Fourteenth International Workshop on Algebraic and Combinatorial Coding Theory
September 7–13, 2014, Svetlogorsk (Kaliningrad region), Russia *pp. 24–31*

Conjectural upper bounds on the smallest size of a complete arc in $PG(2,q)$ based on an analysis of step-by-step greedy algorithms

Daniele Bartoli daniele.bartoli@dmi.unipg.it
Dipartimento di Matematica e Informatica, Università degli Studi di Perugia, Via Vanvitelli 1, Perugia, 06123, Italy
Alexander A. Davydov adav@iitp.ru
Institute for Information Transmission Problems (Kharkevich Institute), Russian Academy of Sciences, Bol'shoi Karetnyi per. 19, GSP-4, Moscow, 127994, Russia
Giorgio Faina faina@dmi.unipg.it
Dipartimento di Matematica e Informatica, Università degli Studi di Perugia, Via Vanvitelli 1, Perugia, 06123, Italy
Alexey A. Kreshchuk krsch@iitp.ru
Institute for Information Transmission Problems (Kharkevich Institute), Russian Academy of Sciences, Bol'shoi Karetnyi per. 19, GSP-4, Moscow, 127994, Russia
Stefano Marcugini, Fernanda Pambianco
{gino,fernanda}@dmi.unipg.it
Dipartimento di Matematica e Informatica, Università degli Studi di Perugia, Via Vanvitelli 1, Perugia, 06123, Italy

Abstract. In the projective plane $PG(2,q)$, an iterative construction of complete arcs by adding a new point on every step is considered. The working mechanism of a step-by-step greedy algorithm constructing small complete arcs in $PG(2,q)$ is explained. It is proven that uncovered points are evenly placed on the plane. For more than half of the steps of the iterative process, an estimation of the number of new covered points is proven. A natural conjecture that the estimation holds for the rest of steps is done. From that, upper bounds on the smallest size $t_2(2,q)$ of a complete arc in $PG(2,q)$ are obtained. In particular,

$$t_2(2,q) < \sqrt{q}\sqrt{3\ln q + \ln\ln q + \ln 3} + \sqrt{\frac{q}{3\ln q}} + 3,$$

$$t_2(2,q) < 1.885\sqrt{q\ln q}.$$

A connection with the Birthday problem is noted. The effectiveness of the new bounds is illustrated by comparison with the smallest known sizes of complete arcs obtained in the recent works of the authors via computer search for a huge region of q.

1 Introduction

Let $PG(2,q)$ be the projective plane over the Galois field F_q. An n-arc is a set of n points no three of which are collinear. An n-arc is called complete if it is not contained in an $(n+1)$-arc of $PG(2,q)$.

In [8] the relationship among the theory of n-arcs, coding theory, and mathematical statistics is presented. In particular, a complete arc in a plane $PG(2,q)$, the points of which are treated as 3-dimensional q-ary columns, defines a parity check matrix of a q-ary linear code with codimension 3, Hamming distance 4, and covering radius 2. Arcs can be interpreted as linear maximum distance separable (MDS) codes [8] and they are related to optimal coverings arrays, to superregular matrices, and to quantum codes.

One of the most important problems in the study of projective planes, which is also of interest in coding theory, is the determination of the *smallest size* $t_2(2,q)$ of a complete arc in $PG(2,q)$. This is a hard open problem. Surveys and results on the sizes of plane complete arcs can be found in [1, 2, 4–6, 8, 9]; see also the references therein. The exact values of $t_2(2,q)$ are known only for $q \le 32$ [6].

This work is devoted to *upper bounds* on $t_2(2,q)$.

Let $t(\mathcal{P}_q)$ be the size of the smallest complete arc in any (not necessarily Galois) projective plane $\mathcal{P}_q$ of order q. In [9], for *sufficiently large* q, the following result is proven by *probabilistic methods*:

$$t(\mathcal{P}_q) \le D\sqrt{q}\ln^C q,\ C \le 300,$$

where C and D are constants independent of q. The authors of [9] conjecture that the constant can be reduced to $C = 10$.

