VERY PRELIMINARY VERSION

FROBENIUS DISTRIBUTIONS AND GALOIS REPRESENTATIONS – NUMERICAL EVIDENCE

Kevin James and V. Kumar Murty[1]

1. Introduction

The purpose of this article is to study one of the conjectures made in [7] from a computational and statistical point of view. Given a modular form $f(z) = \sum_{n\geq 0} a_f(n)q^n \quad (q = e^{2\pi iz})$ and a number a, the conjecture deals with the number $\pi_f(x, a)$ of primes p for which the p-th Fourier coefficient $a_f(p)$ of f is equal to a. As this work is likely to be of interest to researchers in arithmetic, computational mathematics and statistics, we have attempted to write the article in a manner that will be comprehensible to those who have a background in only one of these fields.

The remainder of the paper is laid out as follows. We give a precise statement of the conjectures concerning the asymptotic behavior or $\pi_f(x, a)$ and a similar function $\theta_f(x, a)$ as $x \to \infty$ in section 2. In section 3 we give a discussion of the error terms arising in these asymptotics and, we discuss the calculation of the relevant constants in section 4. In sections 5 and 6 we discuss the computations involved and give the results of these computation in tabular and graphic form. Finally, in section 7 we discuss a statistical analysis of the data and the implications of this analysis.

2. The conjectures

Let f be a normalized eigenform of weight k, level N and character ϵ. For a prime p, write $a_f(p)$ for the p-th Fourier coefficient of f. Let O_f denote the $\mathbb{Z}$ module generated by the $a_f(p)$, and let E_f denote the quotient field. Let $F = F_f$ denote the stable trace field. Let $a \in F_f$ be generic in the sense of [7], Definition (2.14). Then [7], Conjecture (3.1) is the following assertion.

Conjecture 2.1. *Let $a \in O_F$ and assume that f is not of CM-type. Set*

$$\pi_{1/2}(x) \;=\; \sum_{p\leq x} \frac{1}{2\sqrt{p}}.$$

Then, there is a constant $c_{f,a}$ such that

$$\pi_f(x, a) \;=\; (c_{f,a} + \mathbf{o}(1)) \begin{cases} \pi_{1/2}(x) & \text{if } k = 2 \text{ and } F = \mathbb{Q} \\ \log\log x & \text{if } (k, [F : \mathbb{Q}]) = (2, 2) \text{ or } (3, 1) \\ 1 & \text{otherwise.} \end{cases}$$

[1]Research partially supported by a grant from NSERC

If we consider the heuristics that led to (2.1), we may even ask the following stronger question. Define

$$\theta_f(x,a) = \sum_{\substack{p\le x\\ a_f(p)=a}} \log p.$$

As we expect the probability that $a_f(p)=a$ to be proportional to $p^{(1-k)/2}$, we might expect that

$$\theta_f(x,a) \ \sim \ c_{f,a}\int_1^x t^{(1-k)/2}d\theta(t)$$

where we have set as usual

$$\theta(x) = \sum_{p\le x}\log p.$$

Assuming the Riemann Hypothesis, we have

$$\theta(x) \ = \ x \ + \ \mathbf{O}(x^{\frac{1}{2}+\epsilon}).$$

Thus, we might raise the following question.

Question 2.1. *Letting*

$$\theta_{1/2}(x) = \sum_{p\le x}\frac{\log p}{2\sqrt{p}}$$

and with notation as above, is it true that

$$\theta_f(x,a) = \begin{cases} c_{f,a}\theta_{1/2}(x) + \mathbf{O}(x^\epsilon) & \text{if } k=2\\ c_{f,a}\log x + \mathbf{O}(1) & \text{if } k=3\\ c_{f,a} & \text{if } k\ge 4? \end{cases}$$

3. The error term

On probabilistic grounds, it is possible to speculate on a more precise error term in the above. Consider a sequence of Bernoulli trials indexed by the primes $p\le x$. Fix a value a and consider the p-th trial a success if $a_f(p)=a$. Alternately, let X_p be the random variable which takes the value 1 if $a_f(p)=a$ and 0 otherwise. Let us set

$$S(x) = \sum_{p\le x} X_p.$$

We see that $S(x)$ is exactly what we have denoted $\pi_f(x,a)$ and according to the conjecture, it has the expected value

$$c_{f,a}\pi_{1/2}(x).$$

Let us set

$$S^*(x) = \frac{S(x) - c_{f,a}\pi_{1/2}(x)}{\sqrt{c_f(a)\pi_{1/2}(x)}}.$$

Then the law of the iterated logarithm (see [4], Volume 1, Theorem 8.5) gives that

$$\limsup_{x\longrightarrow\infty} \frac{S^*(x)}{\sqrt{2\log\log x}} \;=\; 1.$$

Equivalently, this means that

$$\limsup_{x\longrightarrow\infty} \frac{\pi_f(x,a) - c_{f,a}\pi_{1/2}(x)}{\sqrt{\pi_{1/2}(x)(\log\log x)}} \;=\; \sqrt{2c_f(a)}.$$

In section 6 we give graphs of the ratios of the above two functions for various values of a for two different modular forms.

4. The constant

The major part of the work in the monograph of Lang and Trotter [6] is to make a precise conjecture for the constant $c_{f,a}$. They propose a probabilistic model in which $c_{f,a}$ is expressed as an infinite product over all primes. The factor at the archimedean prime is determined by the constraint that the $a_f(p)$ should satisfy the Sato-Tate distribution law, and the factors at a finite prime ℓ is determined by the image of the ℓ-adic representation of $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ associated to f. They perform this calculation for those f which correspond to elliptic curves (that is, for forms of weight 2 and having rational integer Fourier coefficients). For the distribution of supersingular primes (that is, $a = 0$) and general f of weight 2, there is a conjecture for the constant made by Bayer and Gonzalez [1].

For comparison of the conjectural value of the constant as given by [6] for the curve $X_0(11)$ with our computed approsimation of this computation, we include a table of these values on the following page.

a	$C_{LT}(a)$	$c_f(a)$
-20	0.9847812790	0.9671494409
-18	0.9847812790	1.1353493437
-17	0.4991978833	0.5026883458
-16	0.9847812790	0.9537699032
-15	0.4991978833	0.6670655235
-13	0.4991978833	0.5065110708
-12	0.9847812790	1.1582856941
-11	0.4538162576	0.4491701949
-10	0.9847812790	0.9747948910
-8	0.9847812790	0.9308335528
-7	0.4991978833	0.4759292703
-6	0.9847812790	1.1621084191
-5	0.4991978833	0.4663724577
-3	0.4991978833	0.6192814602
-2	0.9847812790	1.0053766915
-1	0.4991978833	0.5351815088
0	1.3137569279	1.2079811199
1	0.0000000000	0.0038227251
2	0.9847812790	0.9977312414
3	0.4991978833	0.5791428470
4	0.9847812790	1.0474266672
5	0.4991978833	0.5256246962
7	0.4991978833	0.4969542582
8	0.9847812790	1.0760971052
9	0.4991978833	0.5657633093
10	0.9847812790	0.9002517523
12	0.9847812790	1.1869561320
13	0.4991978833	0.4415247448
14	0.9847812790	1.0015539665
15	0.4991978833	0.5886996597
17	0.4991978833	0.4835747205
18	0.9847812790	1.1009448181
19	0.4991978833	0.4701951827
20	0.9847812790	0.9308335528

One is struck by what appears to be a rather large variance in the conjectured constant and the computed approximation to the constant. It may be worthwhile to re-examine the philosophy by which Lang and Trotter arrived at the conjectural value of $c_{f,a}$. They assume that the distribution of the $a_f(p)$ according to the Sato-Tate law is independent of the distribution properties afforded by the ℓ-adic representations. It is not clear to us whether this hypothesis is justified.

It is important to remark thot the heuristics of the previous section for the error term can be expected to hold only for fixed a as $x \longrightarrow \infty$. If a is also allowed to grow as a function of x, then the behavior might be rather different as the result below indicates.

Proposition 4.1. *Let*

$$f(z) \;=\; (\eta(z)\eta(11z))^2$$

and set

$$A \;=\; \prod_{\ell \neq 2,5,11} \frac{\ell(\ell^2-\ell-1)}{(\ell-1)(\ell^2-1)}.$$

There exists a such that

$$\pi_f(x,a) \;<\; \left(\frac{\pi}{8A}+\epsilon\right) c_{f,a}\pi_{1/2}(x).$$

Remark. Computing the product over primes up to 10^7, one finds that

$$A \approx 0.933189264603 \qquad \text{and}$$

$$\frac{\pi}{8A} \approx 0.420813972679$$

Proof. Suppose that for all a, we have

$$\pi_f(x,a) \;\geq\; \left(\frac{\pi}{8A}+\epsilon\right) c_{f,a}\pi_{1/2}(x).$$

Then, summing over a, we deduce that

$$\pi(x) \;\geq\; \left(\frac{\pi}{8A}+\epsilon\right) \pi_{1/2}(x) \sum_{|a|\leq 2\sqrt{x}} c_{f,a}.$$

Thus,

$$\sum_{|a|\leq 2\sqrt{x}} c_{f,a} \;\leq\; \left(\frac{8A}{\pi}-\epsilon\right)\sqrt{x}.$$

The constants $c_{f,a}$ have been given explicitly by Lang and Trotter [6]. In the supersingular case, we have

$$c_{f,0} \;=\; \frac{23\pi}{55} \;\sim\; 1.313755818.$$

Now, for $a \neq 0$, we have

$$c_{f,a} \;=\; \frac{2}{\pi} c_1(a) A$$

where

$$c_1(a) = c_2(a)c_3(a)$$

with

$$c_2(a) = \begin{cases} 5/4 & \text{if } a \not\equiv 1(\mod 5) \\ 0 & \text{otherwise} \end{cases}$$

and

$$c_3(a) = \frac{11}{2^3 \cdot 3^2 \cdot 5^2} c_4(a)$$

with

$$c_4(a) \;=\; \begin{cases} 230 & \text{if } 22|a \\ 217 & \text{if } 2|a,\ 11 \not| a \\ 100 & \text{if } 2 \not| a,\ 11|a \\ 110 & \text{if } (a, 22) = 1. \end{cases}$$

Putting all of this together yields

$$\sum_{|a| \leq 2\sqrt{x}} c_{f,a} \;\sim\; \frac{8}{\pi} A\sqrt{x}.$$

This contradicts the inequality derived above and completes the proof of the Proposition.

5. The computations

From a computational point of view, it is easier to work with forms which can be expressed in terms of θ-series. Indeed, in this case the Fourier coefficients are integers and we have algorithms which will quickly compute them. The essential tool is the discrete *fast Fourier transform*. Before describing the computations done for this paper, we will briefly describe the fast Fourier transform algorithm (FFT) and how it can be used to construct modular forms. For a more detailed account of fast multiplication using FFT see [2].

Let $M = 2^k$ and let $\mathbb{F}$ be a field possessing a primitive M-th root of unity ω. We will denote by W_M the set of all M-th roots of unity in $\mathbb{F}$. The purpose of the fast Fourier transform algorithm is to evaluate a polynomial q of degree smaller than M over $\mathbb{F}$ simultaneously on all of W_M. Suppose that $q(x) = \sum_{i=0}^{M-1} a_i x^i$. Let $q_1 = \sum_{i=0}^{\frac{M-1}{2}} a_{2i} x^i$ and let $q_2 = \sum_{i=0}^{\frac{M-2}{2}} a_{2i+1} x^i$. The key idea behind FFT is the observation that

$$q(\omega^i) = q_1(\omega^{2i}) + \omega^i q_2(\omega^{2i}) \qquad \text{and,} \qquad q(\omega^{M/2+i}) = q_1(\omega^{2i}) - \omega^i q_2(\omega^{2i}).$$

Thus, once we know $q_1(W_{M/2})$ and $q_2(W_{M/2})$, we only need to perform $M/2$ multiplications to compute $q(W_M)$. More precisely, if we let $T(M)$ denote the number of multiplications required to evaluate a polynomial of degree at most $M-1$ on W_M, then we see by using the above observation that $T(M) = 2T(M/2) + M/2$. If we recursively apply the above observation and note that $T(1) = 0$ then we see that $T(M) = \frac{M}{2} \log_2 M/2$. Therefore, we see that the running time of FFT is $O(M \log M)$. Now, we recall that if $p \equiv 1$ modulo M is a prime, then $\mathbb{F}_p$ is a

field possessing a primitive M-th root of unity. In the case that $\mathbb{F} = \mathbb{F}_p$, we refer to the above algorithm as the discrete FFT.

