AAECC (2008) 19:229–239
DOI 10.1007/s00200-008-0075-z

Asymmetric information embedding

Amin Shokrollahi

Received: 5 June 2007 / Revised: 27 December 2007 / Published online: 11 April 2008

Abstract In the Information Embedding Problem one is given a piece of data which can be altered only conditionally, for example only at certain places. One is then asked to embed an arbitrary message into the data by only applying admissible changes to the data. These changes lead to a distortion which is to be kept low. In this short note, we introduce an "asymmetric" version of information embedding in which the file is regarded as a string over a finite alphabet, and admissible changes on the alphabet elements are modeled by a directed graph. We introduce embedding techniques based on list-decoding algorithms for algebraic–geometric codes, and analyze their performance.

1 Introduction

Let $G = (V, E)$ be a directed graph in which for all $v \in V$ we have $(v, v) \in E$, and let $\hat{G}$ be its transitive closure. An example is the *q-line* consisting of q nodes $0, 1, \ldots, q-1$ such that there is an edge between i and $i+1$ for $i = 0, \ldots, q-2$. Figure 1 gives an example for the case $q = 5$.

Let $\mathbf{P}(S)$ denote the set of subsets of a set S. An Information Embedding Scheme (IES) of signature $(n, k; G)$, denoted IES$(n, k; G)$, consists of a pair $(\mathbb{E}, \mathbb{D})$ of polynomial time computable functions $\mathbb{E}\colon \mathbb{F}_q^n \times \mathbb{F}_q^k \to \mathbb{F}_q^n$ and $\mathbb{D}\colon \mathbb{F}_q^n \to \mathbf{P}(\mathbb{F}_q^k)$ called *encoding* and *decoding* functions, such that:

1. For all $v \in \mathbb{F}_q^k$ and $x \in \mathbb{F}_q^n$ and all $i = 1, \ldots, n$ we have that $(y_i, x_i) \in E$, where $y = \mathbb{E}(x, v)$.

A. Shokrollahi (✉)
EPFL, Lausanne, Switzerland
e-mail: amin.shokrollahi@epfl.ch

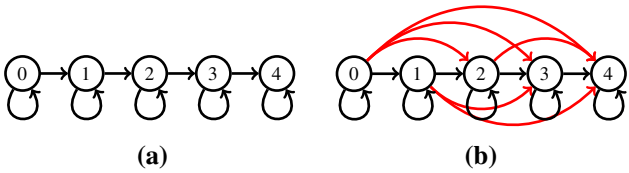

Fig. 1 (**a**) The 5-line and (**b**) its transitive closure

2. There exist $\delta, \gamma > 0$ (depending possibly on q but not on n or k) such that if $v \in \mathbb{F}_q^k$ and $x \in \mathbb{F}_q^n$ are chosen uniformly at random, then $\Pr[v \notin \mathbb{D}(\mathbb{E}(v, x))] \leq \gamma \exp(-n\delta)$.

We call the quantity k/n the "rate" of IES$(n, k; G)$. We are interested in IESs of high rate. More precisely, we would like to find the value of $A(G)$, where

$$A(G) := \sup\{R \mid \exists \text{ Information Embedding scheme of rate } R \text{ for } G\}.$$

It is possible that for some graphs there is no IES. For example, suppose that the graph has only self-loops and no edges between the nodes. Then, the only encoding function is the one mapping (v, x) to x, regardless of v. As a result, a given x does not convey any information about v; the decoder $\mathbb{D}$ can only guess the value of v, with error $1 - 1/2^k$. This error does not decay polynomially in n, hence there is no matching decoding function for $\mathbb{E}$.