Denote by $\bar{t}_2(2,q)$ the *smallest **known** size* of a complete arc in $PG(2,q)$. In [2], the values of $\bar{t}_2(2,q)$ (up to November 2013) are collected for $q \in T$ where

$$T = \{q : 173 \le q \le 49727,\ q \text{ power prime}\} \cup \{q : 173 \le q \le 125003,\ q \text{ prime}\}$$
$$\cup\, \{59 \text{ sporadic prime } q\text{'s in the interval } [125101 \dots 360007]\}.$$

In [3], the values of $\bar{t}_2(2,q)$ are obtained for the following set $T^{\#}$:

$T^{\#} = \{q : 125017 \le q \le 150001,\ q \text{ prime}\} \cup \{290011, 370003, 380041, 390001, 400009\}$. From results of [2–5], we have

$$t_2(2,q) < \sqrt{q}\ln^{0.7295} q \quad \text{for } 109 \le q \le 169 \text{ and } q \in T \cup T^{\#}. \tag{1}$$

In this work, an iterative construction of complete arcs by adding a new point on every step is considered. The working mechanism of a step-by-step greedy algorithm constructing small complete arcs in $PG(2,q)$ is explained. It is proven that uncovered points are evenly placed on the plane. For more than half of the steps of the iterative process, an estimation of the number of new covered points is proven. A natural conjecture that the estimation holds for the rest of steps is done, see Conjecture 7. From that, new upper bounds on $t_2(2,q)$ better than the known ones are obtained, see Theorems 1, 8, and 9.

Also, *in this work*, we propose a form of the upper bound on $t_2(2,q)$ connected with an upper estimate of a *decreasing function* $m(q)$ given as follows:

$$t_2(2,q) = m(q)\sqrt{q\ln q}.$$

Theorem 1. **(i)** *Under Conjecture 7, in $PG(2,q)$ the following hold:*

$$t_2(2,q) < \widehat{t}_2(2,q) = \sqrt{q}\sqrt{3\ln q + \ln\ln q + \ln 3} + \sqrt{\frac{q}{3\ln q}} + 3; \tag{2}$$

$$t_2(2,q) < 1.885\sqrt{q\ln q}; \tag{3}$$

$$m(q) < \sqrt{3 + \frac{\ln\ln q + \ln 3}{\ln q}} + \frac{1}{\sqrt{3}\ln q} + \frac{3}{\sqrt{q\ln q}}. \tag{4}$$

(ii) *The sizes $\overline{t}_2(2,q)$ of the smallest known complete arcs in $PG(2,q)$, $q \in T \cup T^{\#}$, obtained in* [2,3]*, satisfy the upper bounds* (2)*–*(4)*.*

Conjecture 2. *In $PG(2,q)$, the upper bounds* (2), (3) *hold for all q without any extra conditions and conjectures.*

This work can be treated as a development of the paper [1].

2 Upper bounds on $t_2(2,q)$

Assume that in $PG(2,q)$, a complete arc is constructed by a step-by-step algorithm (*Algorithm* for short) adding one new point to the arc on every step. For example, a greedy algorithm on every step adds to the arc a point providing the maximum possible (for the given step) number of new covered points [2,4,5].

A *point* of $PG(2,q)$ is *covered by* an *arc* if the point lies on any of its bisecants.

We consider the $(w+1)$-th step of Algorithm. This step starts from a w-arc. Assume that after the w-th step of Algorithm we obtain a w-arc not covering exactly U_w points. Let $\mathbf{S}_w(U_w)$ be the set of all w-arcs in $PG(2,q)$ each of which does not cover exactly U_w points. As a *starting arc* for the $(w+1)$*-th step* we *randomly choose* a w-arc $\mathcal{K}_w$ of $\mathbf{S}_w(U_w)$ so that for every arc of $\mathbf{S}_w(U_w)$ the probability to be chosen is equal to $\frac{1}{\#\mathbf{S}_w(U_w)}$. So, we consider $\mathbf{S}_w(U_w)$ as an *ensemble of random objects* with the uniform probability distribution.

Lemma 3. *Every point of $PG(2,q)$ may be considered as a random object that can be uncovered by a randomly chosen w-arc $\mathcal{K}_w$ with some probability p_w. The probability p_w is the same for all points; it is equal to*

$$p_w = \frac{U_w}{\#PG(2,q)} = \frac{U_w}{q^2+q+1}. \tag{5}$$

Let $s_w(v)$ be the number of "1" in a sequence of v random and independent "1/0" experiments, each of which yields "1" with probability p_w. We have the binomial probability distribution for the random value $s_w(v)$; the *expected value* of $s_w(v)$ is

$$\mathbf{E}[s_w(v)] = vp_w = v\frac{U_w}{q^2+q+1}. \tag{6}$$

We are interested in the value $s_w(v)$ as it is the number of uncovered points among v random points of $PG(2,q)$, if the events to be uncovered are considered as independent.