We will now describe how the discrete FFT can be used to build modular forms. Recall that we only need $\frac{k}{12}N\prod_{p|N}\left(1+\frac{1}{p}\right)$ Fourier coefficients in order to distinguish the cusp-forms in $S_k(N,\chi)$. Thus for computational purposes we will treat these cuspforms as polynomials modulo x^M for some $M > \frac{k}{12}N\prod_{p|N}\left(1+\frac{1}{p}\right)$. If M is also a power of 2, then we will let p be a prime which is 1 modulo $2M$ and, let ω be a primitive $2M$-th root of unity modulo p. Then we have seen above that the discrete FFT will allow us to evaluate any polynomial f of degree less than $2M$ simultaneously at all of the $2M$-th roots of unity modulo p in time proportional to $M\log M$.

If we wish to multiply two polynomials f and g over $\mathbb{F}_p$ each of degree less than M, then we can use discrete FFT to evaluate each of these at all of the $2M$-th roots of unity. Note that $h = fg$ is the unique polynomial of degree less than $2M$ with $h(\omega^i) = f(\omega^i)g(\omega^i)$ for $i = 1, 2, \ldots, 2M$. Thus, to determine the coefficients of h all we need to do is to solve the linear system:

$$\begin{pmatrix} 1 & \omega & \omega^2 & \ldots & \omega^{2M-1} \\ 1 & \omega^2 & \omega^4 & \ldots & (\omega^2)^{2M-1} \\ & & & \ddots & \\ 1 & \omega^M & (\omega^M)^2 & \ldots & (\omega^M)^{2M-1} \end{pmatrix} \begin{pmatrix} h_0 \\ h_1 \\ \vdots \\ h_{2M-1} \end{pmatrix} = \begin{pmatrix} f(\omega)g(\omega) \\ f(\omega^2)g(\omega^2) \\ \vdots \\ f(\omega^{2M-1})g(\omega^{2M-1}) \end{pmatrix}$$

Now, we note that the matrix on the left is a Vandermonde and its inverse is:

$$\frac{1}{M}\begin{pmatrix} 1 & \omega^{-1} & \omega^{-2} & \ldots & (\omega^{-1})^{2M-1} \\ 1 & \omega^{-2} & \omega^{-4} & \ldots & (\omega^{-2})^{2M-1} \\ & & & \ddots & \\ 1 & \omega^{-M} & \omega^{-2M} & \ldots & (\omega^{-M})^{2M-1} \end{pmatrix}$$

Therefore, in order to determine the coefficients $h_0, \ldots h_{2M-1}$ we simply need to evaluate the polynomial $\sum_{i=0}^{2M-1} f(\omega^i)g(\omega^i)x^i$ at all of the $2M$-th roots of unity and we can once again rely on FFT for this. Thus we can multiply any two polynomials of degree less than M over $\mathbb{F}_p$ in $O(M\log M)$ time.

In order to obtain the exact values of the first M Fourier coefficients of a product of modular forms g_1 and g_2, we select primes $p_1, p_2, \ldots, p_j \equiv 1$ modulo $2M$ so that $\prod_{i=1}^{j} p_i$ is larger than twice the Weil bound for the M-th Fourier coefficient of this product. We then compute the product of g_1 and g_2 (again thought of as polynomials modulo x^M) modulo each of the p_i's using discrete FFT and use the Chinese remainder theorem to determine the exact values of the first M coefficients.

In particular, if we wish to compute the first M Fourier coefficients of a θ-series arising from a quadratic form in one or two variables, we simply use the naive algorithm which is $O(M)$. If we wish to compute the Fourier coefficients of a θ series f arising from a quadratic form of more than two variables, then we first write it as a product $f = \prod \theta_i$ of θ-series each arising from a quadratic form of one or two variables. Using the naive algorithm, we can then compute

M Fourier coefficients of each of the θ_i's in time proportional to M. We then make use of the discrete FFT algor ithm to multiply the θ_i's together modulo some prime p. Using several primes if necessary, we thus obtain M Fourier coefficients of our θ-series f in $O(M \log M)$ time.

In order to shed light on Question 2.1, we initiated a study of the Fourier coefficients of a Hecke eigenform of weight 2 with an inner twist. Let

$$f_1 = \frac{1}{2}\left(2\theta_1\theta_{128} - \theta_1\theta_{32} - 2\theta_4\theta_{128} + \theta_4\theta_{32}\right) \cdot \left(2\theta_{16}\theta_{32} - \theta_8\theta_4\right) \quad \text{and}$$
$$f_2 = \frac{1}{4}\phi_1\phi_2\left(\theta_8\theta_4 - 2\theta_{16}\theta_{32}\right),$$

where

$$\theta_t(z) = \sum_{n\in\mathbb{Z}} q^{tn^2} \qquad (q = e^{2\pi i z}) \quad \text{and}$$
$$\phi_t(z) = \sum_{n\in\mathbb{Z}} \chi_8(n) q^{tn^2}.$$

Then $f = f_1 + \sqrt{2} f_2 \in S_2(512)$ is the normalized eigenform with an inner twist mentioned in [7]. To see that f has an inner twist, note that if σ denotes the nontrivial element of $\mathrm{Gal}(\mathbb{Q}(\sqrt{2})/\mathbb{Q})$ and if χ_4 denotes the nontrivial character modulo 4, then $f^\sigma = f_{\chi_4}$. Thus by Conjecture 2.1, we expect that

$$\pi_f(x,a) \sim c_{f,a}\pi_{1/2}(x), \qquad \text{and}$$
$$\theta_f(x,a) \sim d_{f,a}\theta_{1/2}(x),$$

where $c_{f,a}$ and $d_{f,a}$ are constants.

As outlined above, we used the straight forward $O(M)$ algorithm to compute the Fourier coefficients of each of the factors $(2\theta_1\theta_{128} - \theta_1\theta_{32} - 2\theta_4\theta_{128} + \theta_4\theta_{32})$, $(2\theta_{16}\theta_{32} - \theta_8\theta_4)$ and $\phi_1\phi_2$ of f_1 and f_2 and, then we used fast Fourier transform multiplication to obtain f_1 and f_2. For our computation, we chose $M = 67,108,863 = 2^{26} - 1$ and $p = 2,013,265,921$. Since, the Fourier coefficients of f_1 and f_2 are supported on the arithmetic progressions 1 and 3 (mod 4) respectively, and since $f = f_1 + \sqrt{2} f_2$ is a normalized eigenform of weight 2, it follows that if ℓ is a prime, then the ℓ-th Fourier coefficients of f_1 and f_2 are bounded in absolute value by $2\sqrt{\ell}$ which for $\ell < M = 67,108,863$ is much less than $p/2 = 1006632960.5$. In fact, it is not difficult to deduce from this that for any $n < 67,108,863$, we have $|a_{f_i}(n)| < 1006632960.5$ $(i = 1, 2)$. Thus, our computation of f_1 and f_2 mentioned above actually yields the exact values of the first $67,108,863$ Fourier coefficients of f_1 and f_2 so long as we remember to choose our representatives for the equivalence classes in $\mathbb{Z}/p\mathbb{Z}$ between -1006632960 and 1006632960.

In addition to the above calculations, we also considered a case studied by Lang and Trotter, namely the elliptic curve $X_0(11)$. The corresponding modular form is

$$g(z) \;=\; (\eta(z)\eta(11z))^2$$

where as usual

$$\eta(z) \;=\; \exp(\pi i z/12)\prod_{n=1}^{\infty}\left(1-e^{2\pi i n z}\right).$$

We recall that we can rewrite $\eta(z)\eta(11z)$ as

$$\eta(z)\eta(11z) \;=\; \sum_{i,j\in\mathbb{Z}}(-1)^{i+j}q^{\frac{11(3j^2+j)+3i^2+i}{2}}.$$

Thus it is easy to see that one can compute M Fourier coefficients of $\eta(z)\eta(11z)$ in $O(M)$ time. As before, we can use FFT to compute M Fourier coefficients of g in time proportional to $M\log M$.

There were several reasons for considering a case which was also studied in [6]. Firstly, the monograph [6] provides very limited numerical evidence for the conjectures. They computed the Fourier coefficients for the first 5,000 primes which corresponds to $x = 48,611$. To our knowledge, since the publication of [6], no further numerical studies have been made of the conjecture. Secondly, we wished to verify whether the oscillations that we observed in the case of a modular form with non-rational Fourier coefficients were peculiar to this case, or whether they already occur in the case of elliptic curves. The computations seem to indicate that there are oscillations in this case also, and that the convergence to the conjectured constant is rather slow.

6. The results

Based on the calculations discussed in the previous section, we calculated

$$\begin{aligned} c_f(x,a) &= \frac{\pi_f(x,a)}{\pi_{1/2}(x)},\\ d_f(x,a) &= \frac{\theta_f(x,a)}{\theta_{1/2}(x)},\\ c_g(x,a) &= \frac{\pi_g(x,a)}{\pi_{1/2}(x)} \qquad \text{and}\\ d_g(x,a) &= \frac{\theta_g(x,a)}{\theta_{1/2}(x)}. \end{aligned}$$

According to the conjecture, $c_f(x,a)$ and $d_f(x,a)$ should tend to a constant $c_{f,a}$ as $x\to\infty$ and likewise $c_g(x,a)$ and $d_g(x,a)$ should tend to a constant $c_{g,a}$ as $x\to\infty$. It is strange that until x is about 1.2×10^7, there is considerable oscillation in $c_f(x,a)$.

