VALUE DISTRIBUTION OF ELEMENTARY SYMMETRIC POLYNOMIALS AND ITS PERTURBATIONS OVER FINITE FIELDS

LUIS A. MEDINA, L. BREHSNER SEPÚLVEDA, AND CÉSAR A. SERNA-RAPELLO

Abstract. In this article we establish the asymptotic behavior of generating functions related to the exponential sum over finite fields of elementary symmetric functions and their perturbations. This asymptotic behavior allows us to calculate the probability generating function of the probability that the elementary symmetric polynomial of degree k and its perturbations returns $\beta \in \mathbb{F}_q$ where $\mathbb{F}_q$ represents the field of q elements. Our study extends many of the results known for perturbations over the binary field to any finite field. In particular, we establish when a particular perturbation is asymptotically balanced over a prime field and provide a construction to find such perturbations over any finite field.

In memory of Francis N. Castro.

1. Introduction

Many problems in number theory and combinatorics, as well as in their applications, can be formulated in terms of exponential sums. For example, exponential sums can be used to determine if a system of polynomial equations have solutions over a finite field. Exponential sums also have applications to other scientific fields. For instance, they can be used to detect when a particular function is balanced, which is a property very useful in cryptographic applications [5, 6, 7, 8, 9, 12, 13]. Some classical examples of exponential sums include the number-theoretical Gauss sums, Kloosterman sums, and Weyl sums.

This work is based on the study of exponential sums of the following form. Let $q = p^r$ where p is prime and $r \geq 1$. Let $\mathbb{F}_q$ represents the field of q elements and let $F : \mathbb{F}_q^n \to \mathbb{F}_q$ be a function. The *exponential sum over* $\mathbb{F}_q$ of F is defined as

$$S_{\mathbb{F}_q}(F) = \sum_{\mathbf{x}\in\mathbb{F}_q^n} \xi_p^{\mathrm{Tr}(F(\mathbf{x}))}, \tag{1.1}$$

where $\xi_p = \exp(2\pi i/p)$ and $\mathrm{Tr} = \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}$ is the field trace function. These exponential sums have been extensively studied when the characteristic of the field is 2 because of their cryptographic applications, see [3, 4, 5, 8, 9, 12, 13, 22]. Recently, some cryptographic applications beyond characteristic 2 have been found. That has prompted new research in exponential sums of the type (1.1) and many of the results available for the binary field have been extended to other finite fields [10, 11, 17, 19, 20, 21].

Let $L : \mathbb{F}_q \to \mathbb{F}_q$ be a linear function and X an indeterminate. Consider the generating function given by

$$\mathrm{S}_{\mathbb{F}_q,L}(F;X) = \sum_{\mathbf{y}\in\mathbb{F}_q^n} X^{L(F(\mathbf{y}))}. \tag{1.2}$$

Observe that when $L = \mathrm{Tr}$ and $X = \xi_p$ we recover the regular exponential sum $S_{\mathbb{F}_q}(F)$. Therefore, the study of regular exponential sums is embedded in the study of generating functions of the form (1.2). Thus, from now on, we consider the generating functions (1.2) instead of exponential sums of the form (1.1). Furthermore, in this article we use the term exponential sums to refer to both (1.1) and (1.2).

In [11], closed formulas for exponential sums of type (1.1) of elementary symmetric polynomials were found (extending the results of [3] to every finite field). There is a natural connection between the formulas presented in [11] and the value distribution of elementary symmetric polynomials over $\mathbb{F}_q$. Part of the focus of this article is to explain such connection and to extend it to perturbations of elementary symmetric polynomials.

Date: September 6, 2019.

2010 *Mathematics Subject Classification.* 05E05, 11T23.

Key words and phrases. Exponential sums, symmetric functions, value distribution.

Let k be a natural number. The elementary symmetric polynomial of degree k in the variables $X_1, \ldots, X_n$ is denoted by $\boldsymbol{e}_k(X_1, \ldots, X_n)$. Sometimes we use the more compact notation $\boldsymbol{e}_{n,k}$ to represent that polynomial, that is $\boldsymbol{e}_{n,k}$ also represents the n-variable elementary symmetric polynomial of degree k. In this article, we prefer to use the notation $\boldsymbol{e}_{n,k}$ to represent the elementary symmetric polynomial when it has not been evaluated and the notation $\boldsymbol{e}_k(\mathbf{x})$ when we want to stress that it has been evaluated at $\mathbf{x}$.

Let $F(\mathbf{X}) \in \mathbb{F}_q[X_1, \ldots, X_j]$ (j fixed). The polynomial $\boldsymbol{e}_{n,k} + F(\mathbf{X})$ is called a *perturbation* of the n-variable elementary symmetric polynomial of degree k. These perturbations were introduced in [7] for the binary case and are the main focus of [5, 7]. Perturbations break the symmetry of $\boldsymbol{e}_{n,k}$ and may reduce symmetry attacks in cryptographic implementations.

The value of $\boldsymbol{e}_k(\mathbf{x})$ is important when its exponential sums are studied. Consider the set $A = \{0, a_1, \ldots, a_s\}$, where a_j's are parameters. Suppose that $\mathbf{x} \in A^n$ and that a_j appears m_j times in $\mathbf{x}$. Following [11], the value of $\boldsymbol{e}_k(\mathbf{x})$ will be denoted by $\Lambda_{a_1,\ldots,a_s}(k, m_1, \ldots, m_s)$. In the particular case when the set A is the finite field $\mathbb{F}_q$ we use the notation $\Lambda_{\mathbb{F}_q^\times}(k, m_1, \ldots, m_{q-1})$. A recursive definition for $\Lambda_{a_1,\ldots,a_s}(k, m_1, \ldots, m_s)$, which allows for fast evaluations of it, appears in [11]:

$$\begin{aligned} \Lambda_{a_1}(k,m) &= a_1^k \binom{m}{k} \\ \Lambda_{a_1,a_2,\ldots,a_{l+1}}(k, m_1, m_2, \ldots, m_{l+1}) &= \sum_{j=0}^{m_{l+1}} \binom{m_{l+1}}{j} a_{l+1}^j \Lambda_{a_1,\ldots,a_l}(k-j, m_1, m_2, \ldots, m_l). \end{aligned} \tag{1.3}$$

As mentioned before, one of the main results of [11] are closed formulas for exponential sums of elementary symmetric polynomials over any finite field. For convenience, we include their result next. The result is written in terms of (1.2).

Theorem 1.1 ([11])**.** *Let n and $k > 1$ be positive integers, p be a prime and $q = p^r$ with $r \geq 1$. Let $L : \mathbb{F}_q \to \mathbb{F}_q$ a linear function, X an indeterminate and $D = p^{\lfloor \log_p(k) \rfloor + 1}$. Then,*

$$\mathrm{S}_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k}; X) = \sum_{j_1=0}^{D-1} \sum_{j_2=0}^{j_1} \cdots \sum_{j_{q-1}=0}^{j_{q-2}} c_{j_1,\ldots,j_{q-1};L}(k; X) \left(1 + \xi_D^{-j_1} + \cdots + \xi_D^{-j_{q-1}}\right)^n, \tag{1.4}$$

where

$$c_{j_1,\ldots,j_{q-1};L}(k; X) = \frac{1}{D^{q-1}} \sum_{b_{q-1}=0}^{D-1} \cdots \sum_{b_1=0}^{D-1} X^{L\left(\Lambda_{\mathbb{F}_q^\times}(k, b_1, \ldots, b_{q-1})\right)} \sum_{(j_1', \ldots, j_{q-1}') \in \mathrm{Sym}(j_1, \ldots, j_{q-1})} \xi_D^{j_1' b_{q-1} + \cdots + j_{q-1}' b_1},$$

$\xi_D = \exp(2\pi i/D)$, and $\mathrm{Sym}(j_1, \ldots, j_{q-1})$ is the set of all rearrangements of $(j_1, \ldots, j_{q-1})$.

***Remark* 1.2.** Theorem 1.1 can be generalized without too much effort to linear combinations of elementary symmetric polynomials. See [11] for more details.

Theorem 1.1 is a generalization of the results presented in [3] for the binary field. In [6], Castro and Medina used the closed formulas in [3] to calculate the asymptotic behavior of exponential sums of symmetric Boolean functions. A similar result is now available in every finite field, that is, Theorem 1.1 can be used to study the asymptotic behavior of exponential sums of the form $S_{\mathbb{F}_q}(\boldsymbol{e}_{n,k})$.

