An Approximation Algorithm
for the Number of Zeros of
Arbitrary Polynomials over GF[q]

Dima Grigoriev ¹
Marek Karpinski ²

TR-91-027
April, 1991

Abstract

We design the first polynomial time (for an arbitrary and fixed field GF[q])
(ε, δ)-approximation algorithm for the number of zeros of arbitrary polynomial
f(x₁, . . . , xₙ) over GF[q]. It gives the first efficient method for estimating the
number of zeros and nonzeros of multivariate polynomials over small finite
fields other than GF[q]² (like GF[q³]), the case important for various circuit
approximation techniques. The algorithm is based on the estimation of the
number of zeros of an arbitrary polynomial f(x₁, . . . , xₙ) over GF[q] in the
function on the number m of its terms. The bounding ratio number is proved
to be m(εq−1)log₂ which is the main technical contribution of this paper and
could be of independent algebraic interest.

¹Max-Planck-Institute of Mathematics, 5300 Bonn 1
²Dept. of Computer Science, University of Bonn, 5300 Bonn 1, and the International Computer
Science Institute, Berkeley, California. Supported in part by Leibniz Center for Research in
Computer Science, by the DFG, Grant KA 673/4-1 and by the SERC Grant GR-E 68297
1 Introduction

Recently there has been a progress in design of efficient approximation algorithms for algebraic counting problems. The first polynomial time $(\epsilon, \delta)$-approximation algorithm for the number of zeros of a polynomial $f(x_1, \ldots, x_n)$ over the field $GF[2]$ has been designed by Karpinski and Luby ([KL 91a]) and this result was extended to arbitrary multilinear polynomials over $GF[q]$ by Karpinski and Lhotzky ([KL 91b]).

In this paper we construct the first $(\epsilon, \delta)$-approximation algorithm for the number of zeros of an arbitrary polynomial $f(x_1, \ldots, x_n)$ with $m$ terms over an arbitrary (but fixed) finite field $GF[q]$ working in polynomial time in the size of the input, the ratio $m^{(q-1)\log q}$, and $\frac{1}{\epsilon}, \log(\frac{1}{\delta})$. (The corresponding $(\epsilon, \delta)$-approximation algorithm for the number of nonzeros of a polynomial can be constructed to work in time polynomial in the size of the input, the ratio $m^{\log q}$, and $\frac{1}{\epsilon}, \log(\frac{1}{\delta})$.)

2 Approximation Algorithm

We refer to Karp, Luby and Madras [KLM 89] as well as [KL 91a], [KL 91b] for the more detailed discussion of the abstract structure of the Monte-Carlo method for estimating cardinalities of finite sets and the related techniques.

Given $f \in GF[q][x_1, \ldots, x_n], f = \sum_{i=1}^{m} t_i$, and $c \in GF[q]$. Denote

$$\#_c f = \left| \{(x_1, \ldots, x_n) \in GF[q]^n \mid f(x_1, \ldots, x_n) = c \} \right| .$$

Our $(\epsilon, \delta)$-approximation algorithm will have the following overall structure:

MONTE CARLO APPROXIMATION ALGORITHM
Input $f \in GF[q][x_1, \ldots, x_n], c \in GF[q], \epsilon > 0, \delta > 0, (f \neq 0)$

Output $\hat{Y}$ (such that $\Pr[(1 - \epsilon)\#c f \leq \hat{Y} \leq (1 + \epsilon)\#c f] \geq 1 - \delta$)

1. Construct a universe set $U$ (the size $|U|$ of $U$ must be efficiently computable.)

2. Choose randomly with the uniform probability distribution $N$ members $u_i$ from $U$, $u_i \in U$, $i = 1, 2, \ldots, N$.

3. Construct now from a polynomial $f$ an indicator function $\tilde{f} : U \rightarrow \{0, 1\}$ such that $|\tilde{f}^{-1}(1)| = \#c f$.

4. Compute the number $N = \frac{1}{\beta} \frac{4\log(2/\delta)}{\epsilon^2}$ for $\beta \geq |U|/\#c f$.

5. Compute for all $i$, $1 \leq i \leq N$, the values $\tilde{f}(u_i)$ and set $Y_i \leftarrow |U|\tilde{f}(u_i)$.

6. Compute $\hat{Y} \leftarrow \frac{1}{N} \sum_{i=1}^{N} Y_i$.

7. OUTPUT: $\hat{Y}$.

Correctness of the above algorithm is guaranteed by the following Theorem.