In this paper, we will always assume that the vertex set V of the graph G is a set with q elements, and denote this set either by $\{0, \ldots, q-1\}$ (as in the case of the q-line) or by the field $\mathbb{F}_q$. Moreover, our encoding function $\mathbb{E}$ is composed of two functions: an encoding function φ which maps v to an element $c = \varphi(v)$ of a suitably chosen linear $[n, k]_q$-code C, and a map π which maps the pair (x, c) to the vector $(y_1, \ldots, y_n)$, where

$$y_i = \begin{cases} x_i & \text{if } (c_i, x_i) \notin E \\ c_i & \text{else.} \end{cases}$$

In other words, whenever we are allowed to replace x_i with c_i, we will do so opportunistically. Whether this scheme is successful or not depends on the statistics of x, and that of c. For example, suppose that the underlying graph is the q-line, and that $c = (q-1, q-1, \ldots, q-1)$. Then $\mathbb{E}(x, c) = x$ for all x, and hence $\mathbb{E}(x, c)$ conveys essentially no information about c.

We will therefore assume that we know something about the statistics of c. More precisely, we assume that we know the *signature* of c, defined as the real vector $(s_\beta \mid \beta \in \mathbb{F}_q)$ such that

$$s_\beta = \frac{\#\{i \mid c_i = \beta\}}{n}.$$

Since in our definition of the IES we assume v to be chosen uniformly at random, it follows that the vector $c = \varphi(v)$ is chosen uniformly at random from C. The following theorem which will be proved in Sect. 2 shows that the signature of c is close to uniform.

Theorem 1 *Suppose that C is an $[n, k]_q$-code with dual distance d. Then, for any $\varepsilon > 0$ we have*

$$\Pr\left[\left|\operatorname{sign}(c) - \frac{1}{q}(1, 1, \ldots, 1)\right|_1 > \varepsilon\right] \leq 2q \left\lceil \frac{n}{d-1} \right\rceil \exp\left(-\frac{(d-1)^3\varepsilon^2}{4qn^2}\right),$$

where $\operatorname{sign}(c)$ *is the signature of c, the probability is with respect to the uniform random choice of $c \in C$, and $|x|_1$ is the* 1*-norm of the vector x.*

If the signature of c is not badly chosen, then $\mathbb{E}(x, c)$ conveys some information about c. For example, we are able to determine the statistics of the number of positions in which $\mathbb{E}(x, c)$ and c coincide. Using this statistics, the Guruswami–Sudan list-decoding algorithm [2], and sequences of asymptotically good AG-codes [5,1], we will be able to obtain good decoding algorithms for the proposed encoding algorithm. All this will be done in Sect. 3 along with the proof of the following theorem and its corollary.

Theorem 2 *Suppose that c has signature $(s_\beta \mid \beta \in \mathbb{F}_q)$, and suppose further that q is the square of a prime power. Then, for any $\varepsilon > 0$ there exists n_0 such that for all $n \geq n_0$, if*

$$\frac{k}{n} < \frac{1}{q^2}\left(\sum_{\beta \in \mathbb{F}_q} s_\beta \operatorname{outdeg}(\beta)\right)^2 - \frac{1}{\sqrt{q}-1} - \varepsilon,$$

then there is an explicit $\operatorname{IES}(n, k; G)$*, where* $\operatorname{outdeg}(\beta)$ *is the out-degree of β in the graph G.*

Corollary 1 *Let $G = (\mathbb{F}_q, E)$ be a directed graph, and q be a square of a prime power. Then*

$$A(G) \geq \frac{|E|^2}{q^4} - \frac{1}{\sqrt{q}-1}.$$

The theorem is rather "generic", in the sense that the particular structure of the underlying graph has not been taken into account (except of course the number of edges.) A more refined approach is obtained by considering that different entries of $\mathbb{E}(x, v)$ may provide "soft" information about the entries of v. For example, consider again the case of the q-line, and assume that one of the entries, say the ith, of $\mathbb{E}(x, v)$ is 0. Then, if $c_i \neq 0$, then x_i has to be 0, where $c = \varphi(v)$. Since x_i is zero only with probability $1/q$ (because of the uniform random choice of x), and under the assumption that c is chosen uniformly at random from the code, it follows that the

probability that $c_i = 0$ is $1 - 1/q^2 + 1/q$, which is much larger than the probability that $c_i \neq 0$.