The rest of the article is divided into three sections. In the next one (Section 2) we study the asymptotic behavior of generating functions of the type (1.2) for elementary symmetric polynomials and their perturbations. One of the reasons to study such behavior is to explore the veracity of an open problem related to balancedness. The results presented in Section 2 generalize the results presented in [7] from the binary field to any finite field. In the third section, we study the distribution of the values of elementary symmetric polynomials over $\mathbb{F}_q$. To be more precise, we study the probability that $\boldsymbol{e}_k(\mathbf{x})$ returns $\beta \in \mathbb{F}_q$ when the entries of $\mathbf{x}$ are randomly selected from $\mathbb{F}_q$. We also introduce the concept of asymptotically balanced symmetric polynomial and asymptotically balanced perturbation and show that a perturbation $\boldsymbol{e}_{n,k} + F(\mathbf{X})$ is asymptotically balanced over $\mathbb{F}_p$ (p prime) if and only if $\boldsymbol{e}_{n,k}$ is asymptotically balanced or $F(\mathbf{X})$ is balanced over $\mathbb{F}_p$. We also show that this statement is not true for finite fields in general and provide a way to construct counterexamples. Finally, we finish the article with some concluding remarks.

2. Asymptotic behavior of elementary symmetric polynomials and their perturbations

A function $F:\mathbb{F}_q^n\to\mathbb{F}_q$ is said to be *balanced* if its values are equally distributed. That is, if F takes each value of $\mathbb{F}_q$ exactly q^{n-1} times. Balancedness is important in some cryptographic implementations. That is especially true when the characteristic of the field is 2.

There is an important conjecture proposed by Cusick, Li and Stănică about the balancedness of elementary symmetric polynomials over the binary field [13]. Their conjecture states:

Conjecture 2.1 ([13])**.** *There are no nonlinear balanced elementary symmetric Boolean functions except for degree $k=2^\ell$ and $2^{\ell+1}D-1$-variables, where ℓ, D are positive integers.*

A generalized version of this conjecture for finite fields was presented in [2].

Conjecture 2.2 ([2])**.** *The only nonlinear balanced elementary symmetric polynomial over $\mathbb{F}_q$, $q=p^r$ are those with degree $k=p^\ell$ and $n=p^\ell D-1$ variables, where $\ell, D\in\mathbb{N}$, $D\not\equiv 1 \mod p$.*

It is known that Conjecture 2.1 is true asymptotically [6, 15, 16]. In particular, the argument presented in [6] depends on a calculation of the asymptotic behavior of the exponential sum $S_{\mathbb{F}_2}(\boldsymbol{e}_{n,k})$. Thus, to explore Conjecture 2.2, it is natural to study the asymptotic behavior of $\mathrm{S}_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k};X)$. Theorem 1.1 can be used to do that.

Consider the closed formula (1.4) for $\mathrm{S}_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k},X)$. Observe that q is the biggest modulus of all complex numbers of the form

$$1+\xi_D^{-j_1}+\cdots+\xi_D^{-j_{q-1}},\ \text{for } 0\le j_{q-1}\le j_{q-2}\le\cdots\le j_1\le D-1.$$

This maximum modulus is achieved if and only if $j_1=\cdots=j_{q-1}=0$, which implies

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{q^n}\mathrm{S}_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k},X) &= c_{0,\dots,0;L}(k;X)\\ &= \frac{1}{D^{q-1}}\sum_{b_{q-1}=0}^{D-1}\cdots\sum_{b_1=0}^{D-1}X^{L\left(\Lambda_{\mathbb{F}_q^\times}(k,b_1,\dots,b_{q-1})\right)}.\end{aligned}\tag{2.1}$$

Therefore, the asymptotic behavior of $\mathrm{S}_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k},X)$ is dominated by $c_{0,\dots,0}(k;X)\cdot q^n$. We relabel $c_{0,\dots,0;L}(k;X)$ as $c_{A,L}^{(q)}(k;X)$, i.e.

$$c_{A,L}^{(q)}(k;X)=\frac{1}{D^{q-1}}\sum_{b_{q-1}=0}^{D-1}\cdots\sum_{b_1=0}^{D-1}X^{L\left(\Lambda_{\mathbb{F}_q^\times}(k,b_1,\dots,b_{q-1})\right)}.\tag{2.2}$$

This is done in order to stress that (2.2) is the asymptotic coefficient for $\boldsymbol{e}_{n,k}$ over $\mathbb{F}_q$. Observe that the value of $c_{A,L}^{(q)}(k;X)$ depends on knowing how many times, for each $\alpha\in\mathbb{F}_q$, $\Lambda_{\mathbb{F}_q^\times}(k,b_1,\dots,b_{q-1})=\alpha$ in a $q-1$-hypercube of side length D. That is a very interesting combinatorial problem on its own. For example, if we consider $q=3$, $k=27$ and $L(x)=x$, and color a point in the 81×81 grid $\{(a,b):0\le a,b\le 80\}$ blue if $\Lambda_{\mathbb{F}_3^\times}(27,a,b)=0$, red if $\Lambda_{\mathbb{F}_3^\times}(27,a,b)=1$ and green if $\Lambda_{\mathbb{F}_3^\times}(27,a,b)=2$, then we get the picture in Figure 1.

The above argument can be easily extended to linear combinations of elementary symmetric polynomials. Let $0<k_1<\cdots<k_s$ be integers and $\boldsymbol{k}=(k_1,\dots,k_s)$. Let $\beta_1,\dots,\beta_s\in\mathbb{F}_q$ and $\boldsymbol{\beta}=(\beta_1,\dots,\beta_s)$. Finally, let $D=p^{\lfloor\log_p(k_s)\rfloor+1}$. Then,

$$\lim_{n\to\infty}\frac{1}{q^n}\mathrm{S}_{\mathbb{F}_q,L}\left(\sum_{t=1}^s\beta_t\boldsymbol{e}_{n,k_t};X\right)=c_{A,L}^{(q)}(\boldsymbol{k},\boldsymbol{\beta};X)\tag{2.3}$$

where

$$c_{A,L}^{(q)}(\boldsymbol{k},\boldsymbol{\beta};X)=\frac{1}{D^{q-1}}\sum_{b_{q-1}=0}^{D-1}\cdots\sum_{b_1=0}^{D-1}X^{L\left(\sum_{t=0}^s\beta_t\Lambda_{\mathbb{F}_q^\times}(k_t,b_1,\dots,b_{q-1})\right)}.\tag{2.4}$$

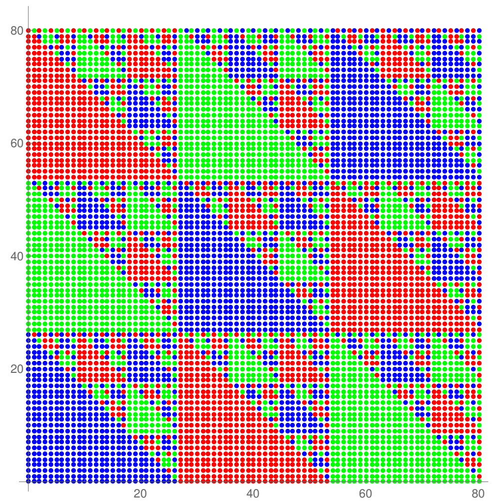

FIGURE 1. Graphical representation of the values of $\Lambda_{\mathbb{F}_3^\times}(27, a, b)$ for $0 \leq a, b \leq 80$.