Theorem 1 (Zero-One Estimator Theorem [KLM 89])

Let $\mu = \frac{\#c f}{|U|}$. Let $\epsilon \leq 2$. If $N \geq \frac{1}{\mu} \frac{4\log(2/\delta)}{\epsilon^2}$, then the above Monte Carlo Algorithm is an $(\epsilon, \delta)$-approximation algorithm for $\#c f$.

We shall distinguish two (technically different) cases:

Case 1. Polynomial $f(x_1, \ldots, x_n)$ over $GF[q]$ is constant free and $c = 0$.

Case 2. Polynomial $f(x_1, \ldots, x_n)$ over $GF[q]$ is arbitrary and $c \neq 0$.  

3
Let us denote \( \tilde{f} = (f - c)^{q-1} - 1 = \sum \tilde{t}_i \).

The corresponding universes and indicator functions will be \( U_1 = GF[q]^n, \tilde{f}_1(s) = 1 \) if and only if \( f(s) = 1 \), and \( U_2 = \{(s,i) \mid \tilde{t}_i(s) \neq 0\} \), \( \tilde{f}_2(s,i) = 1 \) if and only if \( f(s) = c \) and for no \( j < i \), \( (s,j) \in U_2 \).

Let us observe that \( \frac{\#_{c,j}}{\#_{c,j}} \leq m^{q-1} \cdot \frac{|\tilde{G}_{(f-c)^{q-1}}|}{\#_{c,j}} \) for \( \tilde{G}_{(f-c)^{q-1}} = \{(s,i) \mid \tilde{t}_i(s) \neq 0\} \), there is no \( j < i \) such that \( \tilde{t}_j(s) \neq 0 \), see figure 1. (Observe that \( |\tilde{G}_{(f-c)^{q-1}}| = \{|s \mid \text{there is a term } \tilde{t}_i \text{ of } (f - c)^{q-1} - 1 \text{ such that } \tilde{t}_i(s) \neq 0\}| \).

The corresponding bounds \( \beta_i \geq \frac{\#_{c,j}}{\#_{c,j}} \) will be proven to satisfy

\[
\beta_1 \leq (m + 1)(q-1)\log q \\
\beta_2 \leq m^{q-1}(m + 1)(q-1)\log q.
\]

![Figure 1](image)

The rest of the paper will be devoted to the proofs of these two bounds.
We shall denote the corresponding algorithms by $A_1$ and $A_2$.

Let us analyze the bit complexity of both algorithms (for the corresponding subroutines see [KL 91a], [KL 91b], and [KLM 89]).

Denote by $P(q)$ the bit costs of multiplication and powering over $GF[q]$, $P(q) = O(\log^2 q \log \log q \log \log \log q)$ (cf. [We 87]). The evaluation of the polynomial takes time $O(nmP(q))$ and the overall complexity of the algorithm $A_1$ is

$$O(nm(m + 1)^{(g-1)\log q} \log(1/\delta) / \epsilon^2)$$

and of the algorithm $A_2$

$$O(nm(m + 1)^{(g-1)(1+\log q)} \log q P(q) \log(1/\delta) / \epsilon^2).$$

For the fixed finite field $GF[q]$ the running time of both algorithms is bounded by a polynomial of the degree depending on the order of the ground field. The bounds for $\beta_1$ and $\beta_2$ which are proven polynomial in $m$ only, are the main technical contribution of this paper.

Please note that the condition whether $f = 0$ is satisfiable can be checked deterministically for arbitrary polynomial $f \in GF[q][x_1, \ldots, x_n]$ within the bounds stated above because of the following (for a problem of a black-box interpolation of $f$, see [GKS 90]):

**Proposition 1.** Let $f \in GF[q][x_1, \ldots, x_n]$ and $c \in GF[q]$, the equation $f = c$ is satisfiable if and only if $g = (f - c)^{q-1} - 1$ has at least one nonconstant term.