Let the arc $\mathcal{K}_w$ consist of w points $A_1, A_2, \ldots, A_w$. Let A_{w+1} be the point that will be included into the arc on the $(w+1)$-th step. The point A_{w+1} defines a bundle of w tangents $\overline{A_1 A_{w+1}}, \ldots, \overline{A_w A_{w+1}}$ to $\mathcal{K}_w$. Excluding $A_1, \ldots, A_w$, all the points on the tangents are candidates to be new covered points at the $(w+1)$-th step. There are $w(q-1)+1$ candidates in the bundle. Let $\Delta_w(A_{w+1})$ be the number of new covered points on the $(w+1)$-th step. For an arc $\mathcal{K}_w$, denote by $\Delta_w^{\text{aver}}(\mathcal{K}_w)$ the average value of $\Delta_w(A_{w+1})$ by all U_w uncovered points A_{w+1}, i.e.

$$\Delta_w^{\text{aver}}(\mathcal{K}_w) = \frac{\sum\limits_{A_{w+1}} \Delta_w(A_{w+1})}{U_w} \geq 1. \tag{7}$$

In future, we consider continuous approximations of the discrete function $\Delta_w(A_{w+1})$ and other ones keeping the same notations.

Lemma 4. *It holds that*

$$\Delta_w^{\text{aver}}(\mathcal{K}_w) \geq \Delta_w^{\text{low}} = \max\{1, \frac{wU_w}{q+2-w} - w + 1\} \tag{8}$$

where equality $\Delta_w^{aver}(\mathcal{K}_w) = \frac{wU_w}{q+2-w} - w + 1$ *holds if and only if every tangent contains the same number of uncovered points.*

Lemma 5. (i) *Let* $U_w \geq (q+2)(q+2-w)/(w-1)$. *Then*

$$\Delta_w^{\text{aver}}(\mathcal{K}_w) \geq \frac{wU_w}{q+2-w} - w + 1 \geq \mathbf{E}[s_w(w(q-1)+1)] = \frac{(w(q-1)+1)U_w}{q^2+q+1}.$$

(ii) *Let* $U_w < (q+3)/w$. *Then* $\Delta_w^{aver}(\mathcal{K}_w) \geq \mathbf{E}[s_w(w(q-1)+1)]$.

Corollary 6. *Let* $(q+2)(q+2-w)/(w-1) \leq U_w < (q+3)/w$. *Then for any arc* $\mathcal{K}_w$ *of* $\mathbf{S}_w(U_w)$, *there exists an uncovered point* A_{w+1} *providing the following inequality:*

$$\Delta_w(A_{w+1}) \geq \mathbf{E}[s_w(w(q-1)+1)] = \frac{(w(q-1)+1)U_w}{q^2+q+1}. \tag{9}$$

In the region $(q+2)(q+2-w)/(w-1) \leq U_w < (q+3)/w$ we have a *rigorous proof without any conjecture*. This region takes approximately $\frac{1}{\sqrt{3}} \approx 58\%$ of all the steps of the iterative process.

Conjecture 7. *Let*

$$\frac{q+3}{w} < U_w < \frac{(q+2)(q+2-w)}{w}. \tag{10}$$

Then in $PG(2,q)$, *for* q *sufficiently large, there exists a* w*-arc* $\mathcal{K}_w \subset \mathbf{S}_w(U_w)$ *such that one can find an uncovered point* A_{w+1} *providing inequality* (9).

Let $q^* = q + 3$. We denote

$$f_q(w) = \prod_{i=1}^{w} \left(1 - \frac{i}{q^*}\right).$$

The function $f_q(w)$ is used e.g. in the study of the *Birthday problem* [7].

Theorem 8. (**The basic inequality**). *Under Conjecture* 7*, in* $PG(2, q)$ *it holds that*

$$t_2(2, q) \le w + 1 + \xi,$$

where $\xi \ge 1$ *and* w *satisfies "the basic" inequality*

$$f_q(w) \le \frac{\xi}{q^2 + q}.$$

Theorem 9. *Under Conjecture* 7*, in* $PG(2, q)$ *it holds that*

$$t_2(2, q) \le \sqrt{2(q + 3)}\sqrt{\ln \frac{q^2 + q}{\xi}} + 2 + \xi, \quad \xi \ge 1.$$