We wished also to test the implication mentioned at the end of section 3. So, we computed the normalized differences:

$$\frac{\pi_f(x,a)-c_f(2^{26},a)\pi_{1/2}(x)}{\sqrt{2c_f(2^{26},a)\pi_{1/2}(x)(\log\log x)}},$$

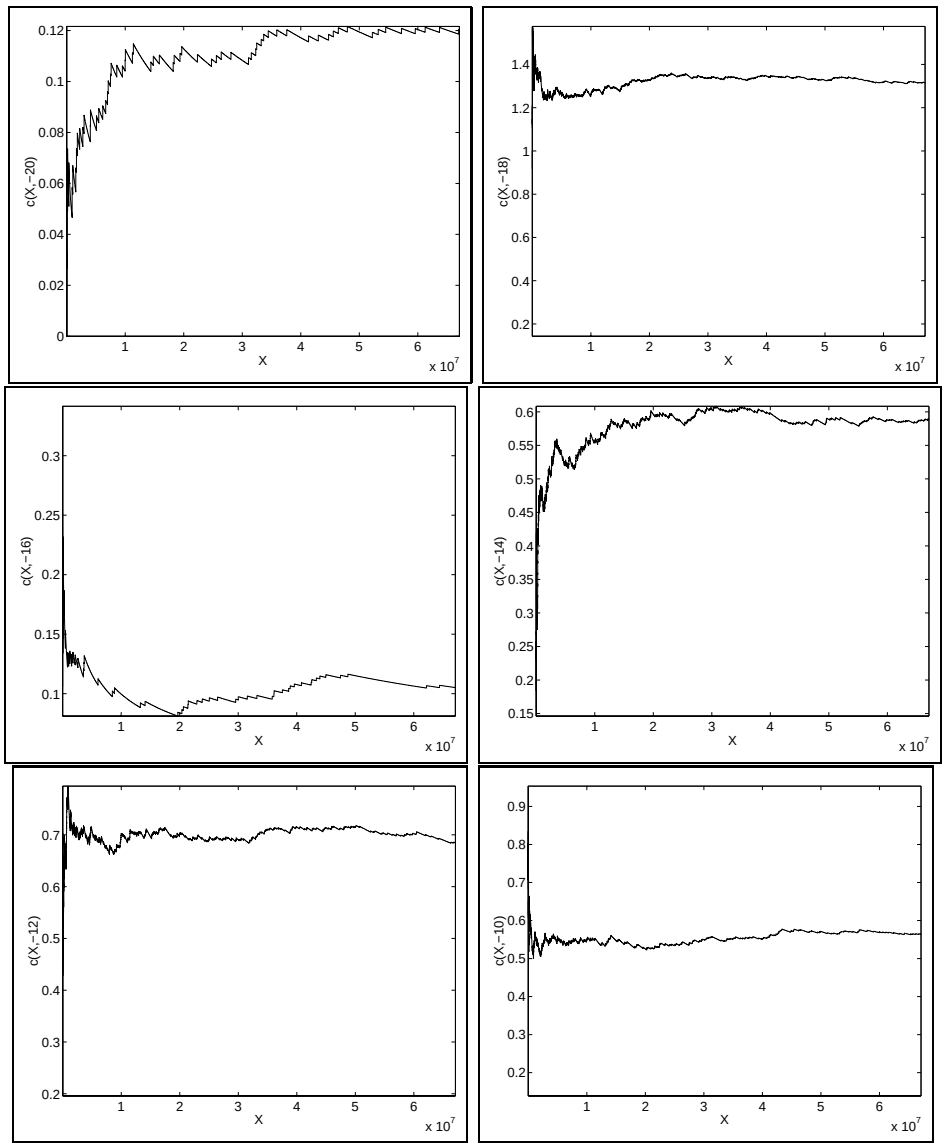
$$\frac{\theta_f(x,a) - d_f(2^{26},a)\theta_{1/2}(x)}{\sqrt{2d_f(2^{26},a)\theta_{1/2}(x)(\log\log x)}},$$

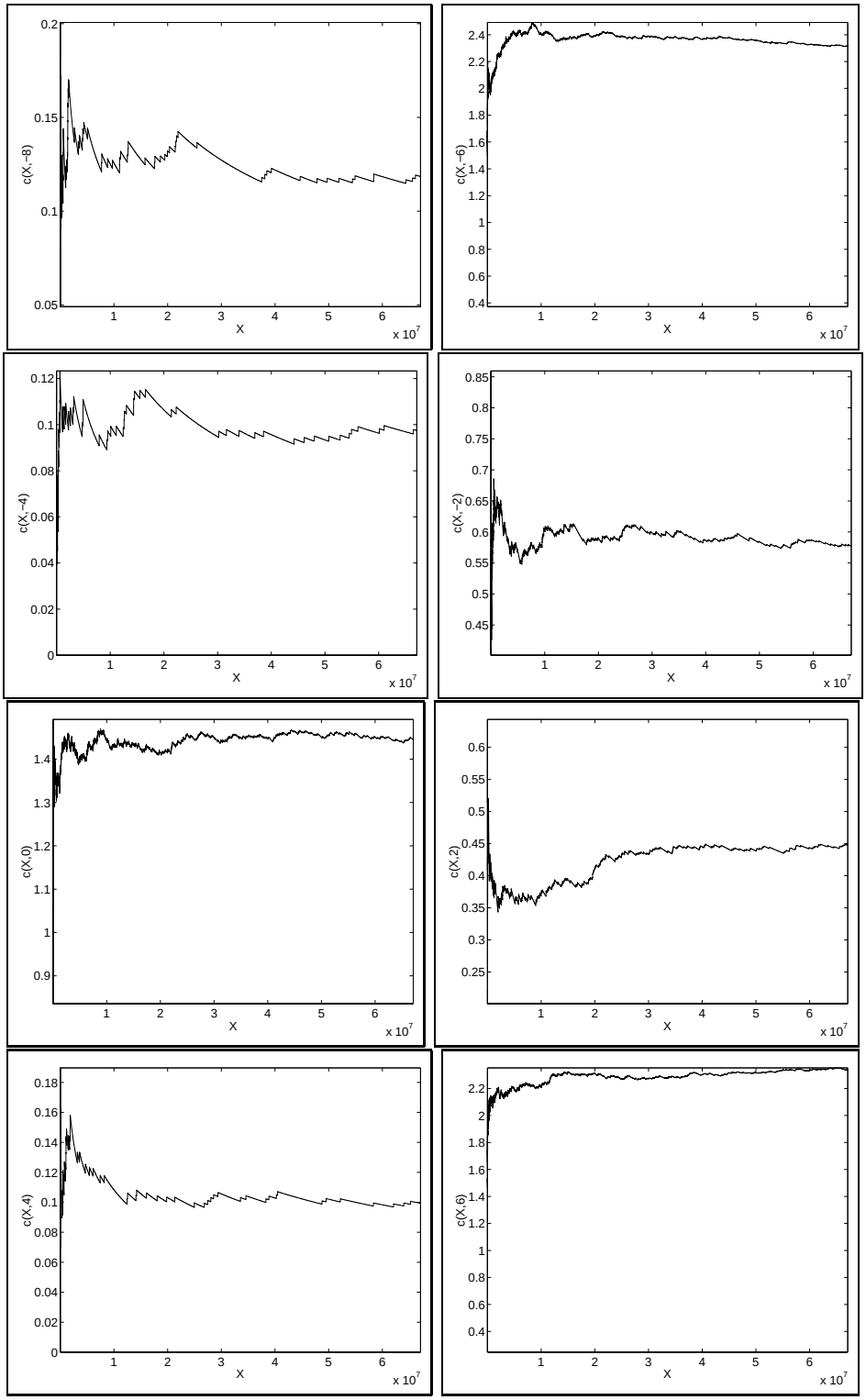
$$\frac{\pi_g(x,a) - c_g(2^{26},a)\pi_{1/2}(x)}{\sqrt{2c_g(2^{26},a)\pi_{1/2}(x)(\log\log x)}} \quad \text{and}$$

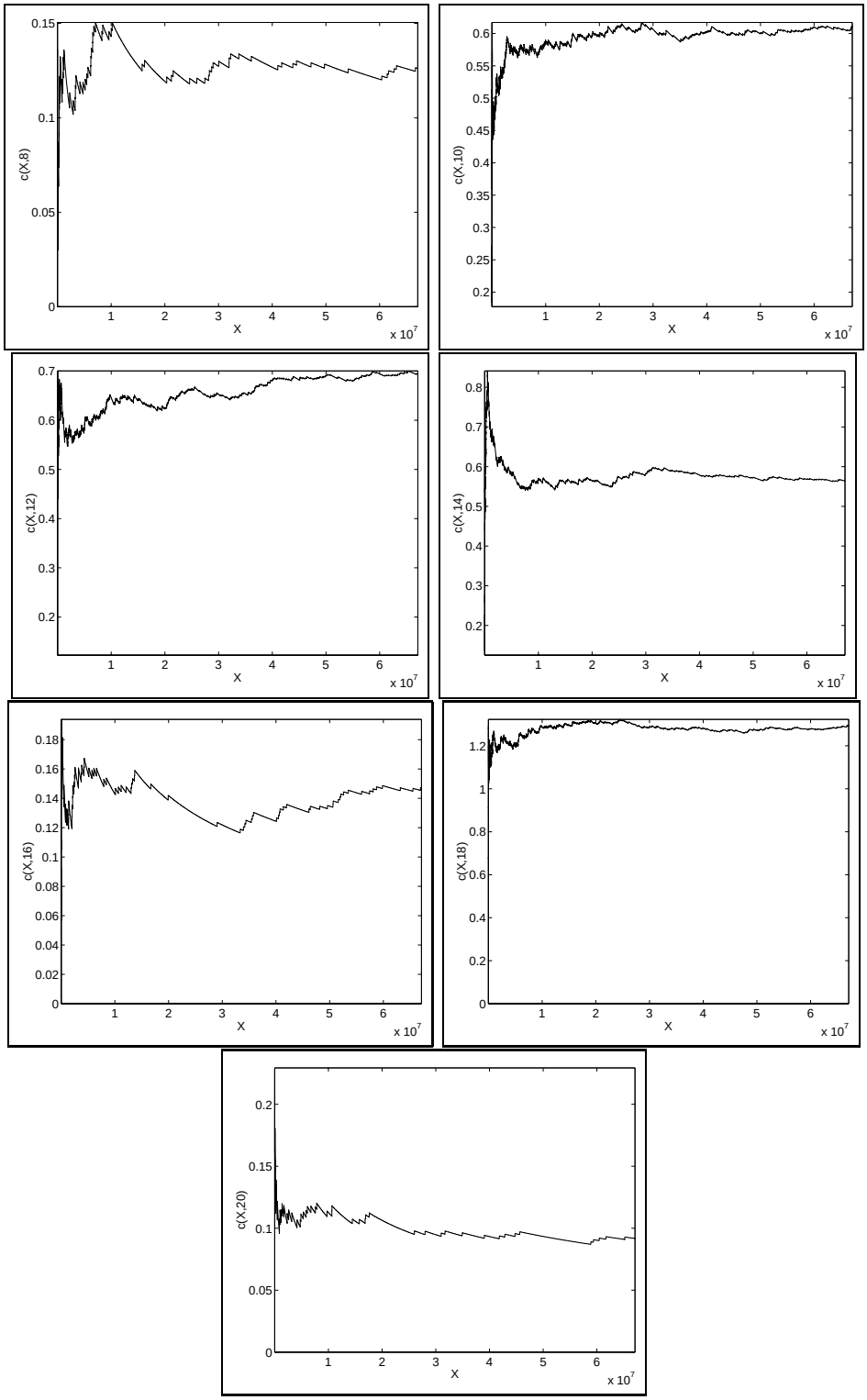
$$\frac{\theta_g(x,a) - d_g(2^{26},a)\theta_{1/2}(x)}{\sqrt{2d_g(2^{26},a)\theta_{1/2}(x)(\log\log x)}}.$$

The data is presented below both in tabular and graphic formats.