**Proof.** $f = c$ is satisfiable iff $(f - c)^{q-1} = 0$ is satisfiable iff the inequality $(f - c)^{q-1} - 1 \neq 0$ is satisfiable. The inequality $(f - c)^{q-1} - 1 \neq 0$ is satisfiable iff there exists in $(f - c)^{q-1} - 1$ at least one nonconstant term. \qed
3 Main Theorem

Given an arbitrary polynomial \( f \in GF[q][X_1, \cdots, X_n] \), \( \deg x_i f \leq q - 1 \), denote \( G = G_f = \{(x_1, \cdots, x_n) \mid f(x_1, \cdots, x_n) \neq 0\} \), \( \tilde{G} = \tilde{G}_f = \{(x_1, \cdots, x_n) \mid \exists t_i \in f : t_i(x_1, \cdots, x_n) \neq 0\} \) (For notational reasons from now on in this section, variables will be written in capital (e.g. \( X_i \)) and values in small (e.g. \( x_i \)).

Denote by \( m = m_f \) the number of terms in \( f \).

By the support of a term \( t \) we mean the set of indices of variables occurring in \( t \).

**Theorem 2** \[ \frac{|\tilde{G}|}{|G|} \leq m^{\log_2 q} \]

**Remark.** This bound is sharp. Example: for \( 0 \leq k \leq n \)

\[ g_k = X_1^{q-1} \cdots X_k^{q-1}(1 - X_{k+1}^{q-1}) \cdots (1 - X_n^{q-1}) . \]

In this case \(|\tilde{G}| = (q - 1)^k q^{n-k} , |G| = (q - 1)^k , m = 2^{n-k} . \)

**Proof.** For any subset \( J \subset \{1, \cdots , n\} \) define an elementary cylinder \( C(J) = \{(x_1, \cdots , x_n) \in GF[q]^n \mid x_j \neq 0 \text{ for } j \in J \text{ and } x_i = 0 \text{ for } i \notin J \} \). Observe that for \( J_1 \neq J_2 \quad C(J_1) \cap C(J_2) = \emptyset \). Define the cone of \( J \)

\[ CON(J) = \{(x_1, \cdots , x_n) \in GF[q]^n \mid x_j \neq 0 \text{ for } j \in J \} = \bigcup_{J_1 \supseteq J} C(J_1) . \]

By \( f_J \in GF[q][\{X_j \}_{j \in J}] \) we denote the polynomial obtained from \( f \) in the following way: multiply \( f \) by the term \( X_J = \prod_{j \in J} X_j \), replace each appeared power \( X_j^2 \) by \( X_j \), make necessary cancellation, denote this intermediate result by \( f : X_J \) and finally, substitute zeroes instead of \( X_i \) for all \( i \notin J \). Remark that each for term of \( f_J \) its support coincides with \( J \), moreover \( m_J \leq m_{f : X_J} \leq m_f \).
Lemma 1. For every $J \subseteq \{1, \cdots, n\}$

a) $G \cap C(J) = G_{f_J}$ (here under equality we mean a canonical isomorphism);

b) $G \cap CON(J) = G_{f \cdot X_J}$.

Proof. Observe that for any point $(x_1, \cdots, x_n) \in C(J)$ (respectively $CON(J)$),

$$f(x_1, \cdots, x_n) \neq 0 \iff f_J(\{x_j\}_{j \in J}) \neq 0 \text{ (respectively } f_X(x_1, \cdots, x_n) \neq 0\text{)},$$

this proves lemma 1.

Lemma 2. a) $G \cap C(J) \neq \emptyset$ iff $f_J \neq 0$;

b) $G \cap CON(J) \neq \emptyset$ iff $f \cdot X_J \neq 0$;

c) if $f_J \neq 0$ then $\tilde{G} \supseteq C(J) = \tilde{G}_{f_J}$ and $\tilde{G} \supseteq CON(J) = \tilde{G}_{f \cdot X_J}$.

Proof. a) (respectively b)) follows from lemma 1a) (respectively 1b)).

c) follows from the statement that if $f_J \neq 0$ then $f$ contains a term with a support being a subset of $J$.

We call $J$ active if $f_J \neq 0$.

Lemma 3. Assume $J$ is active. Then

$$\frac{|\tilde{G}_{f_J}|}{|G_{f_J}|} = \frac{|C(J)|}{|CON(J)|} \leq m_{f_J}^{\log q - 1}(\leq m_{f_J}^{\log q}).$$

Note. This lemma states the theorem for the case of the polynomial $f_J$.