Many of the results presented in [11] and the one presented so far can be extended to perturbations of elementary symmetric polynomials. Let $F(\mathbf{Y}) \in \mathbb{F}_q[Y_1, \ldots, Y_j]$ (j fixed). Consider the identity

$$\boldsymbol{e}_k(Y_1, \ldots, Y_n) = \sum_{m=0}^{j} \boldsymbol{e}_m(Y_1, \ldots, Y_j)\boldsymbol{e}_{k-m}(Y_{j+1}, Y_{j+2}, \ldots, Y_n). \tag{2.5}$$

Observe that

$$\begin{aligned} \mathrm{S}_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k} + F(\mathbf{Y}); X) &= \sum_{(\beta_1,\ldots,\beta_j,y_{j+1},\ldots,y_n)\in\mathbb{F}_q^n} X^{L(F(\beta_1,\ldots,\beta_j))} X^{L(\boldsymbol{e}_k(\beta_1,\ldots,\beta_j,y_{j+1},\ldots,y_n))} \\ &= \sum_{\boldsymbol{\beta}\in\mathbb{F}_q^j} X^{L(F(\boldsymbol{\beta}))} \sum_{(y_{j+1},\ldots,y_n)\in\mathbb{F}_q^{n-j}} X^{L\left(\sum_{m=0}^{j} \boldsymbol{e}_m(\boldsymbol{\beta})\boldsymbol{e}_{k-m}(y_{j+1},\ldots,y_n)\right)} \\ &= \sum_{\boldsymbol{\beta}\in\mathbb{F}_q^j} X^{L(F(\boldsymbol{\beta}))} \mathrm{S}_{\mathbb{F}_q,L}\left(\sum_{m=0}^{j} \boldsymbol{e}_m(\boldsymbol{\beta})\boldsymbol{e}_{n-j,k-m}; X\right). \end{aligned} \tag{2.6}$$

A corollary of (2.6) is that exponential sums of perturbations of $\boldsymbol{e}_{n,k}$ have closed formulas similar to (1.4). Moreover, it is also true that they satisfy the linear recurrences presented in [11]. This implies that a result similar to (2.1) is expected. The next two lemmas are going to be used to prove such claim.

Lemma 2.3. *Consider the set $A = \{0, a_1, \ldots, a_s\}$ where a_j are parameters. Let $m_1, \ldots, m_s, l_1, \ldots, l_s$ be non-negative integers. Suppose that $j \geq m_1 + \cdots + m_s$. Then,*

$$\sum_{m=0}^{j} \Lambda_{a_1,\ldots,a_s}(m, m_1, \ldots, m_s)\Lambda_{a_1,\ldots,a_s}(k-m, l_1, \ldots, l_s) = \Lambda_{a_1,\ldots,a_s}(k, m_1 + l_1, \ldots, m_s + l_s). \tag{2.7}$$

Proof. This is a natural consequence of the equation (2.5) and the fact that if $\mathbf{x} \in A^n$ is such that a_i appears n_i times in $\mathbf{x}$, then $\boldsymbol{e}_k(\mathbf{x}) = \Lambda_{a_1,\ldots,a_s}(m, n_1, \ldots, n_s)$. Observe that j must be bigger than or equal to $m_1 + \cdots + m_s$ in order to have the necessary amount of variables to support $m_1 + \cdots + m_s$ values. □

Lemma 2.4. *Let k be a positive integer. Suppose that $\beta_1, \ldots, \beta_j \in \mathbb{F}_q$ where $q = p^r$ with p prime and that $L : \mathbb{F}_q \to \mathbb{F}_q$ is a linear function. Then,*

$$\lim_{n\to\infty} \frac{1}{q^n} \mathrm{S}_{\mathbb{F}_q,L}\left(\sum_{m=0}^{j} \boldsymbol{e}_m(\beta_1, \ldots, \beta_j)\boldsymbol{e}_{n,k-m}; X\right) = \lim_{n\to\infty} \frac{1}{q^n} \mathrm{S}_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k}; X) = c_{A,L}^{(q)}(k; X). \tag{2.8}$$

Proof. Let $\boldsymbol{\beta}=(\beta_1,\dots,\beta_j)$ and $D=p^{\lfloor \log_p(k)\rfloor+1}$. Recall that

$$
\begin{aligned}
c_{A,L}^{(q)}(k;X) &= \lim_{n\to\infty}\frac{1}{q^n}\mathrm{S}_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k};X) \\
c_{A,L}^{(q)}(k,\boldsymbol{\beta};X) &= \lim_{n\to\infty}\frac{1}{q^n}\mathrm{S}_{\mathbb{F}_q,L}\left(\sum_{m=0}^{j}\boldsymbol{e}_m(\beta_1,\dots,\beta_j)\boldsymbol{e}_{n,k-m};X\right)
\end{aligned}
\tag{2.9}
$$

where

$$
\begin{aligned}
c_{A,L}^{(q)}(k;X) &= \frac{1}{D^{q-1}}\sum_{b_{q-1}=0}^{D-1}\cdots\sum_{b_1=0}^{D-1}X^{L(\Lambda_{\mathbb{F}_q^\times}(k,b_1,\dots,b_{q-1}))} \\
c_{A,L}^{(q)}(k,\boldsymbol{\beta};X) &= \frac{1}{D^{q-1}}\sum_{b_{q-1}=0}^{D-1}\cdots\sum_{b_1=0}^{D-1}X^{L\left(\sum_{m=0}^{j}\boldsymbol{e}_m(\beta_1,\dots,\beta_j)\Lambda_{\mathbb{F}_q^\times}(k-m,b_1,\dots,b_{q-1})\right)}.
\end{aligned}
\tag{2.10}
$$

Let us work with $c_{A,L}^{(q)}(k,\boldsymbol{\beta};X)$. Suppose that α_t appears b_t' times in the entries of the vector $(\beta_1,\dots,\beta_j)$. Then, Lemma 2.3 implies

$$
\begin{aligned}
c_{A,L}^{(q)}(k,\boldsymbol{\beta};X) &= \frac{1}{D^{q-1}}\sum_{b_{q-1}=0}^{D-1}\cdots\sum_{b_1=0}^{D-1}X^{L\left(\sum_{m=0}^{j}\boldsymbol{e}_m(\beta_1,\dots,\beta_j)\Lambda_{\mathbb{F}_q^\times}(k-m,b_1,\dots,b_{q-1})\right)} \\
&= \frac{1}{D^{q-1}}\sum_{b_{q-1}=0}^{D-1}\cdots\sum_{b_1=0}^{D-1}X^{L(\Lambda_{\mathbb{F}_q^\times}(k,b_1+b_1',\dots,b_{q-1}+b_{q-1}'))}.
\end{aligned}
\tag{2.11}
$$

However, we know that $\Lambda_{\mathbb{F}_q^\times}(k,m_1,\dots,m_{q-1})$ is periodic mod p in each of the entries $m_1,\dots,m_{q-1}$ with period length D (see [11]). Since each of the variables b_t runs a full period, i.e. from 0 to $D-1$, then

$$
\begin{aligned}
c_{A,L}^{(q)}(k,\boldsymbol{\beta};X) &= \frac{1}{D^{q-1}}\sum_{b_{q-1}=0}^{D-1}\cdots\sum_{b_1=0}^{D-1}X^{L(\Lambda_{\mathbb{F}_q^\times}(k,b_1+b_1',\dots,b_{q-1}+b_{q-1}'))} \\
&= \frac{1}{D^{q-1}}\sum_{d_{q-1}=0}^{D-1}\cdots\sum_{d_1=0}^{D-1}X^{L(\Lambda_{\mathbb{F}_q^\times}(k,d_1,\dots,d_{q-1}))} \\
&= c_{A,L}^{(q)}(k;X).
\end{aligned}
\tag{2.12}
$$

This concludes the proof. □

Next is the generalization of (2.1). As before, it is stated for perturbations of elementary symmetric polynomials, but the same holds true for perturbations of linear combinations of them.

Theorem 2.5. *Let $k>1$ be an integer, p a prime and $q=p^r$ with $r\geq 1$. Suppose that $F(\mathbf{Y})$ is a polynomial in the variables $Y_1,\dots,Y_j$ (j fixed) with coefficients from $\mathbb{F}_q$ and that $L:\mathbb{F}_q\to\mathbb{F}_q$ is a linear function. Then,*

$$
\lim_{n\to\infty}\frac{1}{q^n}\mathrm{S}_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k}+F(\mathbf{Y});X)=\frac{1}{q^j}c_{A,L}^{(q)}(k;X)\mathrm{S}_{\mathbb{F}_q,L}(F;X).
\tag{2.13}
$$

Proof. Equation 2.6 implies

$$
\mathrm{S}_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k}+F(\mathbf{Y});X)=\sum_{\boldsymbol{\beta}\in\mathbb{F}_q^j}X^{L(F(\boldsymbol{\beta}))}\mathrm{S}_{\mathbb{F}_q,L}\left(\sum_{m=0}^{j}\boldsymbol{e}_m(\boldsymbol{\beta})\boldsymbol{e}_{n-j,k-m};X\right).
\tag{2.14}
$$