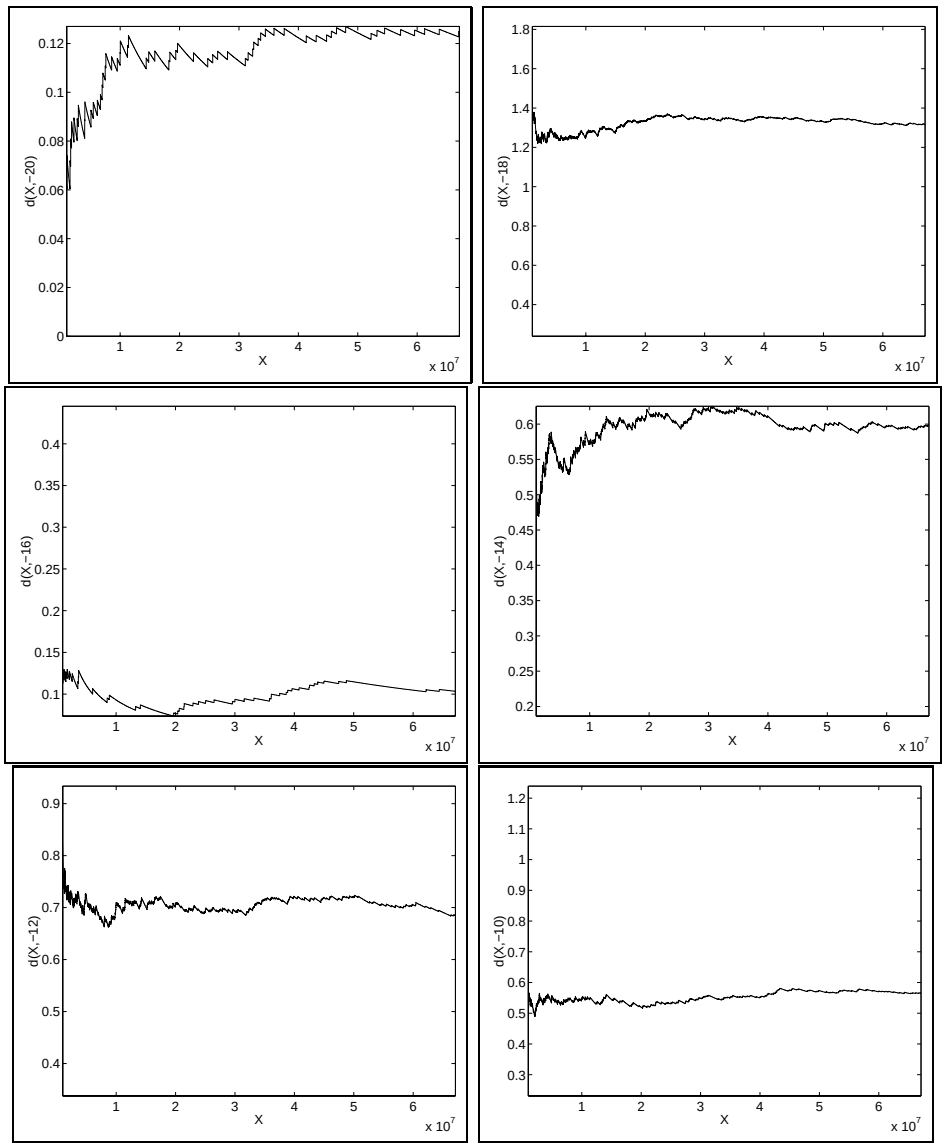
Graphs of $c_f(x, a)$

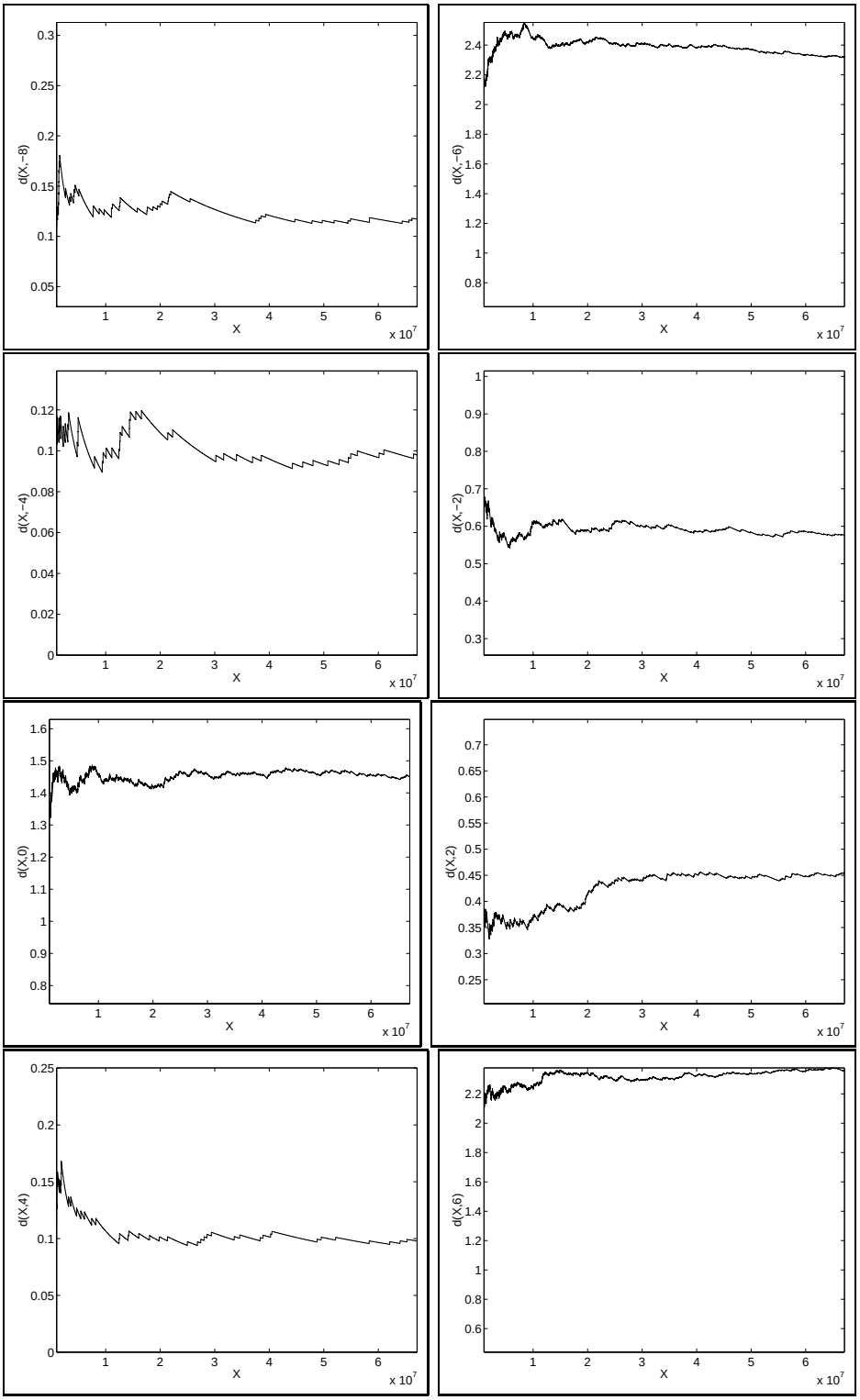


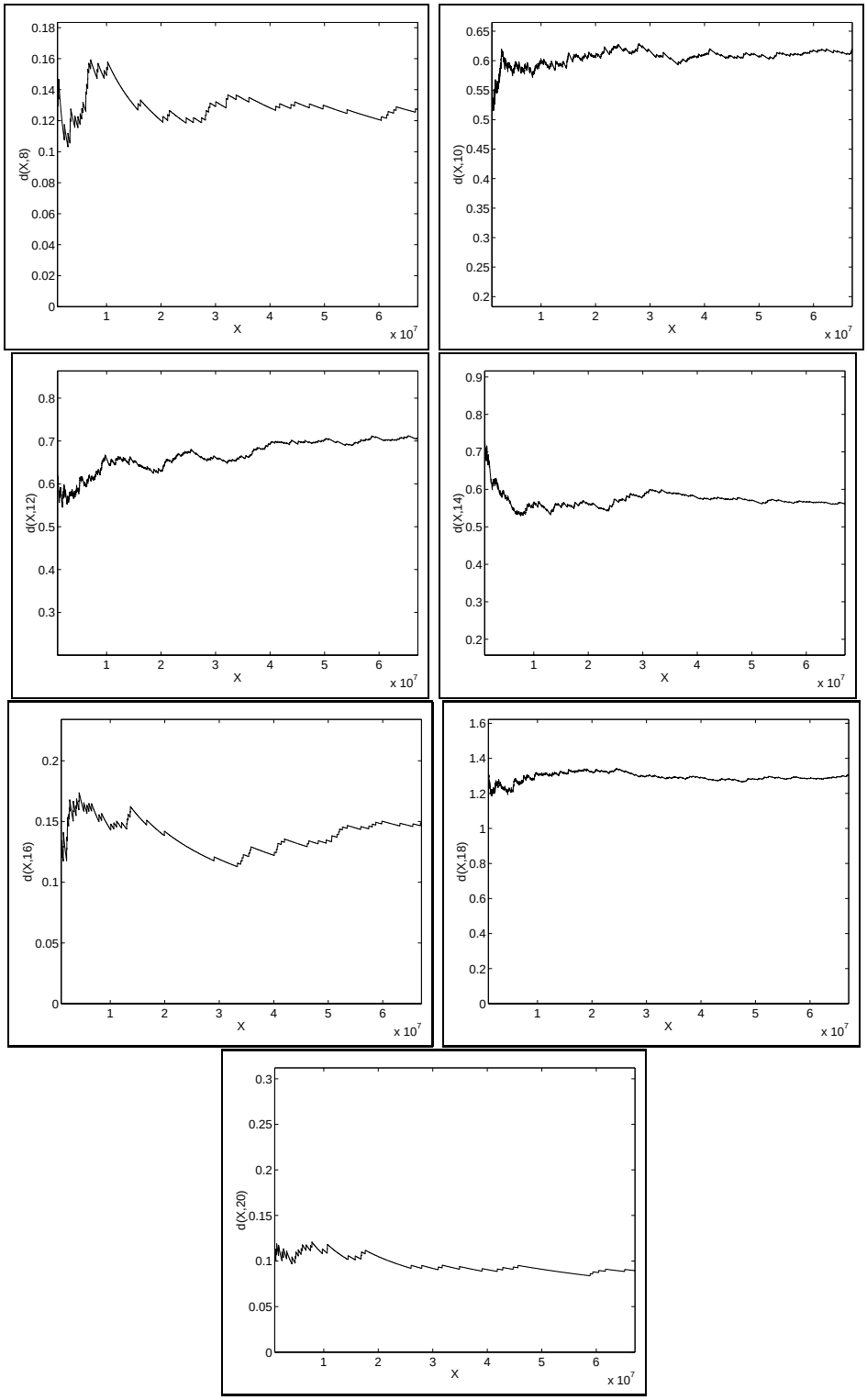

c(X,−8)
c(X,−6)
c(X,−4)
c(X,−2)
c(X,0)
c(X,2)
c(X,4)
c(X,6)
X
x 10^7


c(X,8)
c(X,10)
c(X,12)
c(X,14)
c(X,16)
c(X,18)
c(X,20)
X
x 10^7

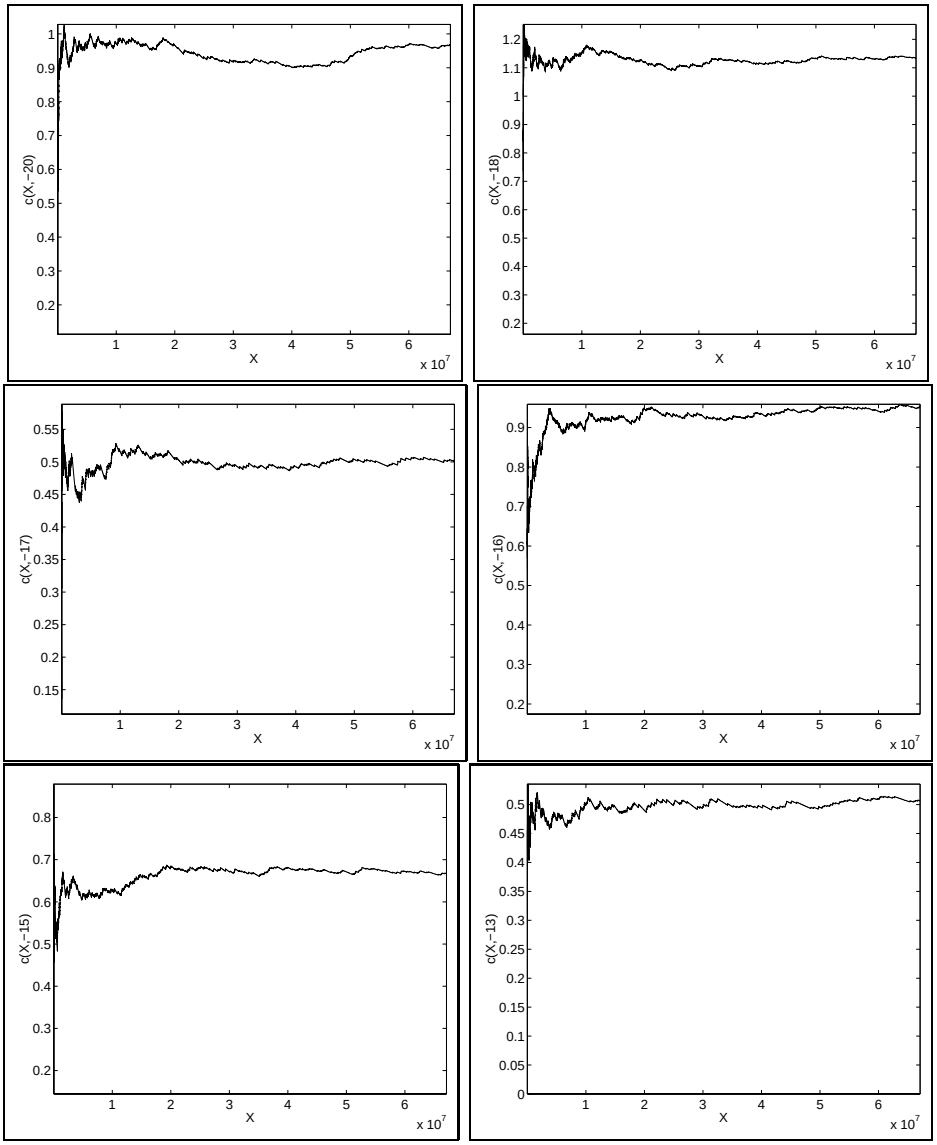
Graphs of $d_f(x, a)$

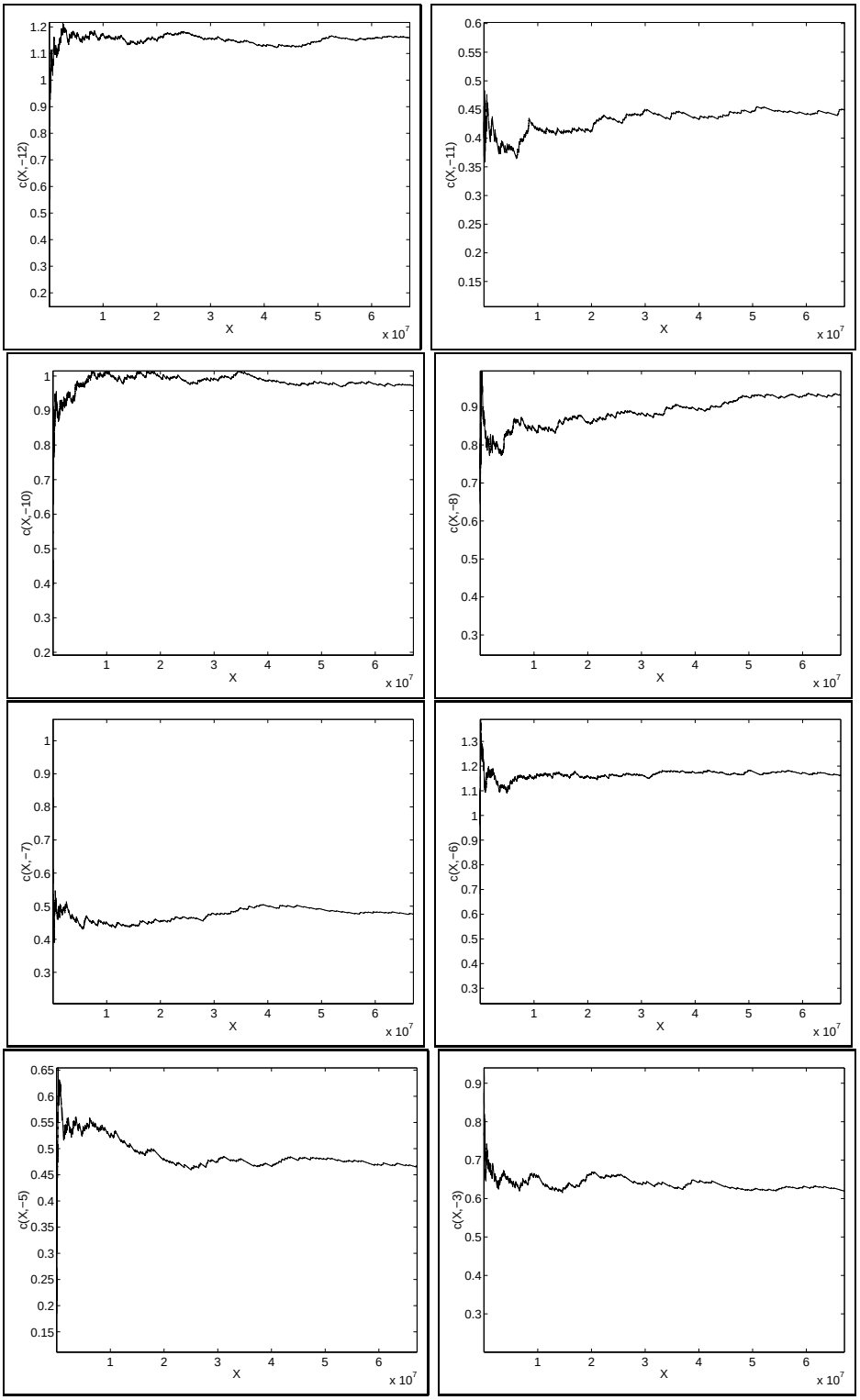


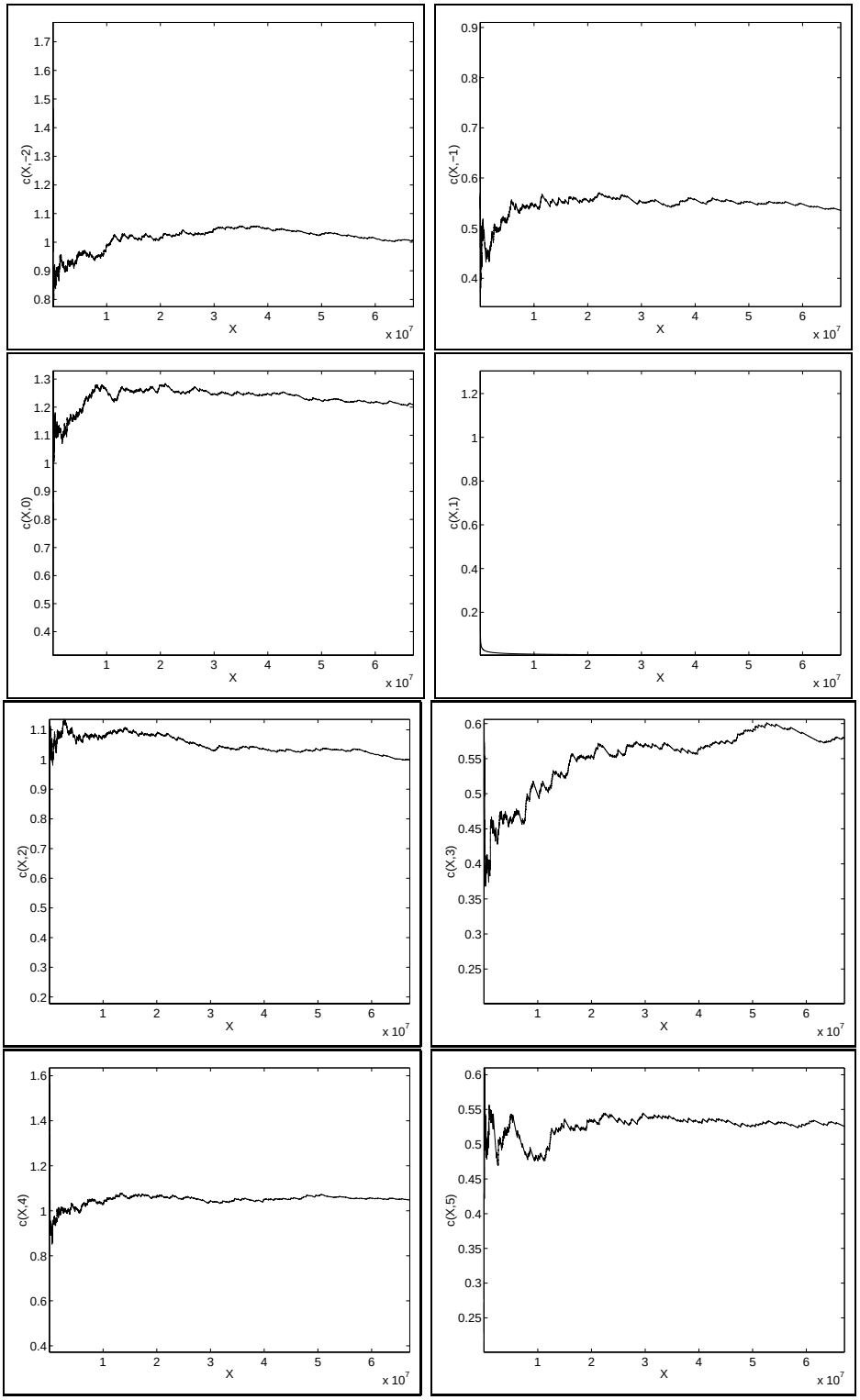

d(X,−8)
d(X,−6)
d(X,−4)
d(X,−2)