This approach and a close-to-optimal assignment of the multiplicities in the Guruswami–Sudan algorithm yield the following more refined result, which will be proved in Sect. 4, along with its corollary.

Theorem 3 *Suppose that c has signature $(s_\beta \mid \beta \in \mathbb{F}_q)$ and that this signature is known to the decoder. Further, suppose that q is a square of a prime power. Then for any $\varepsilon > 0$ there exists n_0 such that for all $n \geq n_0$, if*

$$\frac{k}{n} < \frac{1}{q^2} \sum_{\beta \in \mathbb{F}_q} \frac{s_\beta^2}{\tau_\beta} \mathrm{outdeg}(\beta)^2 - \frac{1}{\sqrt{q} - 1},$$

then there is an explicit $\mathrm{IES}(n, k; G)$*, where*

$$\tau_\beta := \frac{1}{q} \left(s_\beta \mathrm{outdeg}(\beta) + \sum_{(\gamma, \beta) \notin E} s_\gamma \right).$$

Corollary 2 *Let $G = (\mathbb{F}_q, E)$ be a directed graph, and suppose that q is the square of a prime power. Then*

$$A(G) \geq \frac{1}{q^2} \sum_{\beta \in \mathbb{F}_q} \frac{\mathrm{outdeg}(\beta)^2}{q + \mathrm{outdeg}(\beta) - \mathrm{indeg}(\beta)} - \frac{1}{\sqrt{q} - 1},$$

where $\mathrm{indeg}(\beta)$ *is the in-degree of the node β.*

It is not immediately clear that Corollary 2 is stronger than Corollary 1 (though this will be clear after the proofs are presented). We illustrate the difference in strength with the following example: suppose that G is the transitive closure of the q-line. The number of edges in G equals $\sum_{i=0}^{q-1} (q - i) = q(q+1)/2$. Therefore, the bound of Corollary 1 is

$$A(G) \geq \frac{(q+1)^2}{4q^2} - \frac{1}{\sqrt{q} - 1} = \frac{1}{4} - \frac{1}{\sqrt{q} - 1} + \frac{1}{2q} + \frac{1}{4q^2}.$$

The out-degree of $i, i \in \{0, \ldots, q-1\}$ is $q - i$, while its in-degree is $i + 1$. Therefore,

$$\sum_{i=0}^{q-1} \frac{\mathrm{outdeg}(i)^2}{q + \mathrm{outdeg}(i) - \mathrm{indeg}(i)} = \frac{1}{2} \sum_{i=0}^{q-1} \frac{(q-i)^2}{q - i - 0.5} \geq q^2 \left(\frac{1}{4} + \frac{1}{2q} + \frac{\ln(q)}{8q^2} \right).$$

The bound of Corollary 2 implies

$$A(G) \geq \frac{1}{4} - \frac{1}{\sqrt{q}-1} + \frac{1}{2q} + \frac{\ln(q)}{8q^2}.$$

This shows that Corollary 2 is (in this case marginally) better than Corollary 1.

This note is devoted to a theoretical analysis of information hiding, and no claim is made as to whether the introduced asymmetric version has practical applications. In fact, we are aware only of very contrived applications at this point. It would be interesting to explore possible applications of such asymmetric embedding schemes.

2 Proof of Theorem 1

We start with a well-known general Chernoff type bound:

Proposition 1 *Suppose that $T_1, \ldots, T_m$ are i.i.d. random variables over $\{0, 1\}$, and that $\Pr[T_i = 1] = p > 0$. Further, let $T := \sum_{i=1}^{m} T_i$. Then for any positive $\delta < p/2$ we have*

$$\Pr[|T - mp| \geq \delta m] \leq 2 \exp\left(-\frac{m\delta^2}{4p}\right).$$

Proof By [4, Theorem 4.1] we have for any $\varepsilon > 0$

$$\Pr[T > (1+\varepsilon)mp] \leq \left(\frac{\mathrm{e}^{\varepsilon}}{(1+\varepsilon)^{1+\varepsilon}}\right)^{mp}.$$