Remark 10. As all uncovered points lie on tangents to the arc $\mathcal{K}_w$, it can be shown that for every point, the probability to be uncovered, cf. (5), is, in fact,

$$\widetilde{p}_w = \frac{U_w}{(q + 1)(q + 2 - w)}, \tag{11}$$

whence the corresponding expectation, cf. (6) and (9), is

$$\widetilde{\mathbf{E}}[s_w(w(q - 1) + 1)] = \frac{(w(q - 1) + 1)U_w}{(q + 1)(q + 2 - w)}. \tag{12}$$

3 An illustration of the effectiveness of the new upper bounds

We compare the new bounds of Theorem 1 with the smallest known sizes $\overline{t}_2(2, q)$ of complete arcs in $PG(2, q)$ obtained in the recent works [2, 3] of the authors via computer search for $q \in T \cup T^{\#}$.

Figures 1 and 2 **illustrate the reasonableness of Conjecture 7, and the effectiveness of the new upper bounds** even for small q.

Figure 3 explains the working mechanism of a step-by-step greedy algorithm constructing small complete arcs in $PG(2, q)$. Here $\mathbf{E} = \mathbf{E}[s_w(w(q - 1) + 1)]$,

$$b_w^{\max} = \frac{\max\limits_{A_{w+1}} \Delta_w(A_{w+1})}{\mathbf{E}}, \; b_w^{\text{aver}} = \frac{\Delta_w^{\text{aver}}(\mathcal{K}_w)}{\mathbf{E}}, \; b_w^{\text{low}} = \frac{\Delta_w^{low}}{\mathbf{E}}, \; \frac{\widetilde{p}_w}{p_w} \sim \frac{q}{q - w}.$$

The curves $b_w^{\max}$, b_w^{aver} are experimental, while b_w^{low} and $\frac{\widetilde{p}_w}{p_w}$ are calculated by (5), (8), (11). The final region of the process when $U_w < \frac{q+3}{w}$, is not fully shown.

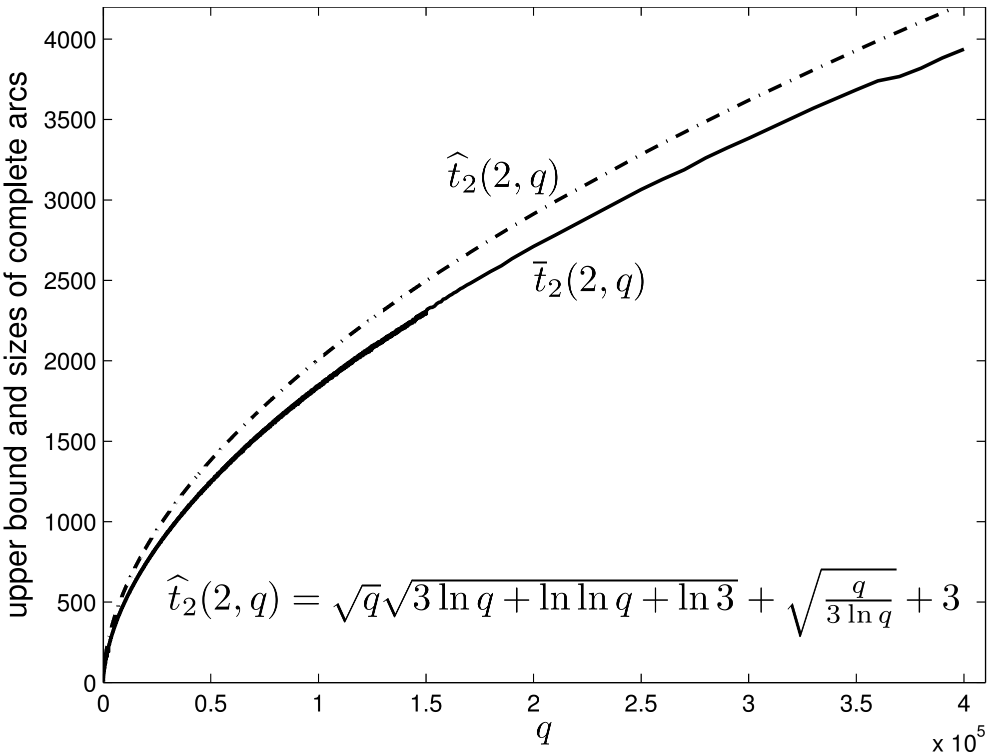

Figure 1: **Upper bound on $t_2(2,q)$ vs greedy algorithms' results.** $\widehat{t}_2(2,q)$ (the top dashed-dotted curve); sizes $\overline{t}_2(2,q)$ of complete arcs obtained by the greedy algorithms, $q \in T \cup T^{\#}$ (the bottom solid curve).