Therefore,

$$
\begin{aligned}
\lim_{n\to\infty}\frac{1}{q^n}\mathrm{S}_{\mathbb{F}_q}(\boldsymbol{e}_{n,k}+F(\mathbf{Y});X) &= \sum_{\boldsymbol{\beta}\in\mathbb{F}_q^j} X^{L(F(\boldsymbol{\beta}))}\left(\lim_{n\to\infty}\frac{1}{q^n}\mathrm{S}_{\mathbb{F}_q,L}\left(\sum_{m=0}^{j}\boldsymbol{e}_m(\boldsymbol{\beta})\boldsymbol{e}_{n-j,k-m};X\right)\right)\\
&= \frac{1}{q^j}\sum_{\boldsymbol{\beta}\in\mathbb{F}_q^j} X^{L(F(\boldsymbol{\beta}))}\left(\lim_{n\to\infty}\frac{1}{q^{n-j}}\mathrm{S}_{\mathbb{F}_q,L}\left(\sum_{m=0}^{j}\boldsymbol{e}_m(\boldsymbol{\beta})\boldsymbol{e}_{n-j,k-m};X\right)\right)\\
&= \frac{1}{q^j}\sum_{\boldsymbol{\beta}\in\mathbb{F}_q^j} X^{L(F(\boldsymbol{\beta}))}c_{A,L}^{(q)}(k;X)\\
&= \frac{1}{q^j}c_{A,L}^{(q)}(k;X)\mathrm{S}_{\mathbb{F}_q,L}(F;X),
\end{aligned}
\tag{2.15}
$$

where the third equation comes from Lemma 2.4. This concludes the proof. □

Theorem 2.5 is also a generalization of the main theorem of [7, Th. 4.4]. In fact, the discussion so far about perturbations of elementary symmetric polynomials generalizes most of the results presented in [7] for the binary field. In the next section we show how the results presented in this section can be used to study the distribution of the values of $\boldsymbol{e}_{n,k}$ (and its perturbations) in finite fields.

3. Distribution of the values of elementary symmetric polynomials and their perturbations over $\mathbb{F}_q$

The generating function $\mathrm{S}_{\mathbb{F}_q;L}(F;X)$ can be written as

$$\mathrm{S}_{\mathbb{F}_q,L}(F;X) = \sum_{\beta\in\mathbb{F}_q} N_{\mathbb{F}_q,L}(F;\beta)X^{\beta}, \tag{3.1}$$

where

$$N_{\mathbb{F}_q,L}(F;\beta) = \left|\{\mathbf{x}\in\mathbb{F}_q^n \,:\, L(F(\mathbf{x}))=\beta\}\right|,\ \ \beta\in\mathbb{F}_q.^1 \tag{3.2}$$

This implies that

$$\frac{1}{q^n}\mathrm{S}_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k};X) = \sum_{\beta\in\mathbb{F}_q} \mathbb{p}_{n,k}^{(q)}(\beta;L)X^{\beta} \tag{3.3}$$

where $\mathbb{p}_{n,k}^{(q)}(\beta;L) = N_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k};\beta)/q^n$ is the probability that $L(\boldsymbol{e}_k(\mathbf{x}))$ returns the value $\beta\in\mathbb{F}_q$ when $\mathbf{x}$ is randomly selected from $\mathbb{F}_q^n$.

Equation (2.1) states that

$$\lim_{n\to\infty}\frac{1}{q^n}\mathrm{S}_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k};X) = c_{A,L}^{(q)}(k;X). \tag{3.4}$$

Expressing $c_{A,L}^{(q)}(k;X)$ as

$$c_{A,L}^{(q)}(k;X) = \sum_{\beta\in\mathbb{F}_q} a_\beta X^{\beta}, \tag{3.5}$$

we see that

$$a_\beta = \lim_{n\to\infty}\mathbb{p}_{n,k}^{(q)}(\beta;L) = \lim_{n\to\infty}\frac{N_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k};\beta)}{q^n} := \mathbb{p}_k^{(q)}(\beta;L). \tag{3.6}$$

The limit in (3.6) exists and we call $\mathbb{p}_k^{(q)}(\beta;L)$ the *probability at infinity* that $L(\boldsymbol{e}_k(\mathbf{x}))$ returns the value $\beta\in\mathbb{F}_q$ when $\mathbf{x}$ is randomly selected with entries from $\mathbb{F}_q$. Clearly, if $\varepsilon>0$, then for all n big enough,

$$\left|\mathbb{p}_{n,k}^{(q)}(\beta;L)-\mathbb{p}_k^{(q)}(\beta;L)\right| < \varepsilon \quad \text{for all } \beta\in\mathbb{F}_q. \tag{3.7}$$

[1]Observe that since $\{\mathrm{S}_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k};X)\}_{n\in\mathbb{N}}$ satisfies the linear recurrence with integer coefficients provided in [11, Th. 5.7], then $\{N_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k};\beta)\}_{n\in\mathbb{N}}$ also satisfies such recurrence. Also, $N_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k};\beta)$ has a closed formula similar to the one in Theorem 1.1.

Also,

$$N_{\mathbb{F}_q,L}(\boldsymbol{e}_{n,k};\beta) \sim \mathbb{p}_k^{(q)}(\beta;L)\cdot q^n. \tag{3.8}$$

Observe that under this setting $c_{A,L}^{(q)}(k;X)$ is the probability generating function for $\mathbb{p}_k^{(q)}(\beta;L)$. Therefore, the study of the distribution of the values of $L(\boldsymbol{e}_k(\mathbf{X}))$ in $\mathbb{F}_q$ is equivalent to the study of $c_{A,L}^{(q)}(k;X)$. We write $\mathrm{G}_{n,k}^{(q)}(L;X)$ for the probability generating function of $\mathbb{p}_{n,k}^{(q)}(\beta;L)$, that is

$$\mathrm{G}_{n,k}^{(q)}(L;X) = \sum_{\beta\in\mathbb{F}_q} \mathbb{p}_{n,k}^{(q)}(\beta;L)X^\beta. \tag{3.9}$$

We also relabel $c_{A,L}^{(q)}(k;X)$ as $\mathrm{G}_k^{(q)}(L;X)$ in an attempt to make the fact that $c_{A,L}^{(q)}(k;X)$ is the probability generating function of $\mathbb{p}_k^{(q)}(\beta;L)$ more clear.

The next theorem summarizes the discussion so far. Again, it is stated for elementary symmetric polynomials, but it can be extended to linear combinations of them.

Theorem 3.1. *Let p be a prime, $q=p^r$ where $r\geq 1$ and $L:\mathbb{F}_q\to\mathbb{F}_q$ be a linear function. Suppose that $k>1$ is an integer and $D=p^{\lfloor \log_p(k)\rfloor+1}$. Then,*

$$\lim_{n\to\infty}\mathrm{G}_{n,k}^{(q)}(L;X) = \frac{1}{D^{q-1}}\sum_{b_{q-1}=0}^{D-1}\cdots\sum_{b_1=0}^{D-1} X^{L\left(\Lambda_{\mathbb{F}_q^\times}(k,b_1,\ldots,b_{q-1})\right)} = \mathrm{G}_k^{(q)}(L;X). \tag{3.10}$$

The study of perturbations of the form $\boldsymbol{e}_{n,k}+F(\mathbf{X})$ follows in an analogous way. We use the notation $\mathrm{G}_{n,k;F}^{(q)}(L;X)$ to represent

$$\mathrm{G}_{n,k;F}^{(q)}(L;X) = \sum_{\beta\in\mathbb{F}_q}\mathbb{p}_{n,k;F}^{(q)}(\beta;L)X^\beta, \tag{3.11}$$

with $\mathbb{p}_{n,k;F}^{(q)}(\beta;L)$ defined in the natural way. As in the previous discussion, the limit

$$\lim_{n\to\infty}\mathbb{p}_{n,k;F}^{(q)}(\beta;L) \tag{3.12}$$

exists. The value of the limit is denoted by $\mathbb{p}_{k;F}^{(q)}(\beta;L)$ and we use $\mathrm{G}_{k;F}^{(q)}(L;X)$ to represent the probability generating function of $\mathbb{p}_{k;F}^{(q)}(\beta;L)$. Observe that the conclusion of Theorem 2.5 can be re-stated as

$$\begin{aligned}\lim_{n\to\infty}\mathrm{G}_{n,k;F}^{(q)}(L;X) &= \mathrm{G}_{k;F}^{(q)}(L;X)\\ &= \frac{1}{q^j}\mathrm{G}_k^{(q)}(L;X)\mathrm{S}_{\mathbb{F}_q,L}(F;X).\end{aligned} \tag{3.13}$$

***Remark* 3.2.** When $L(X)=X$, we drop the "L" in the notation of our functions. For example, we write $\mathbb{p}_{n,k}^{(q)}(\beta)$ instead of $\mathbb{p}_{n,k}^{(q)}(\beta;L)$ or $\mathrm{G}_k^{(q)}(X)$ instead of $\mathrm{G}_k^{(q)}(L;X)$.

