is, the missing pair is $(2,3)$). As the pairs $(1,2)$ and $(1,3)$ each occur exactly twice, there are $(5!/(2!2!))$ $(3,3)$-limited sequences. However, the subgroup of size 2 swapping 2 and 3 fixes the missing pair $(2,3)$. Hence there are $5!/(2!2! \times 2) = 15$ sequences to consider.

To build a $(q^3, q, 3)$-perfect hash family, we need to start with a $(3,3)$-suitable 5-functional sequence Φ. To be $(3,3)$-suitable, we need any $d(t-2) = 3$ functionals in Φ to form a $(q^3, q, 2)$-perfect hash family. So by Lemma 2.6, we need every three functionals in Φ to be linearly independent. This means we can use a special form for Φ to make our calculations easier. We use the fact that without loss of generality, any five elements, no three dependent, in $\mathrm{GF}(q)^3$ can be written as $[1, \alpha, \alpha^2]$ for five distinct values of $\alpha \in \mathrm{GF}(q) \cup \{\infty\}$ (see [12], [4]). Hence without loss of generality we may consider the following 5-functional sequence Φ and the general $(3,3)$-limited representative 5-pair sequence c:

$$\begin{array}{rcccccc} \Phi & = & ([1,v,v^2], & [1,w,w^2], & [1,x,x^2], & [1,y,y^2], & [1,z,z^2]) \\ c & = & ((2,3), & (1,2), & (1,3), & (1,3), & (1,2)). \end{array}$$

So we have

$$W_{c,\Phi} = \left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 1 & v & v^2 & -1 & -v & -v^2 \\ 1 & w & w^2 & -1 & -w & -w^2 & 0 & 0 & 0 \\ 1 & x & x^2 & 0 & 0 & 0 & -1 & -x & -x^2 \\ 1 & y & y^2 & 0 & 0 & 0 & -1 & -y & -y^2 \\ 1 & z & z^2 & -1 & -z & -z^2 & 0 & 0 & 0 \end{array}\right].$$

Our choice of functionals ensures that any $d = 3$ elements in Φ are linearly independent provided v, w, x, y, z are distinct. The remaining condition for Φ to be $(3,3)$-suitable, is that rank $W_{c,\Phi} = 5$ for all 3-limited 5-pair sequences c. Clearly the matrix $W_{c,\Phi}$ has rank 5 for any distinct v, w, x, y, z. This is true for any permutation of c, so rank $W_{c,\Phi} = 5$ for any $(3,3)$-limited 5-pair sequence c. Thus Φ is a $(3,3)$-suitable functional sequence for any distinct v, w, x, y, z.

By inspection, adding the pair $(1,2)$ to c results in the subspace of sub-excluded functionals $\langle [1,w,w^2], [1,z,z^2] \rangle$. Similarly, adding the pair $(1,3)$ to c results in the subspace of sub-excluded functionals $\langle [1,x,x^2], [1,y,y^2] \rangle$.

There is a unique way to complete c to a $(3,3)$-limited 6-pair sequence, namely by adding the pair $(2,3)$. Thus we need to perform row operations on $W_{c,\Phi}$ to find the limit-excluded linear

functional corresponding to this last pair $(2,3)$. Row reductions give us two rows in the right form, namely $(0,0,0,1,v,v^2,-1,-v,-v^2)$ and $(0,0,0,a,b,c,-a,-b,-c)$ where $a = z+w-x-y$, $b = wz - xy$, $c = zw(x+y) - xy(z+w)$. Thus any linear functional which is a linear combination of these is excluded. A general linear functional has form $[\ell, m, n]$, so the limit-excluded functionals are all functionals $[\ell, m, n]$ satisfying

$$\begin{vmatrix} \ell & m & n \\ 1 & v & v^2 \\ a & b & c \end{vmatrix} = 0. \tag{1}$$

We now consider the special case where the final linear functional $[\ell, m, n]$ is of the form $[1, \theta, \theta^2]$. Assume that all parameters v, w, x, y, z, θ are distinct. This will ensure that any three functionals in our perfect hash family are linearly independent. Note that this also means that the sub-excluded functionals are automatically prevented from occurring. Solving equation (1) gives us the limit-excluded functionals $[1, \theta, \theta^2]$ where

$$\theta = -\frac{xy(v-w) + wz(x-v) + yz(w-x)}{z(v-w) + y(x-v) + v(w-x)}.$$

That is, $w_{(2,3),[1,\theta,\theta^2]}$ is in the row space of $W_{c,\Phi}$ if θ has this form.

