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* Alexander A. Davydov@ Giorgio Faina*
Alexey A. Kreshchuk@ Stefano Marcugini*
Fernanda Pambianco*

@ Institute for Information Transmission Problems (Kharkevich Institute), Russian Academy of Science, Moscow, Russia
* Department of Mathematics and Informatics, Perugia University, Perugia, Italy

XIV International Workshop on Algebraic and Combinatorial Coding Theory, ACCT-2014, Svetlogorsk, Russia, September 7-13, 2014

Outline

INTRODUCTION NOTATION

$PG(2, q)$ $\Leftrightarrow$ projective **space** of dimension 2 over Galois field F_q

***n*-arc** $\Leftrightarrow$ a set of n points no three of which are collinear
a line meeting an arc $\Leftrightarrow$ **tangent** or **bisecant**

bisecant $\Leftrightarrow$ a line intersecting an arc in two points

a **point** A of $PG(2, q)$ is **covered** by an arc $\Leftrightarrow$
A lies on a **bisecant** of the arc

complete arc $\Leftrightarrow$ all points of $PG(2, q)$ are covered
by bisecants of the arc
$\Leftrightarrow$ one may not add a new point to a complete arc

INTRODUCTION NOTATION

$PG(2, q)$ ⇔ projective **space** of dimension 2 over Galois field F_q

n-**arc** ⇔a set of n points no three of which are collinear
a line meeting an arc ⇔ **tangent** or **bisecant**

bisecant ⇔ a line intersecting an arc in two points

a **point** A of $PG(2, q)$ is **covered** by an arc ⇔
A lies on a **bisecant** of the arc

complete arc ⇔ all points of $PG(2, q)$ are covered
by bisecants of the arc
⇔ one may not add a new point to a complete arc

INTRODUCTION NOTATION

$t_2(2, q)$ $\Leftrightarrow$ the smallest size of a complete arc in $PG(2, q)$

HARD OPEN CLASSICAL PROBLEM: **1950** $\rightarrow$

upper bound on $t_2(2, q)$

exact values of $t_2(2, q)$ are only for $q \leq 32$

$q = 31, 32$ D. Bartoli, G. Faina, S. Marcugini, F. Pambianco, *Journal of Geometry* **2013**

lower bounds $t_2(2, q) > \sqrt{2q} + 1, \quad \forall q$

$t_2(2, q) > \sqrt{3q} + \frac{1}{2}, \ q = p^h, \ h \leq 3$ (p^3 O.Polverino **1999**)

INTRODUCTION NOTATION

$t_2(2,q) \Leftrightarrow$ the smallest size of a complete arc in $PG(2,q)$

HARD OPEN CLASSICAL PROBLEM: **1950** $\rightarrow$

upper bound on $t_2(2,q)$

exact values of $t_2(2,q)$ are only for $q \leq 32$

$q = 31, 32$ D. Bartoli, G. Faina, S. Marcugini, F. Pambianco, *Journal of Geometry* **2013**

lower bounds $t_2(2,q) > \sqrt{2q} + 1, \quad \forall\, q$

$t_2(2,q) > \sqrt{3q} + \frac{1}{2}, \ q = p^h, \ h \leq 3$ (p^3 O.Polverino **1999**)

INTRODUCTION NOTATION

$t_2(2,q) \Leftrightarrow$ the smallest size of a complete arc in $PG(2,q)$

HARD OPEN CLASSICAL PROBLEM: 1950 $\rightarrow$

upper bound on $t_2(2,q)$

exact values of $t_2(2,q)$ are only for $q \leq 32$

$q = 31, 32$ D. Bartoli, G. Faina, S. Marcugini, F. Pambianco, *Journal of Geometry* **2013**

lower bounds $t_2(2,q) > \sqrt{2q} + 1, \quad \forall\, q$

$t_2(2,q) > \sqrt{3q} + \frac{1}{2}, \; q = p^h, \; h \leq 3$ (p^3 O.Polverino **1999**)