Example 3.3. Consider the polynomial $\boldsymbol{e}_{n,5}$ over $\mathbb{F}_4=\mathbb{F}_2(\alpha)$ with $\alpha^2+\alpha+1=0$. In this case,

$$\mathrm{G}_5^{(4)}(X) = \frac{11}{32}+\frac{7}{32}X+\frac{7}{32}X^\alpha+\frac{7}{32}X^{\alpha+1}. \tag{3.14}$$

This implies that the probability at infinity that $\boldsymbol{e}_5(\mathbf{x})$ returns 0 is $11/32$ and the probabilities that it returns $1,\alpha$ and $\alpha+1$ are all $7/32$.

Let $F(\mathbf{X})=X_1X_2+X_1X_3X_2+X_3X_2+X_1X_3$ and consider the perturbation polynomial $\boldsymbol{e}_{n,5}+F(\mathbf{X})$. Theorem 2.5 implies that

$$\begin{aligned}
\mathrm{G}_{5;F}^{(4)}(X) &= \frac{1}{4^3}\mathrm{G}_5^{(4)}(X)\mathrm{S}_{\mathbb{F}_4}(F;X) \\
&= \frac{1}{64}\left(\frac{11}{32}+\frac{7}{32}X+\frac{7}{32}X^{\alpha}+\frac{7}{32}X^{\alpha+1}\right)\left(17+21X+13X^{\alpha}+13X^{\alpha+1}\right) \\
&= \frac{187}{2048}+\frac{175}{1024}X+\frac{147}{2048}X^2+\frac{131}{1024}X^{\alpha}+\frac{91}{2048}X^{2\alpha}+ \\
&\quad \frac{125}{512}X^{\alpha+1}+\frac{119}{1024}X^{\alpha+2}+\frac{91}{1024}X^{2\alpha+1}+\frac{91}{2048}X^{2\alpha+2} \\
&= \frac{129}{512}+\frac{133}{512}X+\frac{125}{512}X^{\alpha}+\frac{125}{512}X^{\alpha+1},
\end{aligned} \tag{3.15}$$

where the last equation comes from the fact that we are working on $\mathbb{F}_4=\mathbb{F}_2(\alpha)$. Observe that this implies that the probability at infinity that $\boldsymbol{e}_5(\mathbf{x})+F(\mathbf{x})$ returns 0 is 129/512, the probability it returns 1 is 133/512 and the probabilities that it returns α and $\alpha+1$ are 125/512 each.

Example 3.4. Consider now the polynomial $\boldsymbol{e}_{n,4}$ over $\mathbb{F}_9=\mathbb{F}_3(\alpha)$, where $\alpha^2+1=0$. In this case,

$$\mathrm{G}_4^{(9)}(X)=\sum_{\beta\in\mathbb{F}_3}\frac{29}{243}X^{\beta}+\sum_{\beta\in\mathbb{F}_9\setminus\mathbb{F}_3}\frac{26}{243}X^{\beta}. \tag{3.16}$$

Consider now the perturbation $\boldsymbol{e}_{n,4}+F(\mathbf{X})$ where $F(\mathbf{X})=X_1X_2X_3+X_1X_2+X_3$. Observe that

$$\begin{aligned}
\mathrm{G}_{4;F}^{(9)}(X) &= \frac{1}{9^3}\mathrm{G}_4^{(9)}(X)\mathrm{S}_{\mathbb{F}_9}(F;X) \\
&= \frac{1}{729}\left(\sum_{\beta\in\mathbb{F}_3}\frac{29}{243}X^{\beta}+\sum_{\beta\in\mathbb{F}_9\setminus\mathbb{F}_3}\frac{26}{243}X^{\beta}\right)\left(145X^2+\sum_{\beta\in\mathbb{F}_9\setminus\{2\}}73X^{\beta}\right) \\
&= \sum_{\beta\in\mathbb{F}_3}\frac{2203}{19683}X^{\beta}+\sum_{\beta\in\mathbb{F}_9\setminus\mathbb{F}_3}\frac{2179}{19683}X^{\beta}.
\end{aligned} \tag{3.17}$$

One of the first persons to study (if not the first one) the asymptotic distribution of the values of elementary symmetric polynomials over finite fields of *prime order* was N. J. Fine [15]. He proved that $\mathbb{p}_k^{(p)}(t)$ exists for every prime p and calculated the distribution of $\mathbb{p}_k^{(p)}(t)$ for $p=2,3$. He also established that for p equal to 2 or 3 (highlighted by Aberth [1]),

(1) $\mathbb{p}_k^{(p)}(0)\geq 1/p$,
(2) $\mathbb{p}_k^{(p)}(0)=1/p$ only if $k=d\cdot p^l$ where $1\leq d\leq p-1$,
(3) $\mathbb{p}_k^{(p)}(t)=1/p$ if $k=d\cdot p^l$ where $1\leq d\leq p-1$,
(4) $\mathbb{p}_{kp}^{(p)}(t)=\mathbb{p}_k^{(p)}(t)$,
(5) $\mathbb{p}_k^{(p)}(0)\geq \mathbb{p}_k^{(p)}(t)$ with equality only for $k=d\cdot p^l$ where $1\leq d\leq p-1$.

Fine also proved (3) for all p, which implies that the proof of the generalization of the conjecture of Cusick, Li and Stănică presented in [2] is expected to be much harder than the binary counterpart. In particular, when $p>2$, the approach presented in [6] will fail to prove the conjecture asymptotically when $k=d\cdot p^l$ and $1<d\leq p-1$.

Fine proposed as problems the veracity of the other properties for general p. O. Aberth [1] disproved (2) and (5) by showing that $\mathbb{p}_6^{(5)}(0)=1/5$ and $\mathbb{p}_6^{(5)}(2)=26/125$. He also showed that $\mathbb{p}_{30}^{(5)}(0)=15749/78125>1/5=\mathbb{p}_6^{(5)}(0)$ and therefore (4) is also false. In [24], J. D. Smith generalized Aberth's example and showed that if $p>3$ is prime, then

$$\mathbb{p}_{p+1}^{(p)}(t)=\begin{cases}\frac{1}{p}, & t=0\\ \frac{1}{p}+\left(\frac{2t}{p}\right)\frac{1}{p^{\mu}}, & t\neq 0,\end{cases} \tag{3.18}$$

where $\mu=(p+1)/2$ and $\left(\frac{a}{p}\right)$ represents the Legendre symbol. Smith's general formula for $\mathbb{p}_k^{(p)}(t)$ as a multisum coincides with our formula in Theorem 3.1 for $q=p$ and $L(X)=X$.

As mentioned at the beginning of Section 2, one of the reasons the asymptotic behavior of exponential sums of symmetric polynomials was calculated over the binary field was to provide an asymptotic proof of Conjecture 2.1 (see [6]). The concept of *asymptotically balanced symmetric Boolean function* was introduced in [6] to mean that $c_A^{(2)}(k;-1) = 0$. Conjecture 2.1 was proved asymptotically in [6] by showing that $\boldsymbol{e}_{n,k}$ is asymptotically balanced if and only if k is a power of two. Observe that if a polynomial is not asymptotically balanced, then we know that it is not balanced for a sufficiently large number of variables. Thus, asymptotically balanced polynomials are good candidates for regular balancedness. The concept of asymptotically balanced polynomials was extended to perturbations of elementary symmetric polynomials in [7].