We note that this is exactly the same expression for θ as in the $t = 2$, $d = 4$ case. This correspondence was noted in [4] where this case was studied using geometric methods, however the geometric interpretation did not explain the correspondence, whereas the sequence technique does. More specifically, the correspondence between the sequences is:

$d = 3, t = 3$	Φ	=	$([1,v,v^2],$	$[1,w,w^2],$	$[1,x,x^2],$	$[1,y,y^2],$	$[1,z,z^2])$
	c	=	$((2,3),$	$(1,2),$	$(1,3),$	$(1,3),$	$(1,2))$
$d = 2, t = 4$	Φ	=	$([1,v],$	$[1,w],$	$[1,x],$	$[1,y],$	$[1,z])$
	c	=	$((1,2),$	$(1,3),$	$(1,4),$	$(2,3),$	$(2,4))$

So in this special case we can use the construction results from the $d = 2$, $t = 4$ case. This gives us constructions of $(q^3, q, 3)$-perfect hash families for all prime powers q except $q = 2, 3, 5, 7, 8, 9, 13$. Using the general form $[\ell, m, n]$ in equation (1) allows us to construct a $(q^3, q, 3)$-perfect hash family for the additional case $q = 13$. Constructions of $(q^3, q, 3)$-perfect hash families for prime powers $q \geq 11$ are given in [4].

6 Case $d = 4,\ t = 3$

An optimal linear $(q^4, q, 3)$-perfect hash family has $d(t-1) = 8$ elements. Once again we use the algorithm outlined in Section 2.2. We have $T^2 = \{(1,2), (1,3), (2,3)\}$, and we first count the

number of representative $(4,3)$-limited 7-pair sequences on T. Since each pair from T^2 occurs at most 3 times in a $(4,3)$-limited sequence, we have 9 pairs to choose from. We consider the different possibilities for the two pairs that are not in the sequence. Under the symmetric group S_3 acting on $\{1,2,3\}$ there are 2 possibilities for these remaining two pairs.

Case A. The two pairs not in c are $(2,3),(2,3)$. The number of $(4,3)$-limited representative 7-pair sequences c containing the pairs $\{(1,2),(1,2),(1,2),(1,3),(1,3),(1,3),(2,3)\}$ is: $7!/(3!3!1!\times 2) = 70$.

Case B. The two pairs not in c are $(1,3),(2,3)$. The number of $(4,3)$-limited representative 7-pair sequences c containing the pairs $\{(1,2),(1,2),(1,2),(1,3),(1,3),(2,3),(2,3)\}$ is: $7!/(3!2!2!\times 2) = 105$.

In both cases, extending c to an 8-pair sequence will lead to a subspace of excluded functionals. First consider the sub-excluded functionals. Suppose $\Phi = (\phi_1,\ldots,\phi_7)$ is a $(4,3)$-suitable 7-functional sequence. Let $c = ((1,2),(1,2),(1,2),(1,3),(1,3),(1,3),(2,3))$ be a sequence of type A. If we add the pair $(1,2)$ to c then we get an 8-pair sequence that is not $(4,3)$-suitable. The resulting sub-excluded functionals are the functionals in $\langle\phi_1,\phi_2,\phi_3\rangle$. Similarly adding the pair $(1,3)$ gives sub-excluded functionals $\langle\phi_4,\phi_5,\phi_6\rangle$. Permuting the pairs in c will result in sub-excluded functionals being all the functionals that are dependent on 3 functionals of Φ. We get a similar statement if a sequence c of type B is extended to an 8-pair sequence that is not $(4,3)$-limited.

In order to simplify our search for constructions, we will restrict our attention to functionals of a special form. This will make it easier to find constructions, but will mean that we do not completely answer the existence question. We will consider the special case using only linear functionals that have form $[1,\alpha,\alpha^2,\alpha^3]$ for $\alpha\in \mathrm{GF}(q)$. Note that this means that provided we use distinct values of α for our functionals, we automatically have any 4 functionals linearly independent, and so the sub-excluded functionals are automatically excluded.