Note that $(1+\varepsilon)^{1/\varepsilon} \geq \mathrm{e}^{1-\varepsilon/2}$, so that $(1+\varepsilon)^{1+\varepsilon} \geq \mathrm{e}^{\varepsilon+\varepsilon^2(1-\varepsilon)/2}$. Therefore, for $\varepsilon > 1/2$ we have

$$\Pr[T > (1+\varepsilon)mp] \leq \exp\left(\frac{-\varepsilon^2(1-\varepsilon)mp}{2}\right) < \exp\left(\frac{-\varepsilon^2 mp}{4}\right).$$

Setting $\varepsilon = \delta/p$, we see that

$$\Pr[T > (1+\varepsilon)mp] \leq \exp\left(-\frac{m\delta^2}{4p}\right).$$

On the other hand, by [4, Theorem 4.2] we have

$$\Pr[T < (1-\varepsilon)mp] \leq \exp\left(-\frac{mp\varepsilon^2}{2}\right) < \exp\left(-\frac{mp\varepsilon^2}{4}\right).$$

A union bound on the two assertions yields the result. □

The next (also well-known) result shows that if c is a word chosen uniformly at random from C and C has dual distance d, then any set of $d-1$ coordinates of c are independent random variables.

Proposition 2 *Suppose that C is an $[n,k]_q$-code of dual distance d, and let $I \subset \{1,\ldots,n\}$ be any subset of size $d-1$. Further, let c be chosen uniformly at random from C. Then the random variables c_i, $i \in I$, are independent.*

Proof Let the $k \times n$-matrix $\mathcal{G}$ be a generator matrix for C, and let H be the matrix obtained by selecting the columns corresponding to I from $\mathcal{G}$. Then $(c_i \mid i \in I) = (X_1,\ldots,X_k)\cdot H$, where $X_1,\ldots,X_k$ are i.i.d. uniformly distributed random variables over $\mathbb{F}_q$. Since C has dual distance d, the matrix H has rank $d-1$, hence the c_i, $i \in I$, are independent. □

The proof of Theorem 1 proceeds as follows. We subdivide the coordinate positions of a codeword c into $\lceil n/(d-1)\rceil$ disjoint subsets of which all (but at most one) have size $d-1$. Using Chernoff bounds, we prove that if c is chosen uniformly at random, then on each of these subsets the signature of c is close to uniform with very high probability. We then apply a union bound to prove that the signature of the entire vector c is close to uniform with high probability.

To this end, we introduce some notation. For simplicity, we assume that n is divisible by $d-1$. Obvious and easy modifications are necessary if this is not the case. Let $I_j := \{j(d-1), j(d-1)+1,\ldots,(j+1)(d-1)-1\}$ for $0 \le j < n/(d-1)$. For a codeword $c \in C$ let c^j denote the vector obtained by projecting c into the coordinate positions given by I_j. Then $\mathrm{sign}(c^j)$ is the signature of the $(d-1)$-dimensional vector c^j. The first proposition we prove shows that this signature vector is close to uniform.

Proposition 3 *Notation as above, we have for all $0 \le j < n/(d-1)$ and all $\varepsilon > 0$:*

$$\Pr\left[\left|\mathrm{sign}(c^j) - \frac{1}{q}(1,1,\ldots,1)\right|_1 > \varepsilon\right] < 2q\exp\left(-\frac{(d-1)\varepsilon^2}{4q}\right).$$

Proof Let $\beta \in \mathbb{F}_q$ and $T_{i,\beta}$, $i \in I_j$, be i.i.d. random variables over $\{0,1\}$ with $\Pr[T_{i,\beta}=1] = 1/q$. Let $(\sigma_\beta \mid \beta \in \mathbb{F}_q)$ denote the signature of c^j. Then $(d-1)\sigma_\beta = \sum_{i\in i_j} T_{i,\beta}$, and by Proposition 1 we have

$$\Pr[|\sigma_\beta - 1/q| > \varepsilon] \le 2\exp\left(-\frac{(d-1)q\varepsilon^2}{4}\right).$$