A natural generalization for the concept of asymptotically balanced symmetric polynomial over $\mathbb{F}_q$ is to say that a polynomial $\boldsymbol{e}_{n,k}$ is *asymptotically balanced* if and only if

$$\mathbb{p}_k^{(q)}(\beta) = \frac{1}{q}, \quad \text{for every } \beta \in \mathbb{F}_q. \tag{3.19}$$

The concept can also be extended to perturbations in the only natural way, that is, by saying that a perturbation $\boldsymbol{e}_{n,k} + F(\mathbf{X})$ is asymptotically balanced if and only if

$$\mathbb{p}_{k;F}^{(q)}(\beta) = \frac{1}{q}, \quad \text{for every } \beta \in \mathbb{F}_q. \tag{3.20}$$

Observe that Fine [15] proved that $\boldsymbol{e}_{n,k}$ is asymptotically balanced over the prime field $\mathbb{F}_p$ when $k = d \cdot p^l$ where $1 \leq d \leq p-1$. In [2, Th. 2], it was proved that if $q = p^r$, then $\boldsymbol{e}_{n,p^\ell}$ is asymptotically balanced over $\mathbb{F}_q$ for every ℓ.

One of the main goals in [7] was to identify when a particular pertubation is asymptotically balanced over $\mathbb{F}_2$. It was showed [7, Cor. 4.6] that a perturbation $\boldsymbol{e}_{n,k} + F(\mathbf{X})$ is asymptotically balanced over $\mathbb{F}_2$ if and only if $\boldsymbol{e}_{n,k}$ is asymptotically balanced or $F(\mathbf{X})$ is a balanced function. The same result holds true over any prime field, but it is not necessarily true over finite fields in general.

Proposition 3.5. *Let p be a prime. Suppose that $F(\mathbf{X}) \in \mathbb{F}_p[X_1, \dots, X_j]$ (j fixed). Then,*

$$\mathbb{p}_{k;F}^{(p)}(t) = \frac{1}{p}, \quad \textit{for every } t \in \mathbb{F}_p$$

if and only if $\mathbb{p}_k^{(p)}(t) = 1/p$ for every $t \in \mathbb{F}_p$ or $\mathrm{S}_{\mathbb{F}_p}(F;X) = \sum_{t\in\mathbb{F}_p} p^{j-1}X^t$. In other words, $\boldsymbol{e}_{n,k} + F(\mathbf{X})$ is asymptotically balanced if and only if $\boldsymbol{e}_{n,k}$ is asymptotically balanced or $F(\mathbf{X})$ is balanced.

Proof. Theorem 2.5 implies that

$$\mathrm{G}_{k;F}^{(p)}(X) = \frac{1}{p^j}\mathrm{G}_k^{(p)}(X)\mathrm{S}_{\mathbb{F}_p}(F;X). \tag{3.21}$$

Suppose first that $\mathbb{p}_k^{(p)}(t) = 1/p$ for every $t \in \mathbb{F}_p$ or $\mathrm{S}_{\mathbb{F}_p}(F;X) = \sum_{t\in\mathbb{F}_p} p^{j-1}X^t$. Then, the equation

$$\sum_{\beta\in\mathbb{F}_q} a_\beta X^\beta \sum_{\beta\in\mathbb{F}_q} X^\beta = \sum_{\beta\in\mathbb{F}_q}\left(\sum_{\gamma\in\mathbb{F}_q} a_\gamma\right) X^\beta, \tag{3.22}$$

which is true for any finite field $\mathbb{F}_q$, together with (3.21) imply that the coefficients of $\mathrm{G}_{k;F}^{(p)}(X)$ are all equal. But that can only be true if

$$\mathbb{p}_{k;F}^{(p)}(t) = \frac{1}{p}, \quad \text{for every } t \in \mathbb{F}_p.$$

To prove the other direction, let $X = \xi_p = \exp(2\pi i/p)$. That transforms (3.21) into

$$\mathrm{G}_{k;F}^{(p)}(\xi_p) = \frac{1}{p^j}\mathrm{G}_k^{(p)}(\xi_p)S_{\mathbb{F}_p}(F), \tag{3.23}$$

where $S_{\mathbb{F}_p}(F)$ is the regular exponential sum of F (a complex number). If it is true that $\mathbb{p}_{k;F}^{(p)}(t) = 1/p$ for every $t \in \mathbb{F}_p$, then

$$\mathrm{G}_{k;F}^{(p)}(\xi_p) = \sum_{t\in\mathbb{F}_p} \frac{1}{p}\,\xi_p^t = 0.$$

But then

$$\frac{1}{p^j}\mathrm{G}_k^{(p)}(\xi_p)S_{\mathbb{F}_p}(F) = 0,$$

and so $\mathrm{G}_k^{(p)}(\xi_p) = 0$ or $S_{\mathbb{F}_p}(F) = 0$. If the latter is true, then $F(\mathbf{X})$ is balanced over $\mathbb{F}_p$. If $\mathrm{G}_k^{(p)}(\xi_p) = 0$, then the minimal polynomial of ξ_p, i.e. $\Phi_p(X) = 1 + X + X^2 + \cdots + X^{p-1}$, divides the polynomial $\mathrm{G}_k^{(p)}(X)$. Since both polynomials are of the same degree, then $\mathrm{G}_k^{(p)}(X)$ is a constant multiple of $\Phi_p(X)$. We conclude that $\mathbb{p}_k^{(p)}(t) = 1/p$ for every $t \in \mathbb{F}_p$, i.e. $\boldsymbol{e}_{n,k}$ is asymptotically balanced. This concludes the proof. □

Proposition 3.5 is not true for $\mathbb{F}_q$ when q is not prime. The sufficient part still holds and is a consequence of equation (3.22), but the necessary part is not true in general. Next we present a method to construct counterexamples for the statement of Proposition 3.5 over $\mathbb{F}_q$.

3.1. **A construction for counterexamples over $\mathbb{F}_q$.** Let $q = p^r$ with $r > 1$. We want to find an elementary symmetric polynomial $\boldsymbol{e}_{n,k}$ and a polynomial $F(\mathbf{X}) \in \mathbb{F}_q[X_1, \dots, X_j]$, such that $\boldsymbol{e}_{n,k}$ is not asymptotically balanced over $\mathbb{F}_q$ and $F(\mathbf{X})$ is not balanced over $\mathbb{F}_q$, but $\boldsymbol{e}_{n,k} + F(\mathbf{X})$ is asymptotically balanced over $\mathbb{F}_q$.

Suppose that $\boldsymbol{e}_{n,k}$ was selected such that it is not asymptotically balanced over $\mathbb{F}_q$. Recall that

$$\mathrm{G}_{k;F}^{(q)}(X) = \frac{1}{q^j}\mathrm{G}_k^{(q)}(X)\mathrm{S}_{\mathbb{F}_q}(F;X). \tag{3.24}$$

Suppose that

$$\mathrm{G}_k^{(q)}(X) = \sum_{\beta\in\mathbb{F}_q} a_\beta X^\beta \text{ and } \frac{1}{q^j}\mathrm{S}_{\mathbb{F}_q}(F;X) = \sum_{\beta\in\mathbb{F}_q} b_\beta X^\beta, \tag{3.25}$$

where $0 < a_\beta, b_\beta < 1$ and $\sum_{\beta\in\mathbb{F}_q} a_\beta = \sum_{\beta\in\mathbb{F}_q} b_\beta = 1$ and that

$$\mathrm{G}_{k;F}^{(q)}(X) = \sum_{\beta\in\mathbb{F}_q} \frac{1}{q} X^\beta. \tag{3.26}$$

Observe that, by assumption on $\boldsymbol{e}_{n,k}$, not all a_β's are equal. Equation (3.24) can now be expressed as

$$\sum_{\beta\in\mathbb{F}_q} \frac{1}{q} X^\beta = \sum_{\beta\in\mathbb{F}_q} \left(\sum_{\gamma\in\mathbb{F}_q} a_{\beta-\gamma} b_\gamma \right) X^\beta, \tag{3.27}$$

which can be written in matrix form as

$$\frac{1}{q}\mathbf{1} = A_{q,k} \cdot \boldsymbol{b}, \tag{3.28}$$

where $\mathbf{1}$ and $\boldsymbol{b}$ are the column vectors whose entries are all 1's and all the b_β's (resp.), and $A_{q,k}$ is the $q \times q$ matrix $A_{q,k} = (a_{\beta-\gamma})_{\beta,\gamma}$. The problem now is to verify if a solution to (3.28) with $\boldsymbol{b} \neq (1/q)\mathbf{1}$ is possible.