KNOWN UPPER BOUNDS on $t_2(2, q)$

theoretical $t_2(2,q) \leq d\sqrt{q}\log^c q, \quad c \leq 300$

$c, d \Leftrightarrow$ constants independent of q

probabilistic methods J.H. Kim, V. Vu *Combinatorica* **2003**

computer search $t_2(2,q) < \sqrt{q}\ln^{0.7295} q$

HUGE region $t_2(2,q) < 1.745\sqrt{q\ln q}$

ALL $q \leq 150001$ WITHOUT GAPS

41 sporadic q's in $[150503 \ldots 430009]$

D.Bartoli, A.A.Davydov, G.Faina, A.A.Kreshchuk, S.Marcugini, F.Pambianco

J. Geometry, Discrete Mathematics, OC2013, ACCT2014, arXiv **2005-2014**

KNOWN UPPER BOUNDS on $t_2(2, q)$

theoretical $t_2(2,q) \le d\sqrt{q}\log^c q, \quad c \le 300$

$c,\ d \Leftrightarrow$ constants independent of q

probabilistic methods J.H. Kim, V. Vu *Combinatorica* **2003**

computer search $t_2(2,q) < \sqrt{q}\ln^{0.7295} q$

HUGE region $t_2(2,q) < 1.745\sqrt{q\ln q}$

ALL $q \le 150001$ WITHOUT GAPS

41 sporadic q's in $[150503 \ldots 430009]$

D.Bartoli, A.A.Davydov, G.Faina, A.A.Kreshchuk, S.Marcugini, F.Pambianco

J. Geometry, Discrete Mathematics, OC2013, ACCT2014, arXiv **2005-2014**

AIM and APPROACH

AIM $\Rightarrow$ analytical bound

$$t_2(2, q) < c\sqrt{q \ln q}. \qquad c - \text{small constant}$$

WAY $\Rightarrow$ analysis of step-by-step greedy algorithms

A greedy algorithm is an algorithm that makes the *locally optimal choice* at each stage with the hope of finding a global optimum or, at least, a global "good" solution.

"From the first day to this, sheer greed was the driving spirit of civilization" (F. Engels)

AIM and APPROACH

AIM $\Rightarrow$ analytical bound

$$t_2(2, q) < c\sqrt{q \ln q}. \qquad c - \text{small constant}$$

WAY $\Rightarrow$ analysis of step-by-step greedy algorithms

A greedy algorithm is an algorithm that makes the *locally optimal choice* at each stage with the hope of finding a global optimum or, at least, a global "good" solution.

"From the first day to this, sheer greed was the driving spirit of civilization" (F. Engels)

Ensemble of random w-arcs

The w-th step of Algorithm forms a w-arc W.

U_w $\Leftrightarrow$ the number of points not covered by W

$\mathbf{S}_w(U_w) \Leftrightarrow$ the set of all w-arcs in $PG(2,q)$ each of which does not cover exactly U_w points.

Starting arc of the $(w+1)$-*th step* $\Leftrightarrow$ w-arc $\mathcal{K}_w$ randomly chosen of $\mathbf{S}_w(U_w)$.

For every arc of $\mathbf{S}_w(U_w)$ the probability to be chosen $= \frac{1}{\#\mathbf{S}_w(U_w)}$.

$\mathbf{S}_w(U_w)$ $\Leftrightarrow$ an ensemble of random objects with the uniform probability distribution.