Therefore,

$$\begin{aligned}\Pr\left[\left|\mathrm{sign}(c^j) - \frac{1}{q}(1,1,\ldots,1)\right|_1 > \varepsilon\right] &\le \sum_{\beta\in\mathbb{F}_q}\Pr\left[\left|\sigma_\beta - \frac{1}{q}\right| > \frac{\varepsilon}{q}\right]\\ &\le 2q\exp\left(-\frac{(d-1)\varepsilon^2}{4q}\right),\end{aligned}$$

which proves the assertion. □

Proof of Theorem 1 Note that by the previous proposition

$$\begin{aligned}
&\Pr\left[\left|\operatorname{sign}(c) - \frac{1}{q}(1,1,\ldots,1)\right|_1 > \varepsilon\right] \\
&\leq \sum_{j=0}^{n/(d-1)-1} \Pr\left[\left|\operatorname{sign}(c^j) - \frac{1}{q}(1,1,\ldots,1)\right|_1 > \frac{d-1}{n}\varepsilon\right] \\
&\leq 2q\frac{n}{d-1}\exp\left(-\frac{(d-1)^3\varepsilon^2}{4qn^2}\right),
\end{aligned}$$

which is equivalent to the assertion of the theorem when $d-1$ divides n. If not, then obvious modifications to the proof will yield the result. □

3 Proof of Theorem 2

As mentioned in Sect. 1, we are going to use the list-decoding algorithm of Guruswami–Sudan. Suppose that C is an AG-code constructed from the irreducible nonsingular curve X of genus g over $\mathbb{F}_q$ using the divisor αQ, where $Q \in X(\mathbb{F}_q)$. We will use the same multiplicity, r, for all the coordinate positions. We regard $\mathbb{E}(x,c)$ as a corrupted version of the vector c. To apply the list decoding algorithm, we denote by t the number of agreements of $\mathbb{E}(x,c)$ with c. Then, as long as

$$(rt - 1 - g)^2 \geq n\alpha(r+1)^2,$$

we can list-decode $\mathbb{E}(x,c)$ and obtain a short list which will contain c. (See the algorithm on p 1766 of [2].) The condition is equivalent to

$$\frac{\alpha}{n} \leq \left(\frac{r}{r+1}\frac{t}{n} - \frac{g+1}{nr}\right)^2.$$

By choosing r large enough, this condition translates to

$$\frac{\alpha}{n} < \left(\frac{t}{n}\right)^2,$$

or

$$\frac{k}{n} < \left(\frac{t}{n}\right)^2 - \frac{g}{n}.$$

Recall from [5,1] that for q a perfect square, there are sequences of algebraic curves over $\mathbb{F}_q$ for which the ratio between the genus g and the number of $\mathbb{F}_q$-rational points N of the curve converges to the value $1/(\sqrt{q}-1)$. By choosing AG-codes from such curves, the second term on the right hand side can be made arbitrarily close to $1/(\sqrt{q}-1)$.

It suffices now to show that with very high probability

$$\left| \frac{t}{n} - \frac{1}{q} \left(s_\beta \text{outdeg}(\beta) + \sum_{(\gamma,\beta)\notin E} s_\gamma \right) \right| < \delta$$

for any $\delta > 0$ if n is large enough. Fix $\beta \in \mathbb{F}_q$. We will count the number A_β of agreements between $\mathbb{E}(x, c)$ and c on all the positions i for which $c_i = \beta$. Note that if a is chosen uniformly at random from $\mathbb{F}_q$, then $\Pr[(\beta, a) \in E] = \text{outdeg}(\beta)/q =: p$. Let $T_1, \ldots, T_{s_\beta n}$ be i.i.d. random variables on $\{0, 1\}$ such that $\Pr[T_i = 1] = p$. Then $A_\beta = \sum_i T_i$, and Proposition 1 implies that

$$\Pr\left[\left| A_\beta - s_\beta n \frac{\text{outdeg}(\beta)}{q} \right| > \varepsilon n \right] < 2 \exp\left(-\frac{s_\beta n q \varepsilon^2}{\text{outdeg}(\beta)} \right).$$