Observe that $A_{q,k}$ is a doubly stochastic matrix. That means that $(1/q)\mathbf{1}$ is an eigenvector (corresponding to the eigenvalue $\lambda = 1$). It also implies that (see [18])

$$\lim_{N\to\infty} A_{q,k}^N = \frac{1}{q}J_q := \frac{1}{q}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}. \tag{3.29}$$

Suppose that $A_{q,k}$ also happens to be singular. Let $\boldsymbol{v}$ a non-trivial vector in the null space of $A_{q,k}$. Then $A_{q,k}^N\boldsymbol{v} = \mathbf{0}$ for every N and so

$$\mathbf{0} = \lim_{N\to\infty} A_{q,k}^N\boldsymbol{v} = \frac{1}{q}J_q\boldsymbol{v}, \tag{3.30}$$

which implies that $v_1 + \cdots + v_q = 0$. Now choose $\varepsilon > 0$ small enough such that all entries of

$$\frac{1}{q}\mathbf{1} + \varepsilon \boldsymbol{v} = \begin{pmatrix} 1/q + \varepsilon v_1 \\ 1/q + \varepsilon v_2 \\ \vdots \\ 1/q + \varepsilon v_q \end{pmatrix} \tag{3.31}$$

are positive. Observe that

$$\sum_{j=1}^{q} \left(\frac{1}{q} + \varepsilon v_j \right) = 1 + \varepsilon \sum_{j=1}^{q} v_j = 1,$$

which means that $(1/q)\mathbf{1} + \varepsilon \boldsymbol{v}$ is a probability vector different from $(1/q)\mathbf{1}$ that satisfies

$$A_{q,k} \left(\frac{1}{q}\mathbf{1} + \varepsilon \boldsymbol{v} \right) = \frac{1}{q} A_{q,k}\mathbf{1} + \varepsilon A_{q,k}\boldsymbol{v} = \frac{1}{q}\mathbf{1} + \mathbf{0} = \frac{1}{q}\mathbf{1}. \tag{3.32}$$

In other words, $(1/q)\mathbf{1} + \varepsilon \boldsymbol{v}$ is a probability vector different from $(1/q)\mathbf{1}$ that is a solution to (3.28).

To finish off the construction, choose an appropriate ε of the form $1/q^j$. Write

$$\frac{1}{q}\mathbf{1} + \varepsilon \boldsymbol{v} = \frac{1}{q^j} \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_q \end{pmatrix} \tag{3.33}$$

where $m_1 + \cdots + m_q = q^j$ and not all m_t's equal q^{j-1}. Label the finite field as $\mathbb{F}_q = \{\beta_1, \ldots, \beta_q\}$. Construct any function $\mathbb{F}_q^j \to \mathbb{F}_q$ such that in its output table (range) β_t appears m_t times. Let $F(X_1, \ldots, X_j)$ be the polynomial with coefficients in $\mathbb{F}_q$ that represents such function. The polynomial $F(X_1, \ldots, X_j)$ always exists and it is known as the *algebraic normal form* of the function [23]. Observe that

$$\mathrm{S}_{\mathbb{F}_q}(F; X) = \sum_{\beta \in \mathbb{F}_q} N_{\mathbb{F}_q}(F; \beta) X^{\beta} = \sum_{t=1}^{q} m_t X^{\beta_t}. \tag{3.34}$$

and so $F(\mathbf{X})$ is not balanced over $\mathbb{F}_q$. By assumption, $\boldsymbol{e}_{n,k}$ is not asymptotically balanced, but

$$\mathrm{G}_{k;F}^{(q)}(X) \quad = \quad \mathrm{G}_k^{(q)}(X) \cdot \frac{1}{q^j} \mathrm{S}_{\mathbb{F}_q}(F; X) = \sum_{\beta \in \mathbb{F}_q} \frac{1}{q} X^{\beta}$$

by construction of F. Therefore, $\boldsymbol{e}_{n,k}$ is not asymptotically balanced over $\mathbb{F}_q$, $F(\mathbf{X})$ is not balanced over $\mathbb{F}_q$, but $\boldsymbol{e}_{n,k} + F(\mathbf{X})$ is asymptotically balanced over $\mathbb{F}_q$.

***Remark* 3.6.** We know that Proposition 3.5 is true when $q = p$. Therefore, the construction will fail to produce a counterexample over $\mathbb{F}_p$. The step that fails is $A_{p,k}$ being singular. See, when p is prime, the matrix $A_{p,k}$ is not only doubly stochastic, but also a circulant matrix. Therefore, its determinant will be given by [14]

$$\det(A_{p,k}) = \prod_{t=0}^{p-1} \left(a_0 + a_{p-1}\omega_t + a_{p-2}\omega_t^2 + \cdots + a_1 \omega_t^{p-1} \right), \tag{3.35}$$

where $\omega_t = \exp(2\pi i t/p)$. But then $\det(A_{p,k}) = 0$ if and only if $a_0 = a_1 = \cdots = a_{p-1}$, i.e. if and only if $\boldsymbol{e}_{n,k}$ is asymptotically balanced over $\mathbb{F}_p$. However, $\boldsymbol{e}_{n,k}$ was specifically chosen to be not asymptotically balanced.

Example 3.7. Consider $\mathbb{F}_4 = \mathbb{F}_2(\alpha)$ with $\alpha^2 + \alpha + 1 = 0$. Select $\boldsymbol{e}_{n,3}$ and observe that

$$\mathrm{G}_3^{(4)} = \frac{5}{16} + \frac{5}{16}X + \frac{3}{16}X^{\alpha} + \frac{3}{16}X^{\alpha+1}. \tag{3.36}$$

For this particular example, the 4×4 matrix $A_{4,3}$ is given by

$$A_{4,3} = \begin{pmatrix} 5/16 & 5/16 & 3/16 & 3/16 \\ 5/16 & 5/16 & 3/16 & 3/16 \\ 3/16 & 3/16 & 5/16 & 5/16 \\ 3/16 & 3/16 & 5/16 & 5/16 \end{pmatrix}, \tag{3.37}$$

which is a singular 4×4 doubly stochastic matrix. Therefore $(1/4)\mathbf{1}$ is an eigenvector for $A_{4,3}$.

The null space of $A_{4,3}$ is spanned by the vectors

$$\boldsymbol{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{v}_2 = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}.$$

Observe that the entries of both vectors add up to 0, as predicted by the above discussion. Choose $\varepsilon = 1/4^2 = 1/16$. Then,

$$\frac{1}{4}\mathbf{1} + \varepsilon \boldsymbol{v}_1 = \begin{pmatrix} 3/16 \\ 5/16 \\ 1/4 \\ 1/4 \end{pmatrix} = \frac{1}{16}\begin{pmatrix} 3 \\ 5 \\ 4 \\ 4 \end{pmatrix}.$$

Now choose a polynomial in two variables over $\mathbb{F}_4$ such that it returns the value 0 three times, the value 1 five times, the value α four times and the value $\alpha+1$ four times. Such polynomials exist and

$$F(X_1, X_2) = X_1^3X_2^3 + X_2^2 + X_1^2$$

is an example.

Note that

$$\mathrm{S}_{\mathbb{F}_4}(F; X) = 3 + 5X + 4X^{\alpha} + 4X^{\alpha+1}$$

and

$$\begin{aligned} \mathrm{G}^{(4)}_{3;F}(X) &= \mathrm{G}^{(4)}_3(X) \cdot \frac{1}{16}\mathrm{S}_{\mathbb{F}_4}(F; X) \\ &= \left(\frac{5}{16} + \frac{5}{16}X + \frac{3}{16}X^{\alpha} + \frac{3}{16}X^{\alpha+1}\right)\left(\frac{3}{16} + \frac{5}{16}X + \frac{4}{14}X^{\alpha} + \frac{4}{16}X^{\alpha+1}\right) \\ &= \frac{1}{4} + \frac{1}{4}X + \frac{1}{4}X^{\alpha} + \frac{1}{4}X^{\alpha+1}. \end{aligned} \tag{3.38}$$

Therefore, $\boldsymbol{e}_{n,3} + F(\mathbf{X})$ is asymptotically balanced over $\mathbb{F}_4$ even though $\boldsymbol{e}_{n,3}$ is not asymptotically balanced and $F(\mathbf{X}) = X_1^3X_2^3 + X_2^2 + X_1^2$ is not balanced over $\mathbb{F}_4$. This proves that Proposition 3.5 is not true in general.