Uniform distribution of uncovered points

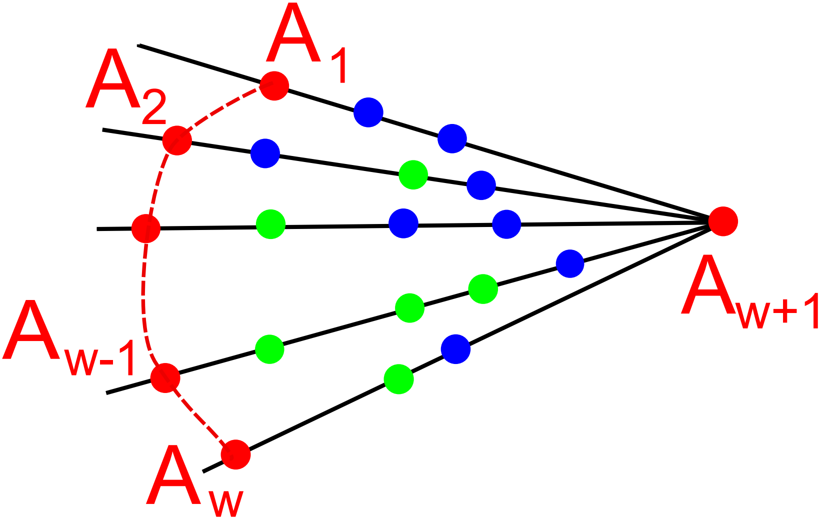
Lemma

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:

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

the **proportion** of uncovered points =
the **probability** that a point is uncovered

One step of a greedy algorithm



the number of new covered points on the $(w+1)$-th step

A_i – point of $PG(2,q)$. $\#PG(2,q) = q^2 + q + 1$

arc $\mathcal{K}_w = \{A_1, A_2, \ldots, A_w\}$

point A_{w+1} will be included in the arc on the $(w+1)$-th step

A_{w+1} defines a bundle $\mathcal{B}_w(A_{w+1})$ of w tangents to $\mathcal{K}_w$

$w(q-1)+1$ points of $\mathcal{B}_w(A_{w+1}) \setminus \{A_1, \ldots, A_w\}$ are candidates to be new covered points at the $(w+1)$-th step

$\Delta_w(A_{w+1})$ – the number of new covered points on $(w+1)$-th step

U_w uncovered points $\Rightarrow$ U_w distinct bundles

the main idea for bounds

if events to be uncovered are independent, the expected value of the number new covered points among $w(q-1)+1$ random points =

$$\mathbf{E} = p_w \cdot (w(q-1)+1) = \frac{U_w}{q^2+q+1}(w(q-1)+1)$$

MAIN IDEA $\Rightarrow$ there exists an uncovered point A_{w+1} providing

$$\boxed{\Delta_w(A_{w+1}) \geq \mathbf{E}}$$

RIGOROUS PROOF $\approx 58\%$ of all the steps of the process

CONJECTURE rest of the steps

Theorem and Conjecture

Theorem

$$\text{Let} \qquad U_w > \frac{(q+2)(q+2-w)}{w} \quad \text{or} \quad \frac{q+3}{w} > U_w$$

Then for any arc $\mathcal{K}_w$ *of* $\mathbf{S}_w(U_w)$, *there exists an uncovered point* A_{w+1} *providing the inequality* $\Delta_w(A_{w+1}) \geq \mathbf{E}$

Conjecture

$$\text{Let} \qquad \frac{(q+2)(q+2-w)}{w} > U_w > \frac{q+3}{w}$$

Then for any arc $\mathcal{K}_w$ *of* $\mathbf{S}_w(U_w)$, *there exists an uncovered point* A_{w+1} *providing the inequality* $\Delta_w(A_{w+1}) \geq \mathbf{E}$

Theorem and Conjecture

Theorem

$$\text{Let} \qquad U_w > \frac{(q+2)(q+2-w)}{w} \quad \text{or} \quad \frac{q+3}{w} > U_w$$

Then for any arc $\mathcal{K}_w$ of $\mathbf{S}_w(U_w)$, there exists an uncovered point A_{w+1} providing the inequality $\Delta_w(A_{w+1}) \geq \mathbf{E}$