Proof. We conduct by induction on $|J|$. Remark that $|\tilde{G}_{f_J}| = |C(J)| = (q-1)|J|$. Assume that for a certain $j_0 \in J$ the polynomial $f_J$ does not divide by $(X_{j_0} - \alpha)$ for each $\alpha \in GF[q]^*$. Then $f_{J_{j_0}} = f_J(X_{j_0} = \alpha) \neq 0$. Then by lemma 2a) we can apply inductive hypothesis to each of these polynomials $f_{J_{j_0}}$. Since $|G_{f_J}| = \sum_{\alpha \in GF[q]^*} |G_{f_{J_{j_0}}}|$ and $m_{f_{J_{j_0}}} \leq m_{f_J}$, we get by induction the statement of the lemma in this case.
Assume now that $\prod_{j \in J} (X_j - \alpha_j) | f_J$ for some $\alpha_j \in GF[q]^*$, $j \in J$. We claim in this case that $m_{f_J} \geq 2^{|J|}$. By lemma 1a) this would prove lemma 3. We prove the claim by induction on $|J|$.

Fix some $j_0 \in J$ and write (uniquely) $f_J = \sum h_{J_1}(X_{j_0})M_{J_1}$ where $M_{J_1}$ are terms in the variables $\{X_j\}_{j \in J \setminus \{j_0\}}$ and $h_{J_1}(X_{j_0}) \in GF[q][X_{j_0}]$. Then $(X_{j_0} - \alpha_{j_0})|h_{J_1}(X_{j_0})$ for each $M_{J_1}$, hence $h_{J_1}(X_{j_0})$ contains at least two terms.

Take a certain $x_{j_0} \in GF[q]^*$ such that $0 \neq f_J(X_{j_0} = x_{j_0}) \in GF[q][\{X_j\}_{j \in J \setminus \{j_0\}}]$ and apply inductive hypothesis of the claim to $f_J(X_{j_0} = x_{j_0})$, taking into account that $m_{f_J} \geq 2m_{f_J(x_{j_0} = x_{j_0})}$. Lemma 3 is proved.

**Lemma 4** If $J \subseteq \{1, \cdots, n\}$ is a minimal (w.r.t. inclusion relation) support of the terms in $f$ then $J$ is active.

**Proof.** Represent (uniquely) $f = f_1 + f_2$ where $f_1$ is the sum of all terms occurring in $f$ with the support $J$. Then the polynomial $f_J = X_Jf_1 \neq 0$ has the same number of terms as $f_1$, this proves lemma 4.

**Corollary 1** $\bar{G}$ coincides with the union of the cones $CON(J)$ for all (minimal) active $J$.

Now we consider the lattice $\mathcal{L} = 2^{\{1, \cdots, n\}}$ and for $J \in \mathcal{L}$ we denote its cone $con(J) \subseteq \mathcal{L}$, $con(J) = \{J'|J \subseteq J'\}$. We'll construct a partition $\mathcal{P}$ of the union $\mathcal{G}$ of $con(J)$ for all active $J$.

Take any linear ordering $\prec$ of the active elements with the only property that if $J_1 \nsubseteq J_2$ for two active elements then $J_1 \succ J_2$ (e.g. as the first element one can take arbitrary maximal one, then a maximal in the rest set etc.).

Associate with any element $J_1 \in \mathcal{G}$ an active element $J$ minimal w.r.t. ordering $\prec$
with the property \( J \subseteq J_1 \). Then as an element of the partition \( \mathcal{P} \) which is attached to an active element \( J \) (denote it by \( \mathcal{P}(J) \)) consists of all such elements of \( \mathcal{G} \) which are associated with \( J \).

For any \( J_1 \) call a subset \( S \subseteq \text{con}(J_1) \) a relative principal ideal with the generator \( J_1 \) if for any \( J_2 \supseteq J_3 \supseteq J_1 \) and \( J_2 \in S \) we have \( J_3 \in S \).