The number of agreements between $\mathbb{E}(x, c)$ and c equals $\sum_\beta A_\beta$. We have

$$\begin{aligned} &\Pr\left[\left| \sum_\beta A_\beta - \frac{n}{q} \sum_\beta s_\beta \text{outdeg}(\beta) \right| > \varepsilon n \right] \\ &\le \sum_\beta \Pr\left[\left| A_\beta - s_\beta n \frac{\text{outdeg}(\beta)}{q} \right| > \frac{\varepsilon n}{q} \right] \\ &\le 2q \exp\left(-\frac{s_\beta n \varepsilon^2}{q \text{outdeg}(\beta)} \right). \end{aligned}$$

It follows that t/n is sharply concentrated around its mean value which is $\sum_\beta s_\beta$ $\text{outdeg}(\beta)/q$, and the result of the theorem follows.

Proof of Corollary 1 The AG-code C used in the proof of Theorem 2 is the image of the linear space of αQ under evaluation on points $P_1, \ldots, P_n$ of $X(\mathbb{F}_q)$. The dual distance of this code is at least $k - g + 1$, where k is the dimension of the code C, and g is the genus of the curve X. Since we are using sequences of optimal AG-codes, and since the rate of the original code is fixed, it follows that for any positive $\varepsilon > 0$ the signature of a randomly chosen codeword in the code is ε-close to the uniform signature, with very high probability. As a result, by making ε small enough, the upper bound for the rate in Theorem 2 can be made arbitrarily close to

$$\frac{1}{q^2} \left(\sum_\beta \frac{1}{q} \text{outdeg}(\beta) \right)^2 - \frac{1}{\sqrt{q} - 1} = \frac{|E|^2}{q^4} - \frac{1}{\sqrt{q} - 1},$$

since $\sum_\beta \text{outdeg}(\beta)$ is the number of edges in the graph G. □

4 Proof of Theorem 3

The idea of the proof is to assign different multiplicities to different symbols of $\mathbb{E}(x, c)$. In that respect, the approach is somewhat similar to that of Kötter and Vardy [3]. We will assign higher multiplicities to regions of $\mathbb{E}(x, c)$ with lower fidelity, and lower multiplicities to regions with higher fidelity. By optimizing the multiplicity assignments we will be able to obtain the best upper bound attainable by this method.

More precisely, let $y := \mathbb{E}(x, c)$, and let t_β be the number of positions i such that $y_i = c_i$, given that $y_i = \beta$. Furthermore, let $\mu_\beta n$ be the number of positions i such that $y_i = \beta$. Then, using multiplicity r_β whenever $y_i = \beta$, we obtain the following bounds for successful list decoding of the vector $\mathbb{E}(x, c)$:

$$\sum_{\beta \in \mathbb{F}_q} t_\beta r_\beta > \ell$$

$$(\ell - g)^2 > 2n\alpha \sum_{\beta \in \mathbb{F}_q} \mu_\beta \binom{r_\beta + 1}{2}$$

where g is the genus of the curve X. Choosing

$$r_\beta := r \frac{t_\beta}{n\mu_\beta} - 1,$$

we obtain the inequality

$$\sqrt{\frac{\alpha}{n}} \leq \frac{\sum_\beta \frac{t_\beta}{n} \left(\frac{t_\beta}{n\mu_\beta} - \frac{1}{r} \right)}{\sqrt{\sum_\beta \frac{t_\beta^2}{n^2 \mu_\beta}}} - \frac{g+1}{nr\sqrt{\sum_\beta \frac{t_\beta^2}{n^2\mu_\beta}}}.$$

Fix $\varepsilon > 0$. By choosing r large enough, we can assume that

$$\frac{\alpha}{n} \leq \sum_\beta \frac{t_\beta^2}{n^2 \mu_\beta} - \varepsilon.$$

(A more detailed analysis reveals that $r = O(1/\varepsilon)$; since the decoding algorithm is polynomial in r, we need to make sure that ε is not too small in order to keep the running time polynomial.)