d(X,0)
d(X,2)
d(X,4)
d(X,6)
X
x 10^7

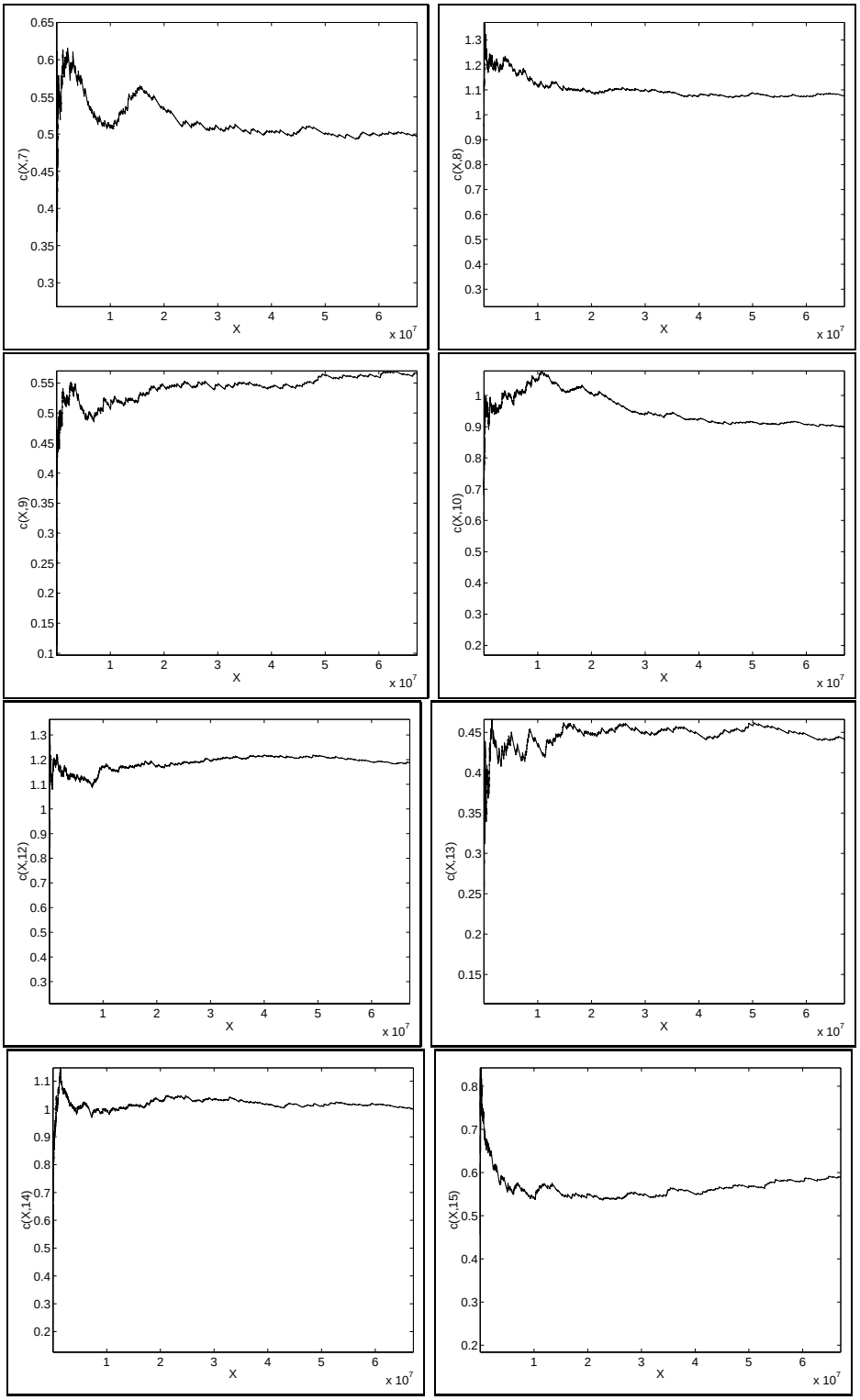

d(X,8)
d(X,10)
d(X,12)
d(X,14)
d(X,16)
d(X,18)
d(X,20)
X
x 10^7

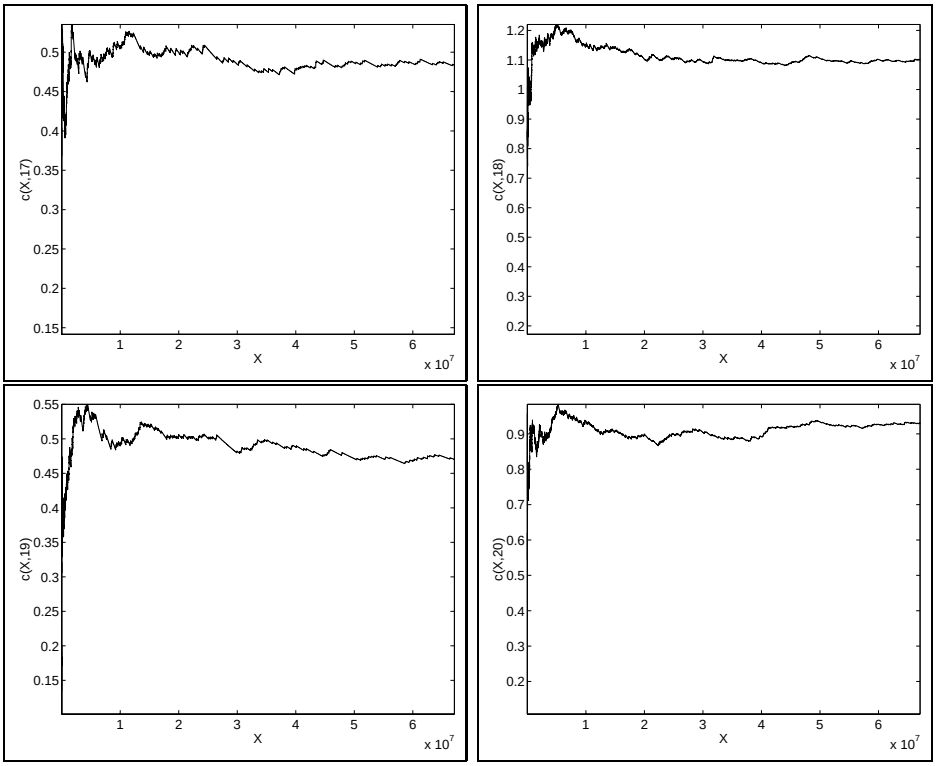
Graphs of $c_g(x, a)$



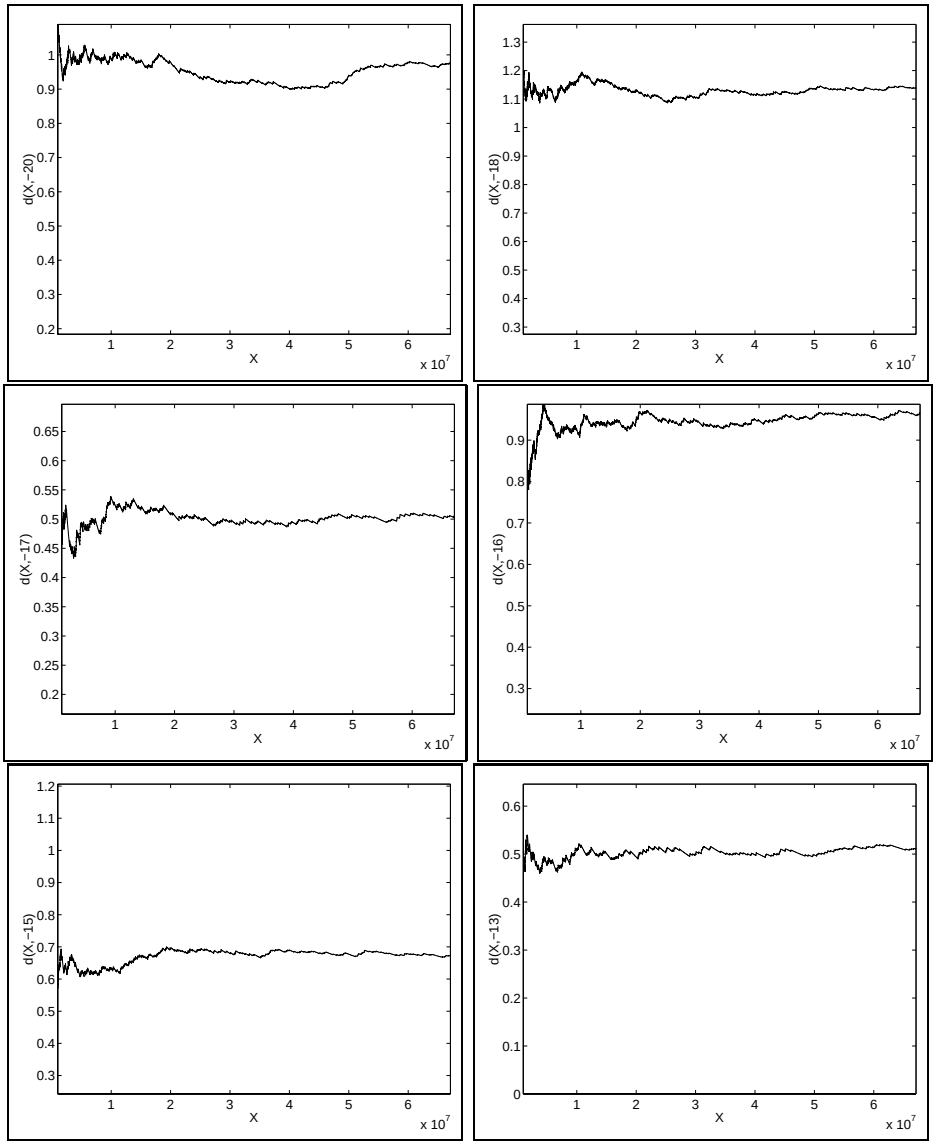

c(X,−12)
c(X,−11)
c(X,−10)
c(X,−8)
c(X,−7)
c(X,−6)
c(X,−5)
c(X,−3)
X
x 10^7

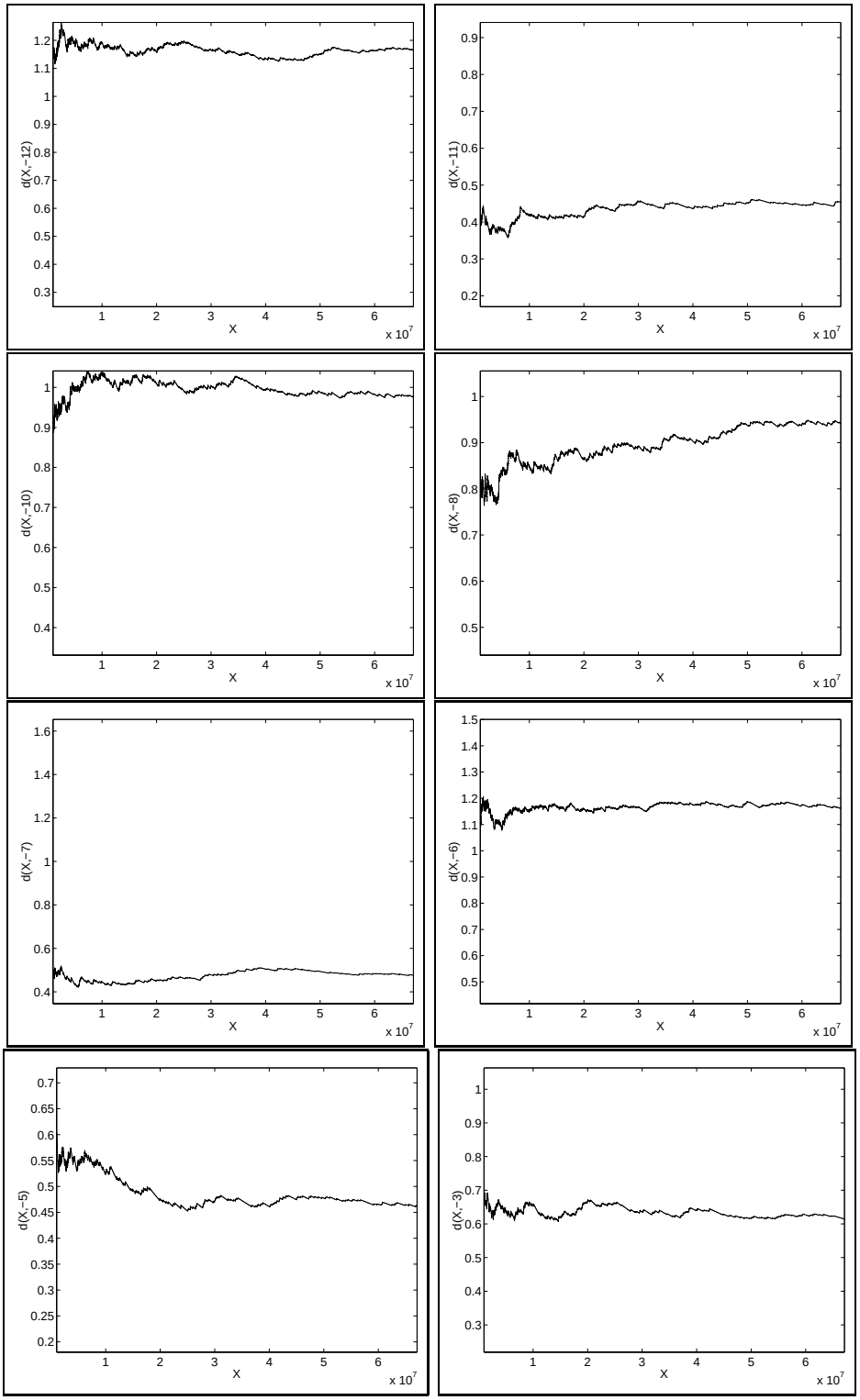

c(X,−2)
c(X,−1)
c(X,0)
c(X,1)
c(X,2)
c(X,3)
c(X,4)
c(X,5)
X
x 10^7

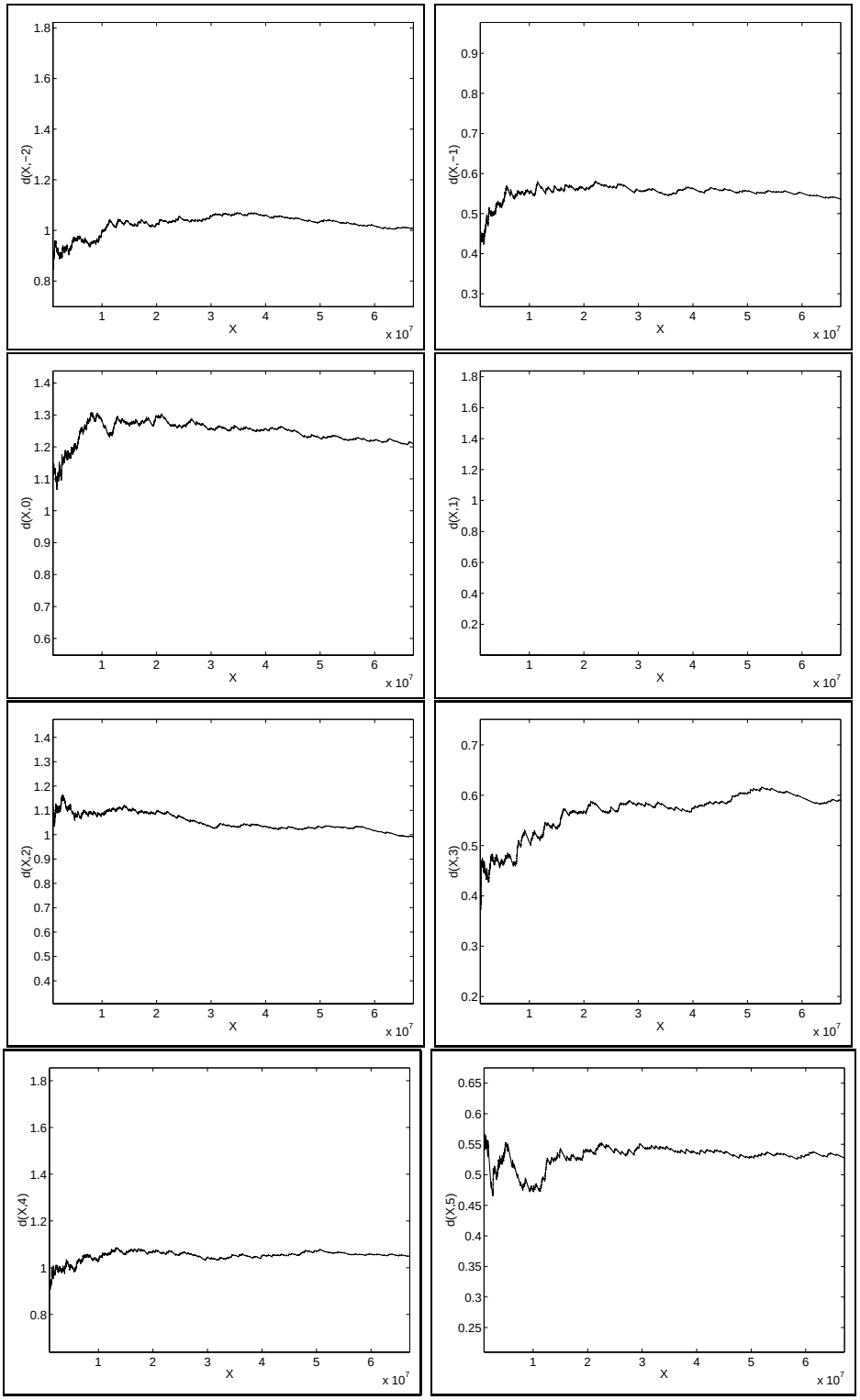