Conjecture

$$\text{Let} \qquad \frac{(q+2)(q+2-w)}{w} > U_w > \frac{q+3}{w}$$

Then for any arc $\mathcal{K}_w$ of $\mathbf{S}_w(U_w)$, there exists an uncovered point A_{w+1} providing the inequality $\Delta_w(A_{w+1}) \geq \mathbf{E}$

tools for rigorous proof

the average value of $\Delta_w(A_{w+1})$ by all U_w uncovered points A_{w+1}

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

there exists $\Delta_w(A_{w+1}) \geq \Delta_w^{aver}(\mathcal{K}_w)$

Lemma

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

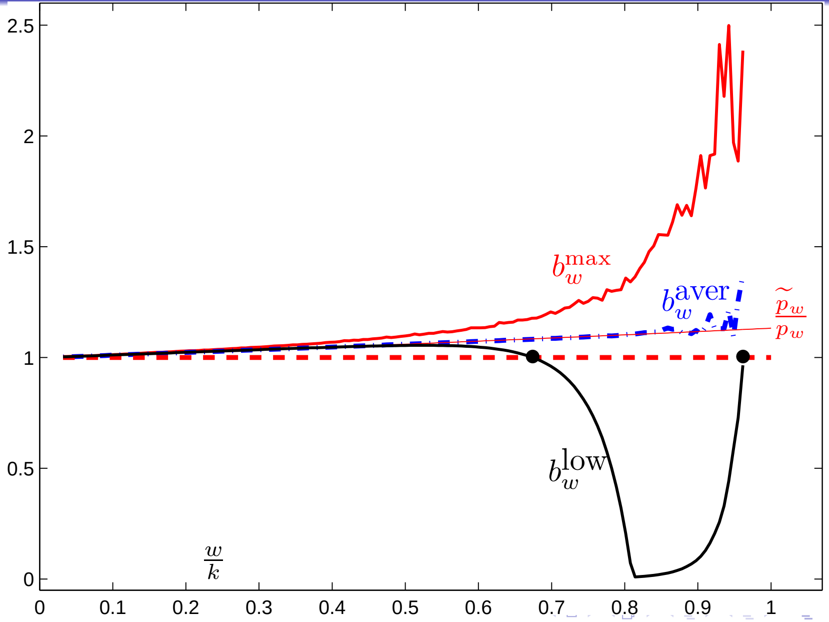
where equality holds if and only if every tangent contains the same number of uncovered points.

Conjecture is reasonable

1. The rigorous proof uses the equality $\Delta_w^{aver}(\mathcal{K}_w) = \Delta_w^{low}$. Formally, we have no right to take inequality $\Delta_w^{aver}(\mathcal{K}_w) > \Delta_w^{low}$. The equality is sufficient only for the 1-st 58% steps of Algorithm. The 2-nd part formally needs in conjecture. But, in this 2-nd part: the numbers of uncovered points on tangents are essentially distinct & there is many random factors affecting the process $\Rightarrow$ the variance of the random value $\Delta_w(A_{w+1})$ increases $\Rightarrow$ $\Delta_w(A_{w+1}) > \mathbf{E}$ exists

2. In fact: the probability that a point is uncovered $> p_w$ $\Rightarrow$ the expected value of the number new covered points among $w(q-1)+1$ random points $> \mathbf{E}$

RIGOROUS proof vs CONJECTURE



new upper bounds (under Conjecture)