To show the assertion of Theorem 3, it suffices to show that, with high probability for any $\delta > 0$

$$\left| \frac{t_\beta}{n} - \frac{s_\beta \text{outdeg}(\beta)}{q} \right| < \delta \tag{1}$$

$$\left| \mu_\beta - \frac{1}{q} \left(s_\beta \mathrm{outdeg}(\beta) + \sum_{(\gamma,\beta)\notin E} s_\gamma \right) \right| < \delta. \tag{2}$$

To this end, let $y := \mathbb{E}(x, c)$, $\gamma, \beta \in \mathbb{F}_q$, and let

$$T_{\beta,\gamma} := \#\{i \mid c_i = \gamma \ \&\ y_i = \beta\}.$$

Then we have

$$t_\beta = T_{\beta,\beta}$$

$$\mu_\beta = \frac{1}{n} \sum_{\gamma\in\mathbb{F}_q} T_{\beta,\gamma}.$$

Let $p_{\beta,\gamma} := \Pr[y_i = \beta \mid c_i = \gamma]$, and let $T_1, \ldots, T_{s_\gamma n}$ be i.i.d. random variables on $\{0, 1\}$ with $\Pr[T_\ell = 1] = p_{\beta,\gamma}$. Then $T_{\beta,\gamma} = \sum_\ell T_\ell$, and by The Chernoff bound of Proposition 1

$$\Pr\left[|T_{\beta,\gamma} - s_\gamma n p_{\beta,\gamma}| > \delta s_\gamma n\right] \le 2\exp\left(-\frac{s_\gamma n\delta^2}{2p_{\beta,\gamma}}\right).$$

Assertions (1) and (2) would therefore follow from

$$p_{\beta,\beta} = \frac{\mathrm{outdeg}(\beta)}{q} \tag{3}$$

$$p_{\beta,\gamma} = \frac{1}{q} \text{ for } (\gamma, \beta) \notin E \tag{4}$$

$$p_{\beta,\gamma} = 0 \text{ for } \gamma \neq \beta, (\gamma, \beta) \in E. \tag{5}$$

All these assertions follow from the fact that the vector x has been chosen uniformly at random: $\Pr[y_i = \beta \mid c_i = \beta]$ is the probability that x_i is chosen so that $(\beta, x_i) \in E$, which is equal to $\mathrm{outdeg}(\beta)/q$. The probability $\Pr[y_i = \beta \mid c_i = \gamma]$ is clearly 0 if $(\gamma, \beta) \in E$, and $\gamma \neq \beta$, because of the way our encoding works. The same probability is $1/q$ if $(\gamma, \beta) \notin E$. This concludes the proof of Theorem 3.

Proof of Corollary 2 The proof follows the same lines as that of Corollary 1. Since c is chosen uniformly from C, its signature is very close to the uniform one. Plugging the uniform signature into the formula of Theorem 3 yields the result. □

Acknowledgments The author is indebted to Rüdiger Urbanke for his help in proving Theorem 1.

References

1. Garcia, A., Stichtenoth, H.: A tower of Artin–Schreier extensions of function fields attaining the Drinfeld-Vladut bound. Invent. Math. **121**, 211–222 (1995)

2. Guruswami, V., Sudan, M.: Improved decoding of Reed–Solomon and algebraic–geometric codes. IEEE Trans. Inform. Theor **45**, 1757–1767 (1999)
3. Kötter, R., Vardy, A.: Algebraic soft decision decoding of Reed–Solomon codes. IEEE Trans. Inform. Theor. **49**, 2809–2825 (2003)
4. Motwani, R., Raghavan, P.: Randomized Algorithms. Cambridge University Press, Cambridge (1995)
5. Tsfasman, M., Vladut, S., Zink, T.: Modular curves, Shimura curves, and Goppa codes better than the Varshamov–Gilbert bound. Math. Nachrichten **109**, 21–28 (1982)