c(X,7)
c(X,8)
c(X,9)
c(X,10)
c(X,12)
c(X,13)
c(X,14)
c(X,15)
X
x 10^7

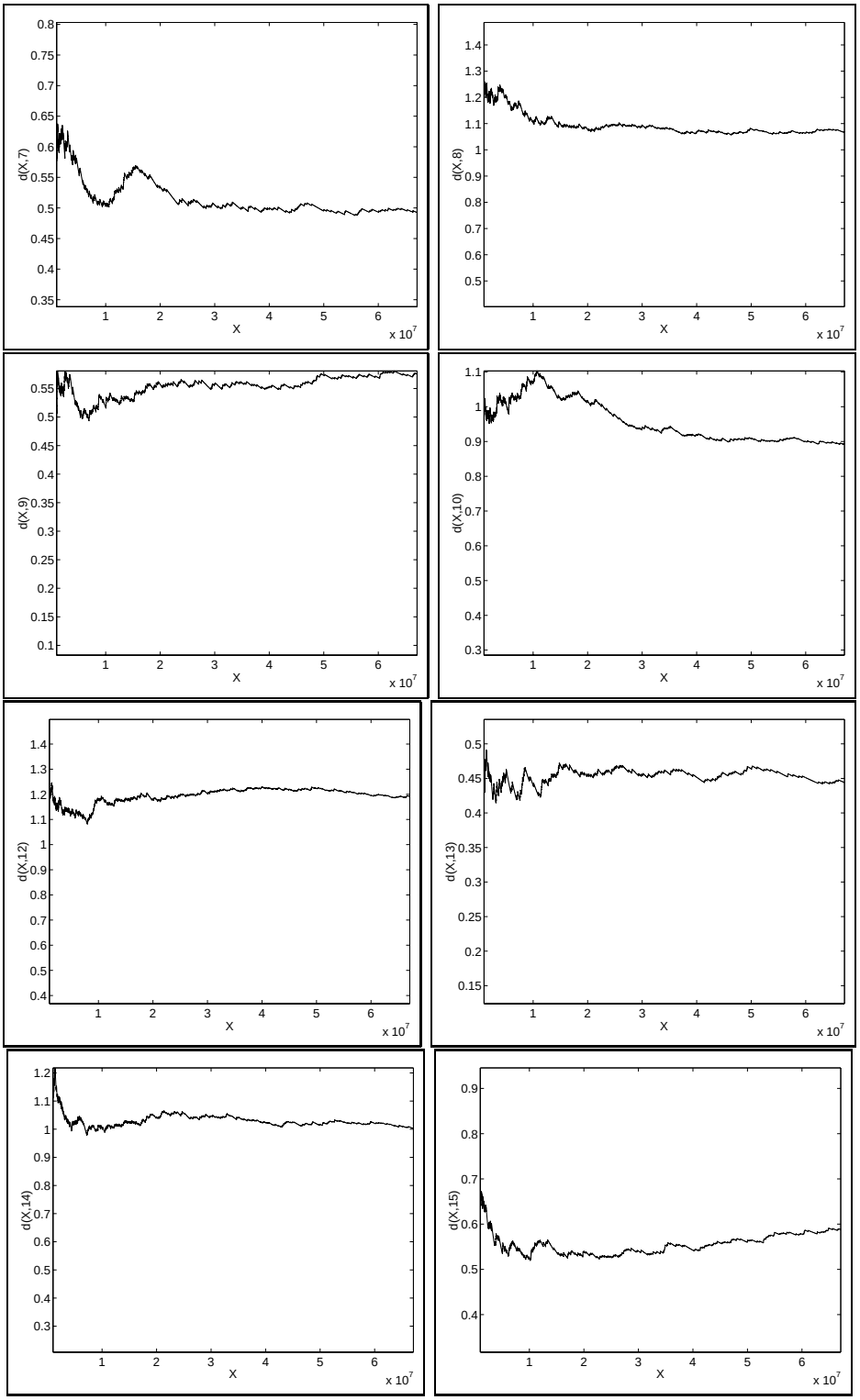

c(X,17)
X
x 10^7
c(X,18)
X
x 10^7
c(X,19)
X
x 10^7
c(X,20)
X
x 10^7

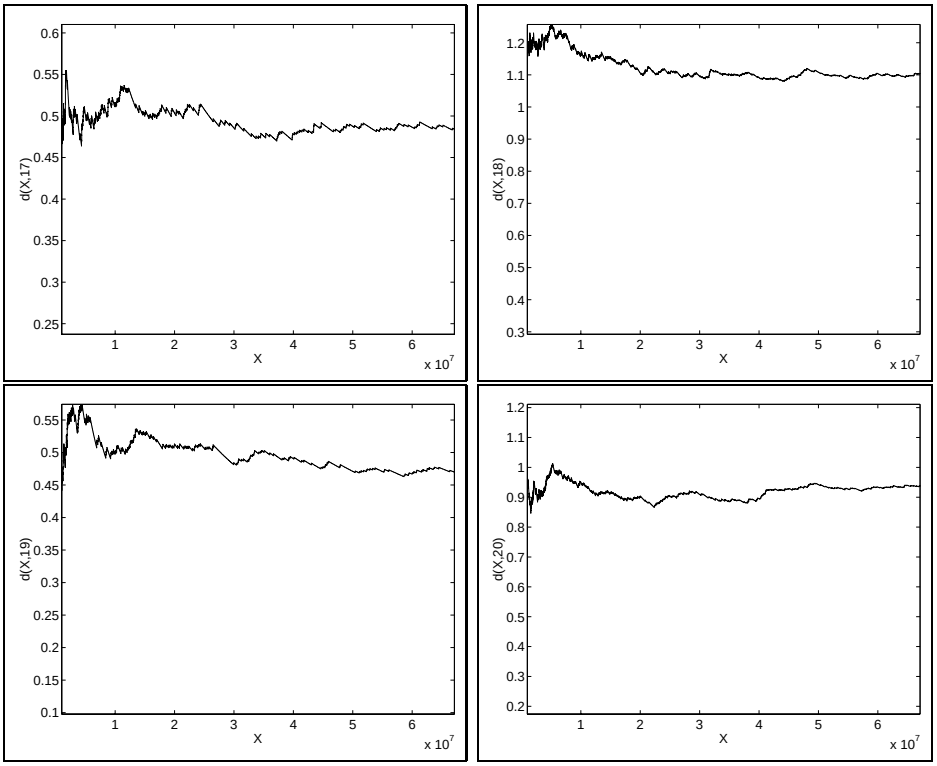
Graphs of $d_g(x, a)$



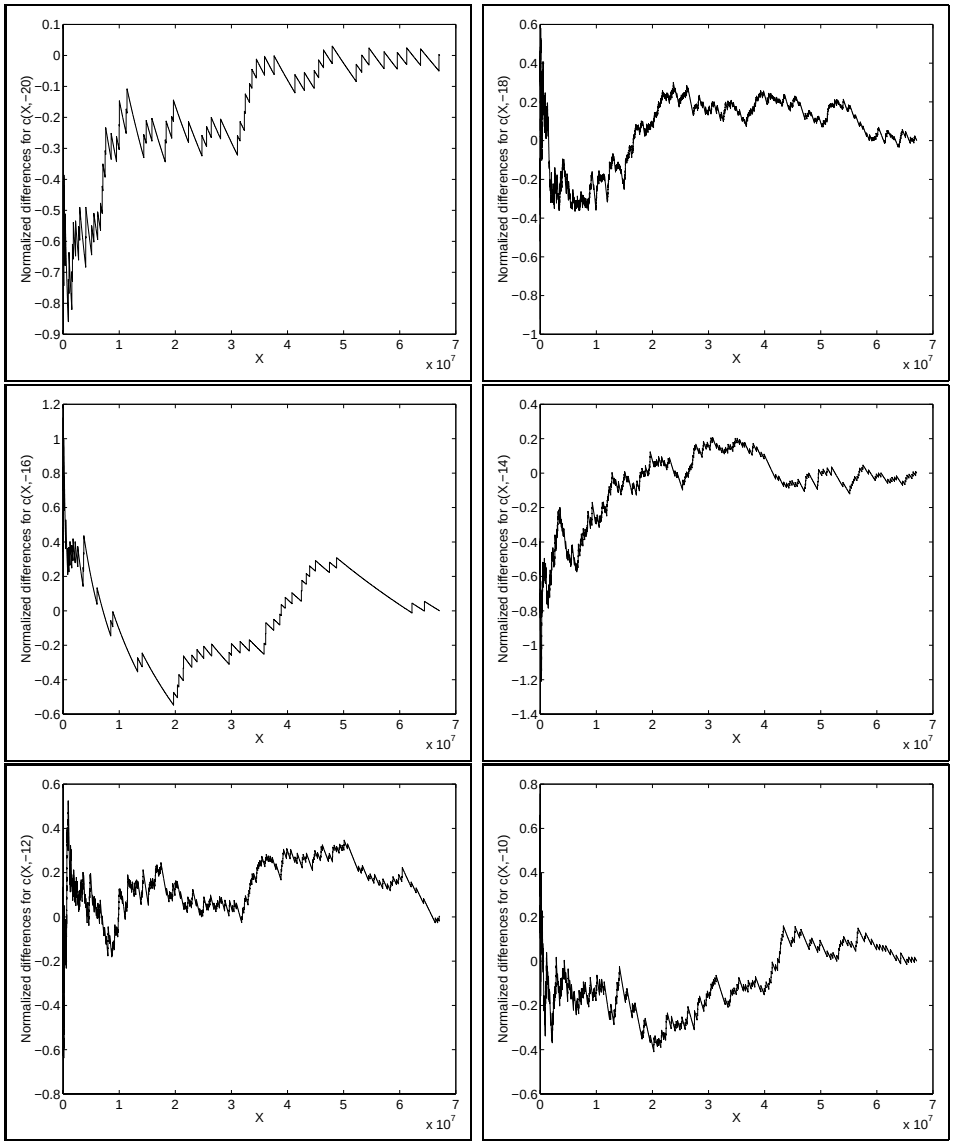

d(X,−12)
d(X,−11)
d(X,−10)
d(X,−8)
d(X,−7)
d(X,−6)
d(X,−5)
d(X,−3)
X
x 10^7

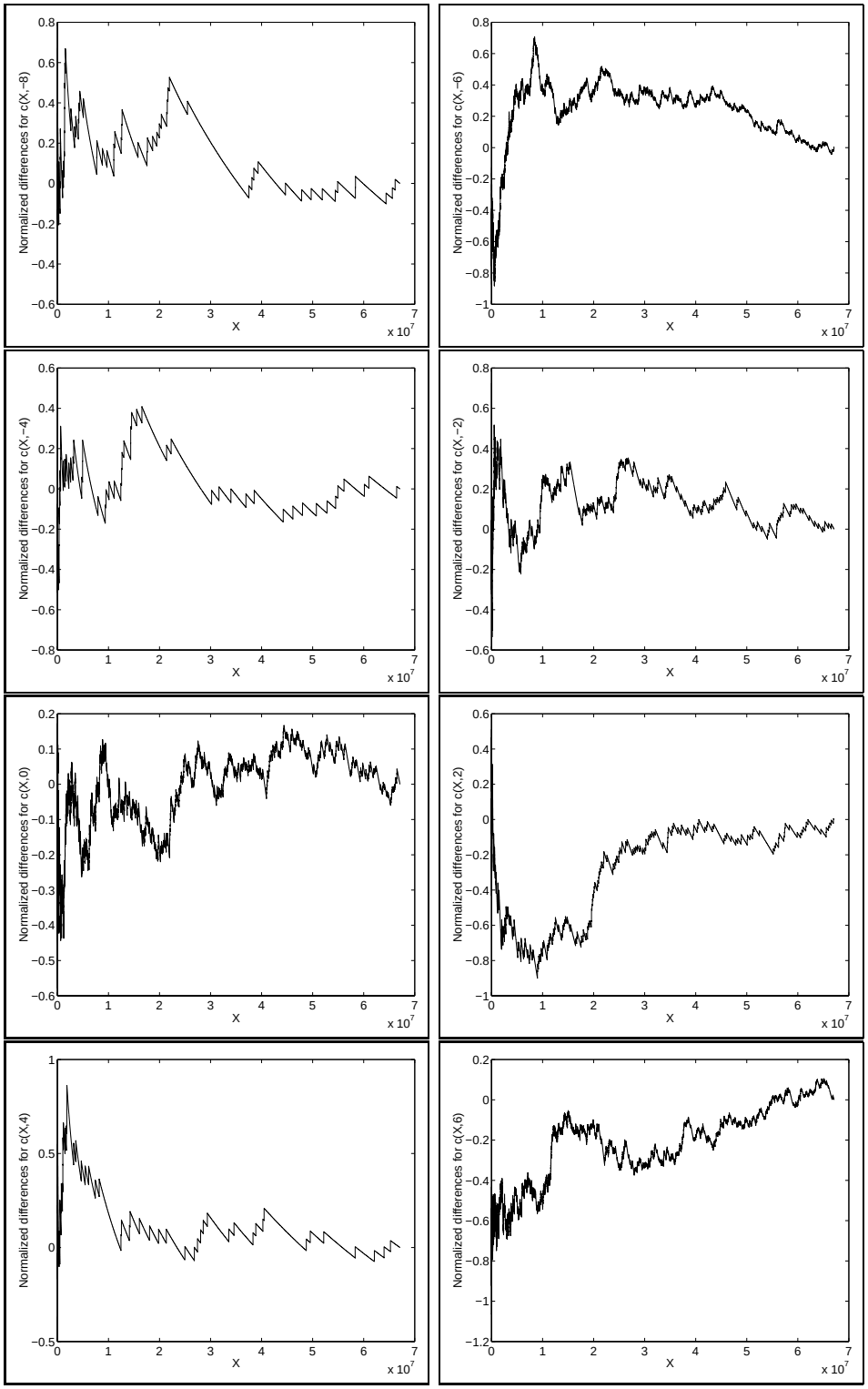

d(X,−2)
d(X,−1)
d(X,0)
d(X,1)
d(X,2)