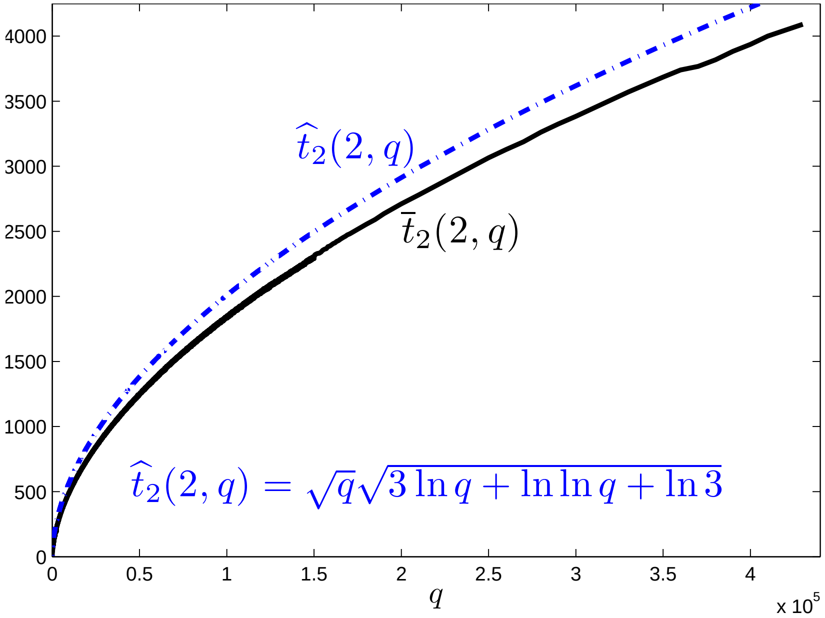
Theorem

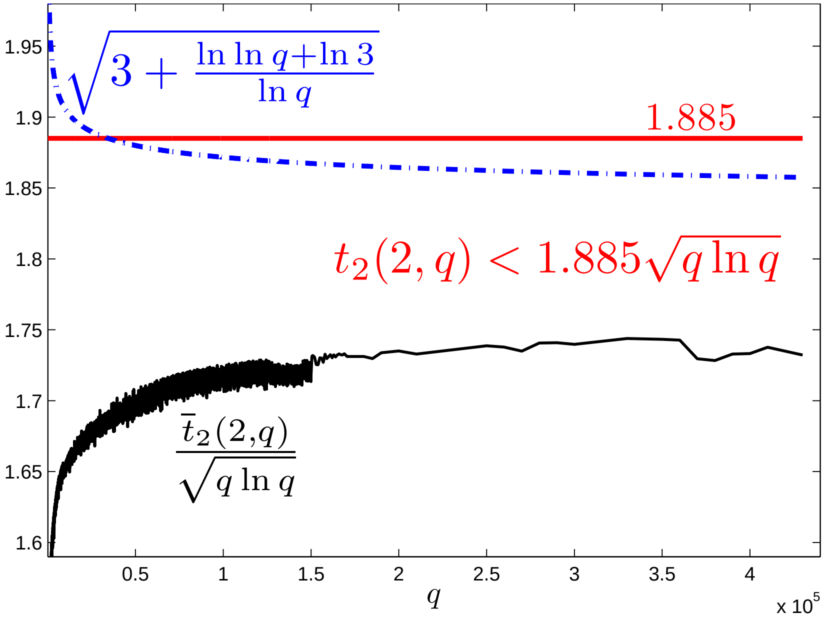
*Let a constant $\xi \geq 1$. **Under Conjecture,***

$$t_2(2, q) \leq w + 1 + \xi, \quad w \text{ satisfies } \prod_{i=1}^{w}\left(1 - \frac{i}{q+3}\right) \leq \frac{\xi}{q^2+q}.$$

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

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


4000
3500
3000
2500
2000
1500
1000
500
0
0
0.5
1
1.5
2
2.5
3
3.5
4
q
x 10^5
$\widehat{t}_2(2,q)$
$\overline{t}_2(2,q)$
$\widehat{t}_2(2,q) = \sqrt{q}\sqrt{3\ln q + \ln\ln q + \ln 3}$


$\sqrt{3+\frac{\ln\ln q+\ln 3}{\ln q}}$
1.885
$t_2(2,q) < 1.885\sqrt{q\ln q}$
$\frac{\overline{t}_2(2,q)}{\sqrt{q\ln q}}$
1.95
1.9
1.85
1.8
1.75
1.7
1.65
1.6
0.5
1
1.5
2
2.5
3
3.5
4
q
x 10^5

Thank you Spasibo

Mille grazie

Premnogo blagodarya

!'Muchas gracias

Toda raba

Merci beaucoup

Dankeschön

Dank u wel

Domo arigato