Example 3.8. With $\mathbb{F}_4$ as in the previous example. Consider $\boldsymbol{e}_{n,9}$ and observe that

$$\mathrm{G}^{(4)}_9(X) = \frac{45}{128} + \frac{29}{128}X + \frac{27}{128}X^{\alpha} + \frac{27}{128}X^{\alpha+1}. \tag{3.39}$$

That implies that $A_{4,9}$ is non-singular and so (3.28) has only the trivial solution. We conclude that a perturbation $\boldsymbol{e}_{n,9} + F(\mathbf{X})$ is asymptotically balanced if and only if $F(\mathbf{X})$ is balanced over $\mathbb{F}_4$.

4. Concluding remarks

In this article we studied the asymptotic behavior of exponential sums of elementary symmetric polynomials and their perturbations over finite fields. One of the purposes of doing so was to explore the veracity of Conjecture 2.2. We extended most of the results that appear in [7] to arbitrary finite fields. We also linked the asymptotic behavior of exponential sums of elementary symmetric polynomials and their perturbations to the value distribution of these polynomials over finite fields. The concept of asymptotically balanced symmetric polynomial (or perturbation) was also extended to general finite fields. In the particular case of a perturbation $\boldsymbol{e}_{n,k} + F(\mathbf{X})$, we showed that it is asymptotically balanced over $\mathbb{F}_p$ (p prime) if and only if $\boldsymbol{e}_{n,k}$ is asymptotically balanced or $F(\mathbf{X})$ is balanced over $\mathbb{F}_p$. We also show that this result does not hold in finite fields in general and provided a way to construct counterexamples.

The asymptotic behavior of the exponential sums considered in this work is dominated by a counting problem over a $q-1$-hypercube of length a power of p. Working on this problem over general finite fields can be difficult and counterintuitive. For example, consider $\boldsymbol{e}_3(\mathbf{X})$ in $\mathbb{F}_8 = \mathbb{F}_2(\alpha)$, with $\alpha^3 + \alpha + 1 = 0$. Then,

$$\mathrm{G}^{(8)}_3(X) = \sum_{\beta \in \mathbb{F}_8} \frac{1}{8}X^{\beta}, \tag{4.1}$$

i.e. $\mathbb{p}_3^{(8)}(\beta) = 1/8$ for every $\beta \in \mathbb{F}_8$. That is quite surprising given that for $\boldsymbol{e}_3(\mathbf{X})$ we have $\mathbb{p}_3^{(q)}(0) > \mathbb{p}_3^{(q)}(\beta)$, $\beta \neq 0$ for $\mathbb{F}_2$ and $\mathbb{F}_4$. Furthermore, $\boldsymbol{e}_3(\mathbf{X})$ is asymptotically balanced over $\mathbb{F}_8$, but in this case the degree of the elementary polynomial is not of the form $k = dp^l$ with $1 \leq d \leq p-1$. Moreover, for the first seven elementary symmetric polynomials, i.e. for $1 \leq k \leq 7$, we have

$$\mathrm{G}_k^{(8)}(X) = \begin{cases} \sum_{\beta\in\mathbb{F}_8} \frac{1}{8} X^\beta & k \neq 5,7 \\ \frac{71}{512} + \sum_{\beta\in\mathbb{F}_8^\times} \frac{63}{512} X^\beta & k = 5 \\ \frac{67}{512} + \frac{67}{512} X + \sum_{\beta\in\mathbb{F}_8\setminus\mathbb{F}_2} \frac{63}{512} X^\beta & k = 7. \end{cases} \tag{4.2}$$

This example provides further evidence about the difficulty to determine the veracity of the generalized conjecture of Cusick, Li and Stănică.

Acknowledgments. The authors would like to thank Oscar E. González for reading a previous version of this article.

References

[1] O. Aberth. The elementary symmetric functions in a finite field of primer order. *Illinois J. Math.* **8(1)** (1964), 132–138.
[2] R. A. Arce-Nazario, F. N. Castro, O. E. González, L. A. Medina, and I. M. Rubio. New families of balanced symmetric functions and a generalization of Cuscik, Li and P. Stănică. *Designs, Codes and Cryptography* **86** (2018), 693–701.
[3] J. Cai, F. Green and T. Thierauf. On the correlation of symmetric functions. *Math. Systems Theory* **29** (1996) 245–258.
[4] A. Canteaut and M. Videau. Symmetric Boolean Functions. *IEEE Trans. Inf. Theory* **51(8)** (2005) 2791–2881.
[5] F. N. Castro, O. E. González, and L. A. Medina. Diophantine equations with binomial coefficients and perturbations of symmetric Boolean functions. *IEEE Trans. Inf. Theory* **64(2)** (2018) 1347–1360.
[6] F. N. Castro and L. A. Medina. Linear Recurrences and Asymptotic Behavior of Exponential Sums of Symmetric Boolean Functions. *Elec. J. Combinatorics* **18** (2011) #P8.
[7] F. N. Castro and L. A. Medina. Asymptotic Behavior of Perturbations of Symmetric Functions. *Annals of Combinatorics* **18** (2014) 397–417.
[8] F. N. Castro and L. A. Medina. Modular periodicity of exponential sums of symmetric Boolean functions. *Discrete Appl. Math.* **217** (2017) 455–473.
[9] F. N. Castro, L. A. Medina, and P. Stănică. Generalized Walsh transforms of symmetric and rotation symmetric Boolean functions are linear recurrent. *Appl. Algebra Eng. Commun. Comput.* **29(5)** (2018) 433–453.
[10] F. N. Castro, R. Chapman, L. A. Medina, and L. B. Sepúlveda. Recursions associated to trapezoid, symmetric and rotation symmetric functions over Galois fields. *Discrete Mathematics*, **341(7)** (2018) 1915–1931.
[11] F. N. Castro, L. A. Medina, and L. B. Sepúlveda. Closed formulas for exponential sums of symmetric polynomials over Galois fields *J. Algebr. Comb.* **50(1)** (2019) 73-98.
[12] T. W. Cusick. Hamming weights of symmetric Boolean functions. *Discrete Appl. Math.* **215** (2016) 14–19.
[13] T. W. Cusick, Y. Li, and P. Stănică. Balanced Symmetric Functions over $GF(p)$. *IEEE Trans. Inf. Theory* **54 (3)** (2008) 1304–1307.
[14] Philip Davis. Circulant Matrices. Chelsea publishing, Second Edition,1994.
[15] N. J. Fine. On the asymptotic distribution of the elementary symmetric functions (mod p). *Trans. Amer. Math. Soc.* 69(1) (1950), 109–129.
[16] G. Gao, Y. Guo, and Y. Zhao. Recent Results on Balanced Symmetric Boolean Functions. *IEEE Trans. Inf. Theory* **62 (9)** (2016) 5199–5203.
[17] Y. Hu and G. Xiao. Resilient Functions Over Finite Fields. *IEEE Trans. Inf. Theory* **49** (2003) 2040–2046.
[18] O. Ibe. Markov Processes for Stochastic Modeling (Elsevier Insights), Second Edition (2013), Elsevier, Boston, MA.
[19] Y. Li and T. W. Cusick. Linear Structures of Symmetric Functions over Finite Fields. Inf. Processing Letters **97** (2006) 124–127.
[20] Y. Li and T. W. Cusick. Strict Avalanche Criterion Over Finite Fields. *J. Math. Cryptology* **1(1)** (2007) 65–78.
[21] M. Liu, P. Lu and G.L. Mullen. Correlation-Immune Functions over Finite Fields. *IEEE Trans. Inf. Theory* **44** (1998), 1273–1276.
[22] C. Mitchell. Enumerating Boolean functions of cryptographic significance. *J. Cryptology* **2(3)** (1990) 155–170.
[23] V. S. Pless, W. C. Huffman, R. A. Brualdi, Eds. Handbook of Coding Theory, (1998) Elsevier, Amsterdam, the Netherlands.
[24] J. D. Smith. Probability and the elementary symmetric functions. *Proc. Camb. Phil. Soc.* **74** (1973) 133–139.

Department of Mathematics, University of Puerto Rico, 17 Ave. Universidad STE 1701, San Juan, PR 00925
Email address: `luis.medina17@upr.edu`

Department of Mathematics, University of Puerto Rico, 17 Ave. Universidad STE 1701, San Juan, PR 00925
Email address: `leonid.sepulveda1@upr.edu`

Department of Mathematics, University of Puerto Rico, 17 Ave. Universidad STE 1701, San Juan, PR 00925
Email address: `cesar.serna@upr.edu`