d(X,3)
d(X,4)
d(X,5)
X
x 10^7

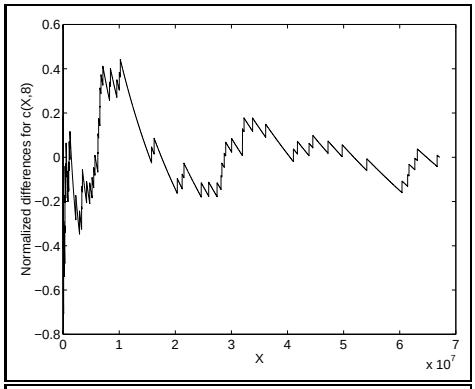

d(X,7)
d(X,8)
d(X,9)
d(X,10)
d(X,12)
d(X,13)
d(X,14)
d(X,15)
X
x 10^7

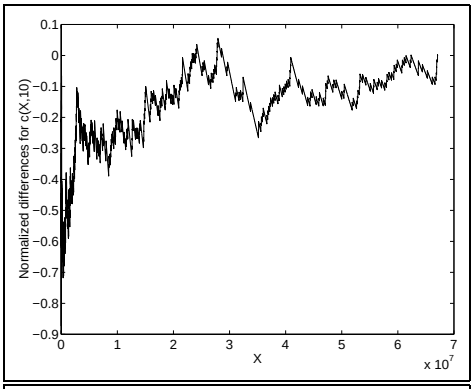

d(X,17)
X
x 10^7
d(X,18)
X
x 10^7
d(X,19)
X
x 10^7
d(X,20)
X
x 10^7

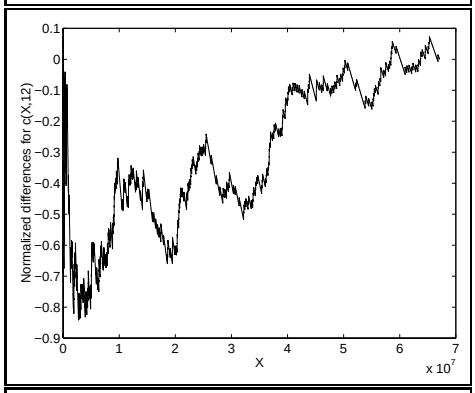
Normalized Differences for $\pi_f(x, a)$

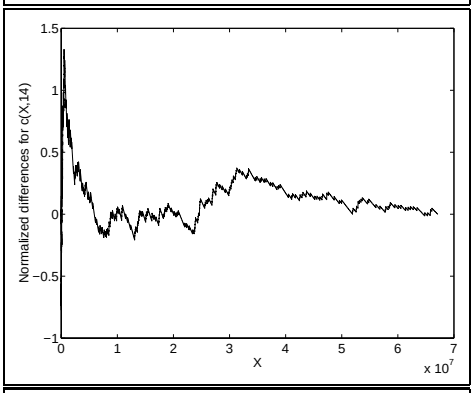


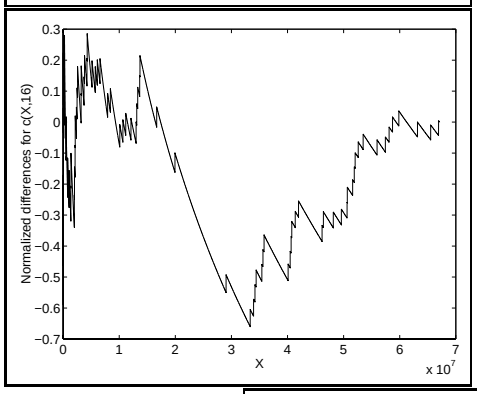

Normalized differences for c(X,−8)