Let $\Phi = ([1,s,s^2,s^3],[1,u,u^2,u^3],[1,v,v^2,v^3],[1,w,w^2,w^3],[1,x,x^2,x^3],[1,y,y^2,y^3],[1,z,z^2,z^3])$ for distinct s,u,v,w,x,y,z. We check that Φ is $(4,3)$-suitable. By Lemma 2.6, every 4 elements of Φ form a $(q^4,q,2)$-perfect hash family since any 4 elements of Φ are independent. To check the rank condition, consider the $(4,3)$-limited 7-pair sequence

$$c = ((1,2),(1,2),(1,2),(1,3),(1,3),(1,3),(2,3)).$$

The resulting matrix $W_{c,\Phi}$ has rank 7 as s,u,v,w,x,y,z are distinct. This is true for any permutation of c. Similarly, the rank condition holds for sequences of type B. Hence Φ is $(4,3)$-suitable.

We now use this c and Φ and perform row reductions on $W_{c,\Phi}$ to find the excluded lin-

ear functionals. Completing the sequence c with the pair $(1,2)$ or $(1,3)$ repeats the condition that every four linear functionals are independent. Completing the sequence c with the pair $(2,3)$ gives us the limit-excluded linear functionals $(0,0,0,0,1,z,z^2,z^3,-1,-z,-z^2,-z^3)$, $(\underline{0},\phi_1,-\phi_1)$, $(\underline{0},\phi_2,-\phi_2)$, for certain ϕ_1,ϕ_2. (The functionals ϕ_1,ϕ_2 are easy to calculate, but have a complex form, so we omit the detailed description here.) Any linear combination of these 3 functionals is limit-excluded. However, we are considering the special case where we only want to add a functional to Φ that has form $[1,\alpha,\alpha^2,\alpha^3]$. To find the limit-excluded functionals of this form we solve

$$\begin{vmatrix} 1 & \alpha & \alpha^2 & \alpha^3 \\ 1 & z & z^2 & z^3 \\ & \phi_1 & & \\ & \phi_2 & & \end{vmatrix} = 0.$$

This equation has 3 solutions for $\alpha \in \mathrm{GF}(q)$, one solution being $\alpha = z$ which corresponds to a functional already in Φ . Thus there are at most two further excluded functionals of the form $[1,\alpha,\alpha^2,\alpha^3]$.

We undertake a similar process for a sequence c of type B. Here the two pairs missing from c are different, so we will get a different subspace of excluded functionals for each missing pair. Each of these subspaces excludes three linear functionals of the form $[1,\alpha,\alpha^2,\alpha^3]$. Two of these are functionals are already in Φ. Hence for each of the two pairs missing from c, we have one further excluded functional of the form $[1,\alpha,\alpha^2,\alpha^3]$.

In total there are $70 + 105 = 175$ representative $(4,3)$-limited 7-pair sequences, each with 2 excluded linear functionals of the form $[1,\alpha,\alpha^2,\alpha^3]$. So there are at most $175 \times 2 = 350$ functionals excluded in this way. Since there are q linear functionals of the form $[1,\alpha,\alpha^2,\alpha^3]$, $\alpha \in \mathrm{GF}(q)$, then provided $q > 7+350 = 357$, an optimal $(q^4,q,3)$-perfect hash family will exist. We have proved:

Theorem 6.1 *Optimal linear $(q^4,q,3)$-perfect hash families exist for all prime power $q > 357$.*

In order to construct optimal linear $(q^4,q,3)$-perfect hash families, we begin with a $(4,3)$-suitable 7-functional sequence Φ; by the above argument, any choice of distinct s,u,v,w,x,y,z for Φ will do. We chose $0,1,2,3,4,5,6$, that is,

$$\Phi = ([1,0,0,0],[1,1,1,1],[1,2,2^2,2^3],[1,3,3^2,3^3],[1,4,4^2,4^3],[1,5,5^2,5^3],[1,6,6^2,6^3]).$$