**Lemma 5**  
\[ a) \mathcal{P} \text{ is a partition of } \mathcal{G}; \]
\[ b) \text{For each active element } J, \mathcal{P}(J) \text{ is a relative principal ideal with the generator } J \text{ (with the unique active element } J). \]

**Proof.** Part a) is clear. To prove part b) consider \( J_1 \in \mathcal{P}(J) \) and \( J_1 \supseteq J_2 \supseteq J \), then \( J_2 \in \mathcal{G} \) (since \( \mathcal{G} \) is a union of the cones). We have to prove that \( J \) corresponds to \( J_2 \).

Assume the contrary and let \( J_0 \subseteq J_2 \) for some active \( J_0 \) such that \( J_0 \prec J \), hence \( J_0 \subseteq J_1 \) and we get a contradiction with \( J_1 \in \mathcal{P}(J) \) which proves lemma 5.

**Lemma 6**  
For any active element \( J \) and each \( J_1 \in \mathcal{P}(J) \) the sum \( M_{J_1} \) of the terms occurring in \( fX_J \) with the support \( J_1 \) equals to
\[
f_J(\frac{X_{J_1}}{X_J})^{q-1}(-1)^{\lvert J_1 \setminus J \rvert}. \]

**Proof.** We prove it by induction on \( |J_1 \setminus J| \).

The base for \( J_1 = J \) is clear. Take any \( J_1 \in \mathcal{P}(J) \), then for each \( J_1 \supseteq J_2 \supseteq J \) we have \( J_2 \in \mathcal{P}(J) \) by lemma 5 and by inductive hypothesis \( M_{J_2} = f_J(\frac{X_{J_2}}{X_J})^{q-1}(-1)^{\lvert J_2 \setminus J \rvert} \).

Since \( J_1 \) is not active we have \( f_{J_1} = 0 \). Observe that \( f_{J_1} = (\sum_{J_2 \subseteq J_1, J_2 \neq J} M_{J_2}) \frac{X_{J_2}}{X_J} \).

Therefore \( f_{J_1} = \frac{X_{J_1}}{X_J}(-f_J(\frac{X_{J_2}}{X_J})^{q-1}(-1)^{\lvert J \setminus J_2 \rvert} + M_{J_2}) \) and we obtain
\[
M_{J_1} = f_J(\frac{X_{J_1}}{X_J})^{q-1}(-1)^{\lvert J \setminus J_1 \rvert}.
\]
taking into account that each term in $f_J$ has a support equal to $J$.
Induction and lemma 6 are proved.

**Corollary 2** For any active element $J$

$$m_f \geq m_{f.x} \geq m_{f.J} \cdot |P(J)|.$$

**Lemma 7** For any relative principal ideal $S \subseteq \text{con}(J)$ with the generator $J$ the weight $K$ of $S$

$$K = \sum_{s \in S} (q - 1)^{|s \setminus J|} \leq |S|^{\log_2 q}.$$

**Proof.** We prove by induction on $n - |J|$.
The base for $n = |J|$ (then $|S| = 1$) is obvious. For the inductive step take some index $i_0 \notin J$. Consider a partition of $S = S_0 \cup S_1$ where $S_1$ (respectively $S_0$) consists of all elements containing (respectively not containing) $i_0$. Then $S_0$ can be considered as a relatively principal ideal with the generator $J$ in the lattice $2^{\{1, \ldots, n\}\setminus\{i_0\}}$. By $S_1'$ denote a subset of $2^{\{1, \ldots, n\}\setminus\{i_0\}}$ obtained from $S_1$ by deleting $i_0$ from each element. Then $S_1'$ is also a relative principal ideal (may be empty) with the generator $J$ and $S_1' \subset S_0$, in particular $|S_1| \leq |S_0|$.

According to this partition represent $K = K_0 + (q - 1)K_1$ where $K_0 = \sum_{s_1 \in S_1} (q - 1)^{|s_1 \setminus J|}$, $K_1 = \sum_{s_0 \in S_0} (q - 1)^{|s_0 \setminus J|}$. By inductive hypothesis

$$K \leq |S_0|^{\log_2 q} + (q - 1)|S_1|^{\log_2 q} \leq (|S_0| + |S_1|)^{\log_2 q}.$$

the latter inequality follows from the convexity of the function $X \rightarrow X^{\log_2 q}$ (on the ray $IR_+$ of nonnegative reals), namely rewrite this inequality in the form

$$|S_0|^{\log_2 q} + (2|S_1|)^{\log_2 q} \leq |S_1|^{\log_2 q} + (|S_0| + |S_1|)^{\log_2 q}.$$
This completes the proof of the induction and lemma 7.