It is interesting (and expected) that, for almost all the steps of the iterative process, the curve b_w^{aver} and the line $\frac{\widetilde{p}_w}{p_w}$ practically coincide with each other. It means, that in fact, we have $\Delta_w^{\text{aver}}(\mathcal{K}_w) \approx \widetilde{\mathbf{E}}[s_w(w(q-1)+1)]$, see (7), (12).

References

[1] D. Bartoli, A. Davydov, G. Faina, A. Kreshchuk, S. Marcugini, F. Pambianco, Two Types of Upper Bounds on the Smallest Size of a Complete Arc in the Plane $PG(2,q)$, in *Proc. VII Int. Workshop on Optimal Codes and Related Topics, OC2013, Albena, Bulgaria, 2013*, Sofia, 2013, 19–25. http://www.moi.math.bas.bg/oc2013/a3.pdf

[2] D. Bartoli, A. A. Davydov, G. Faina, A.A. Kreshchuk, S. Marcugini, and F. Pambianco, Tables of Sizes of Small Complete Arcs in

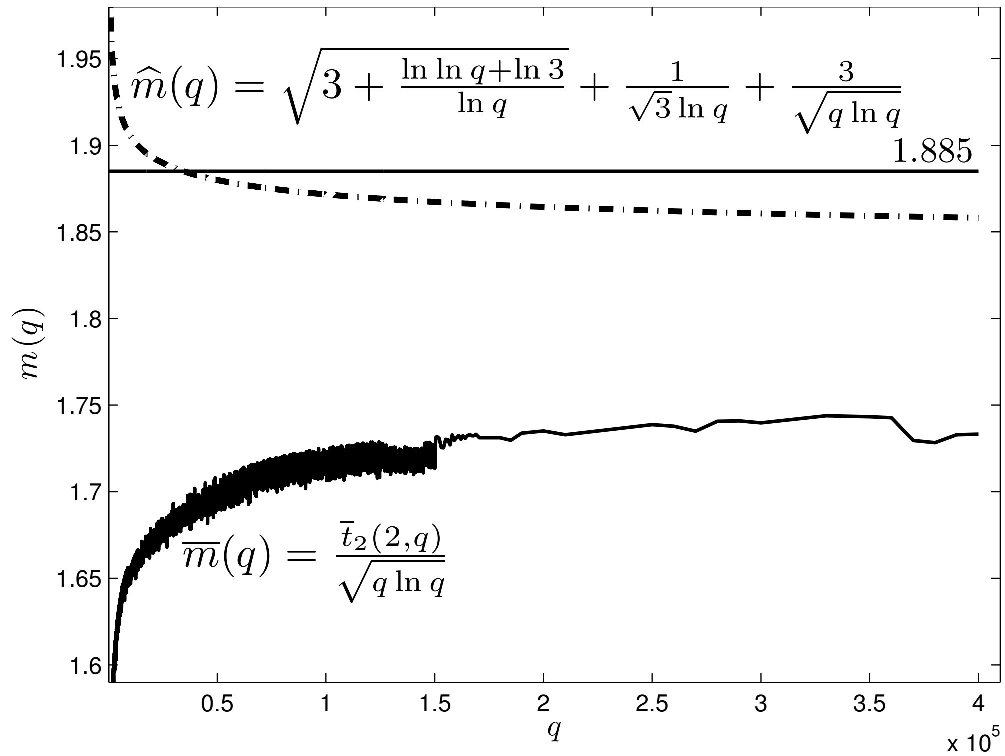

Figure 2: **A new type of the upper bound based on functions** $m(q)$**.** $\widehat{m}(q)$ (the top dashed-dotted curve); $\overline{m}(q) = \frac{\overline{t}_2(2,q)}{\sqrt{q \ln q}}$ from greedy algorithms' complete arcs, $q \in T \cup T^{\#}$ (the bottom solid curve); bound (3) (the line $y = 1.885$).

the Plane $PG(2,q)$, $q \leq 360007$, 2013, arXiv:1312.2155 [math.CO], http://arxiv.org/abs/1312.2155

[3] D. Bartoli, A. Davydov, G. Faina, A. A. Kreshchuk, S. Marcugini, F. Pambianco, I. A. Tkachenko, Upper Bounds on the Smallest Sizes of a Complete Arc in $PG(2,q)$ Based on Computer Search, in *Proc. XIV Int. Workshop on Algebraic and Combin. Coding Theory, ACCT2014, Svetlogorsk , Russia*, 2014.