For each of the 7! arrangements of $\{s,u,v,w,x,y,z\} = \{0,1,2,3,4,5,6\}$, we calculate the 2 solutions for α obtained from Case A, and the 2 solutions for α from Case B. Each arrangement excludes (at most) 4 values of α. Once these are calculated, any value $\beta \in \mathrm{GF}(q)\backslash\{0,1,2,3,4,5,6\}$

not excluded will give us a linear functional $[1, \beta, \beta^2, \beta^3]$ which we can add to Φ to make a perfect hash family. We did some computer searches and found that optimal linear $(q^4, q, 3)$-perfect hash families exist for q much smaller than the bound of 357. Our program searched for such perfect hash families for each prime $q < 2000$. Table 5 lists an example of a perfect hash family containing $S = \{[1,0,0,0], [1,1,1,1],\ [1,2,2^2,2^3],\ [1,3,3^2,3^3],\ [1,4,4^2,4^3],\ [1,5,5^2,5^3],\ [1,6,6^2,6^3]\}$ for each prime $q < 2000$ where they exist.

Table 5: Examples of $(q^4, q, 3)$-perfect hash families $S \cup \{[1, \beta, \beta^2, \beta^3]\}$ for prime $q < 2000$

q	β
47	24
59	21
67	34
79	37
83	42
89	43
97,109	26
101,137	11
103,107,151,163,173,179,191,193,197,199,211,227,241,251,263,269	12
113,127,131,223,229,233,239,271,277,281,349,521,541,599	13
139,157,257	17
149	19
$283 \leq q \leq 2000,\ q \neq 349, 521, 541, 599$	12

Table 5 shows that the set $\{[1,0,0,0], [1,1,1,1], [1,2,2^2,2^3], [1,3,3^2,3^3], [1,4,4^2,4^3], [1,5,5^2,5^3], [1,6,6^2,6^3],\ [1,12,12^2,12^3]\}$ is an optimal linear $(q^4, q, 3)$-perfect hash family for all prime q, $599 < q < 2000$. We conjecture that this set will be a perfect hash family for all primes $q > 599$.

7 Conclusion

In this paper we have developed a useful technique to construct optimal linear (q^d, q, t)-perfect hash families. In Method 1 in Section 2.2 we demonstrate how to check whether a given set of functionals is a linear (q^d, q, t)-perfect hash family. In Method 2 in Section 2.2, we give an algorithm that can be used recursively to construct optimal linear (q^d, q, t)-perfect hash families for any parameters q, d, t where they exist. We have illustrated this algorithm by constructing new optimal linear $(q^2, q, 5)$- and $(q^4, q, 3)$-perfect hash families. This algorithm is straightforward to use for small d and t. It can be used for larger d, t, but the size of d and t will be limited by computational capacity.

References

[1] M. Atici, S. S. Magliveras, D. R. Stinson and W.-D. Wei. Some recursive constructions for perfect hash families. *J. Combin. Designs*. 4 (1996) 353-363.

[2] M. Atici, D. R. Stinson and W.-D. Wei. A new practical algorithm for the construction of a perfect hash function. *J. Combin. Math. Combin. Comput.* 35 (2000), 127-145.

[3] S. G. Barwick, W.-A. Jackson and C. T. Quinn. Optimal linear perfect hash families with small parameters. *J. Combin. Designs*, 12 (2004) 311-324.

[4] S. G. Barwick and W.-A. Jackson. Geometrical constructions of optimal linear perfect hash families. Preprint. Available at Cryptology ePrint Archive: Report 2006/002.

[5] A. Beutelspacher and U. Rosenbaum. *Projective Geometry: From Foundations to Applications*. Cambridge University Press, 1998.

[6] S. R. Blackburn. Combinatorics and threshold cryptography. Combinatorial designs and their applications, *CRC Research Notes in Mathematics*, 403 (1999) 49-70.

[7] S. R. Blackburn. Perfect hash families: probabilistic methods and explicit constructions. *J. Combin. Theory Ser A.* 92 (2000) 54-60.

[8] S. R. Blackburn, M. Burmester, Y. Desmedt and P. R. Wild. Efficient multiplicative sharing schemes. Advances in Cryptology – EUROCRYPT '96. *Lecture Notes in Computer Science* 1070 (1996) 107-118.

[9] S. R. Blackburn and P. R. Wild. Optimal linear perfect hash families, *J. Combin. Theory Ser. A*, 83 (1998) 233–250.

[10] L.R.A. Casse. *A Guide to Projective Geometry.* To appear, Oxford University Press.