### Corollary 3
For any active element \( J \)

\[
|\tilde{G} \cap \bigcup_{J_1 \in \mathcal{P}(J)} C(J_1)| \leq |G \cap C(J)| \left( m_{f, X_J} \right)^{\log_2 q} \leq |G \cap C(J)| \left( m_f \right)^{\log_2 q}.
\]

**Proof.**

\[
|\tilde{G} \cap \bigcup_{J_1 \in \mathcal{P}(J)} C(J_1)| = (q - 1)^{|J|} \cdot \sum_{J_1 \in \mathcal{P}(J)} (q - 1)^{|J_1 \setminus J|}.
\]

By lemma 3 \((q - 1)^{|J|} \leq \left| \mathcal{P}(J) \right| \left( m_{f, J} \right)^{\log_2 q} \). By lemma 5b \( \mathcal{P}(J) \) is a relative principal ideal, hence \( |\tilde{G} \cap \bigcup_{J_1 \in \mathcal{P}(J)} C(J_1)| \leq |\mathcal{P}(J)|^{\log_2 q} \) by lemma 7. Therefore we get the corollary 3 applying corollary 2.

Finally, we complete the proof of the theorem summing left and right sides of the inequalities from corollary 3 ranging over all active elements \( J \), taking into account corollary 1, lemma 5a) and lemma 2a).

### 4 Bounds for \( \beta_1 \) and \( \beta_2 \)

We shall apply now Theorem 2 to derive upper bounds for \( \beta_1 \) and \( \beta_2 \).

### Theorem 3
Given any polynomial \( f \in GF[q][x_1, \ldots, x_n] \) with \( m \) terms and without constant terms. Then

\[
\frac{q^n}{\#0_f} \leq \beta_1 = (m^{q-1} + 1)^{\log q} \leq (m + 1)^{(q-1) \log q}.
\]

**Proof.** Consider the polynomial \( g = f^{q-1} \).

For \( s \in GF[q]^n, f(s) = 0 \iff (f^{q-1} - 1)(s) \neq 0 \). Apply Theorem 2
to the polynomial $f^{q-1} - 1 \in GF[q][x_1, \cdots, x_n]$, $|\bar{G}| = q^n$, $|G| = \#_c f$, and the number of terms of $f^{q-1} - 1$ is $m^{q-1} + 1$. So the exact bound is $(m^{q-1} + 1)^{\log q}$.

\[ \frac{|\tilde{G}_{(f - c)^{q-1} - 1}|}{\#_c f} \leq \beta_2 / m^{q-1} = ((m + 1)^{q-1} - 1)^{\log q} \leq (m + 1)^{(q-1)^{\log q}}. \]

**Proof.** For $s \in GF[q]^n$, $f(s) = c \Leftrightarrow (f - c)^{q-1}(s) = 0 \Leftrightarrow (f - c)^{q-1}(s) - 1 \neq 0$. Observe that $(f - c)^{q-1} - 1$ polynomial is constant free. Apply Theorem 2 to the polynomial $(f - c)^{q-1} - 1$ with $|G| = \#_c f$ and $m^{q-1} - 1$ terms which results in $\beta_2 = ((m + 1)^{q-1} - 1)^{\log q}$.

\[ \frac{|\tilde{G}_{f}|}{\#_c f} = \frac{q^{n-1}}{(q-1)^n} \text{ tends to infinity with growing } n \text{ and does not satisfy the inequality } \leq q^{q-1}. \]

The bounds proven in Theorems 3, and 4 are almost optimal (cf. [GK 90]).

5 Open Problem

Our method yields the first polynomial time $(\epsilon, \delta)$-approximation algorithm for the number of zeros of arbitrary polynomials $f \in GF[q][x_1, \cdots, x_n]$ for the fixed field
$GF[q]$. Degree of the polynomial bounding the running time of the algorithm depend on the order of the ground field.

Is it possible to remove dependence of the degree on $q$ in the approximation algorithm?

Acknowledgements.

We are thankful to Dick Karp, Hendrik Lenstra, Barbara Lhotzky, Mike Luby, Andrew Odlyzko and Mario Szegedy for the number of fruitful discussions.

References