[4] D. Bartoli, A. A. Davydov, G. Faina, S. Marcugini, and F. Pambianco, On Sizes of Complete Arcs in $PG(2,q)$, *Discrete Math.*, **312**, 680–698, 2012.

[5] D. Bartoli, A. A. Davydov, G. Faina, S. Marcugini, and F. Pambianco, New Upper Bounds on the Smallest Size of a Complete Arc in a Finite Desarguesian Projective Plane, *J. Geom.*, **104**, 11–43, 2013.

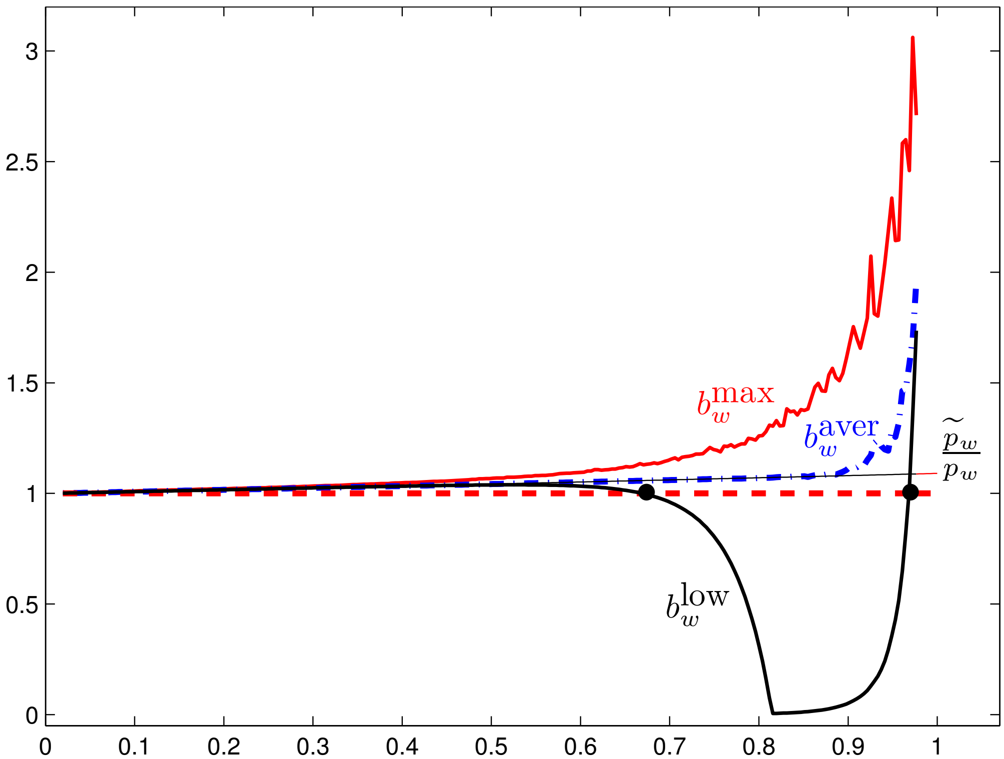

Figure 3: **The working mechanism of a step-by-step greedy algorithm constructing small complete k-arc in $PG(2,q)$, $k = 255, q = 3023$.** $b_w^{\max}$ (the top solid curve); b_w^{aver} (the second dashed-dotted curve); $\frac{\widetilde{p}_w}{p_w}$ (the solid line close to b_w^{aver}); b_w^{low} (the bottom solid curve); the dashed line is $y = 1$; the horizontal axis shows values of $\frac{w}{k}$.

[6] D. Bartoli, G. Faina, S. Marcugini, and F. Pambianco, On the Minimum Size of Complete Arcs and Minimal Saturating Sets in Projective Planes, *J. Geom.*, **104**, 409–419, 2013.

[7] D. Brink, A (Probably) Exact Solution to the Birthday Problem, *The Ramanujan J.*, **28**, 223–238, 2012.

[8] J. W. P. Hirschfeld and L. Storme, The Packing Problem in Statistics, Coding Theory and Finite Projective Spaces, *J. Statist. Planning Infer.*, **72**, 355–380, 1998.

[9] J. H. Kim and V. Vu, Small Complete Arcs in Projective Planes, *Combinatorica,* **23**, 311–363, 2003.