[11] A. Fiat and M. Naor. Broadcast Encryption. Advances in Cryptology – CRYPTO '93. *Lecture Notes in Computer Science* 773 (1994) 480-491.

[12] J. W. P. Hirschfeld. *Projective Geometry over Finite Fields*. Oxford Mathematical Monographs, Oxford, UK, 1998.

[13] D. R. Hughes and F. C. Piper. *Projective planes.* Springer-Verlag, Berlin-Heidelberg-New York, 1973.

[14] K. Mehlhorn. *Data Structures and Algorithms 1: Sorting and Searching.* Springer Verlag, Berlin, 1984.

[15] D. R. Stinson, T. van Trung and R. Wei. Secure frameproof codes, key distribution patterns, group testing algorithms and related structures. *J. Statist. Plan. Infer.*, 86 (2000) 595–617.

[16] H. Wang and C. Xing. Explicit constructions of perfect hash families from algebraic curves over finite fields. *J. Combin. Theory* Ser A. 93 (2001) 112-124.

Appendix: Notation Index

This appendix contains an index of the notation and definitions introduced in Section 2.1.

$V = \mathrm{GF}(q)^d$

$V^* = \{\phi : V \to \mathrm{GF}(q)\}$

$P = (p_1, \ldots, p_t) \in V^t$

$U_P = \left\{(\psi_1, \ldots, \psi_t) \in (V^*)^t : p_1^{\psi_1} + \cdots + p_t^{\psi_t} = 0\right\}$

$T = \{1, \ldots, t\}$

$T^2 = \{(a, b) \in T \times T : 1 \le a < b \le t\}$

$w_{(a,b),\phi} = (\psi_1, \ldots, \psi_t) \in (V^*)^t$ where

$$\psi_i = \begin{cases} \phi & \text{if } i = a, \\ -\phi & \text{if } i = b, \\ 0 & \text{if } i \neq a, b \end{cases} \qquad (\text{for } (a,b) \in T^2,\ \phi \in V^*)$$

$V_{(a,b)} = \{w_{(a,b),\phi} : \phi \in V^*\} \subseteq (V^*)^t, \quad \text{for } (a,b) \in T^2$

$W_{c,\Phi} = \left\langle w_{(a_i,b_i),\phi_i} : 1 \le i \le k \right\rangle \subseteq (V^*)^t,$
where $c = ((a_1, b_1), \ldots, (a_k, b_k)) \in (T^2)^k, \Phi = (\phi_1, \ldots, \phi_k) \in (V^*)^k$

rank $W_{c,\Phi}$ means the rank of the matrix $W_{c,\Phi}$ with k rows $w_{(a_i,b_i),\phi_i}$, $i = 1, \ldots, k$..

$R = \{(p, \ldots, p) : p \in V\} \subseteq V^t$

$Y = \{(\psi_1, \ldots, \psi_t) \in (V^*)^t : \psi_1 + \cdots + \psi_t = 0\}$

$\dim L$ refers to the dimension of the subspace L over $\mathrm{GF}(q)$

A *k-pair sequence* $c = ((a_1, b_1), \ldots, (a_k, b_k))$ *on* T is a sequence of k pairs from T^2.

A *k-functional sequence* $\Phi = (\phi_1, \ldots, \phi_k) \in (V^*)^k$ is a sequence of k functionals from V^*.

c_I, Φ_I: Let c be a k-pair sequence, Φ a k-functional sequence and $I \subseteq T$. Then c_I is the subsequence of c consisting of the pairs in c that lie in I^2. Φ_I is the subsequence of Φ consisting

of the functional corresponding to the pairs in c_I.

A (d,t)-*limited* k-pair sequence is a k-pair sequence on T where for each proper subset I of T we have $|c_I| < d(|I|-1)$.

A (d,t)-*suitable* sequence Φ is a $(d(t-1)-1)$-functional sequence such that (i) every $d(t-2)$ elements of Φ form a $(q^d, q, t-1)$-perfect hash family, and (ii) for every (d,t)-limited $(d(t-1)-1)$-pair sequence c, we have rank $W_{c,\Phi} = d(t-1)-1$.

The *representative* k-pair sequences are the collection of one sequence from each isomorphism group under the action of the symmetry group S_t.