Normalized differences for c(X,−6)
Normalized differences for c(X,−4)
Normalized differences for c(X,−2)
Normalized differences for c(X,0)
Normalized differences for c(X,2)
Normalized differences for c(X,4)
Normalized differences for c(X,6)
X
x 10^7

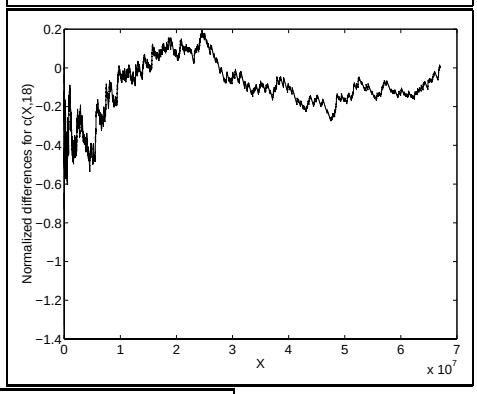

Normalized differences for c(X,8)
X
x 10^7

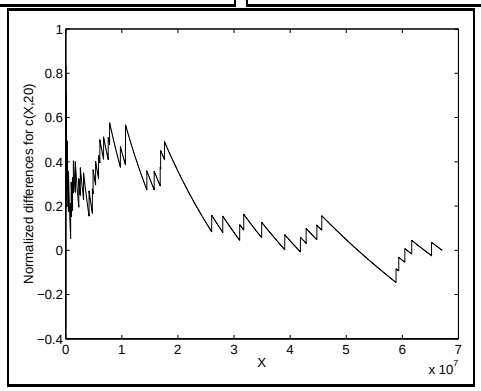

Normalized differences for c(X,10)
X
x 10^7


Normalized differences for c(X,12)
X
x 10^7

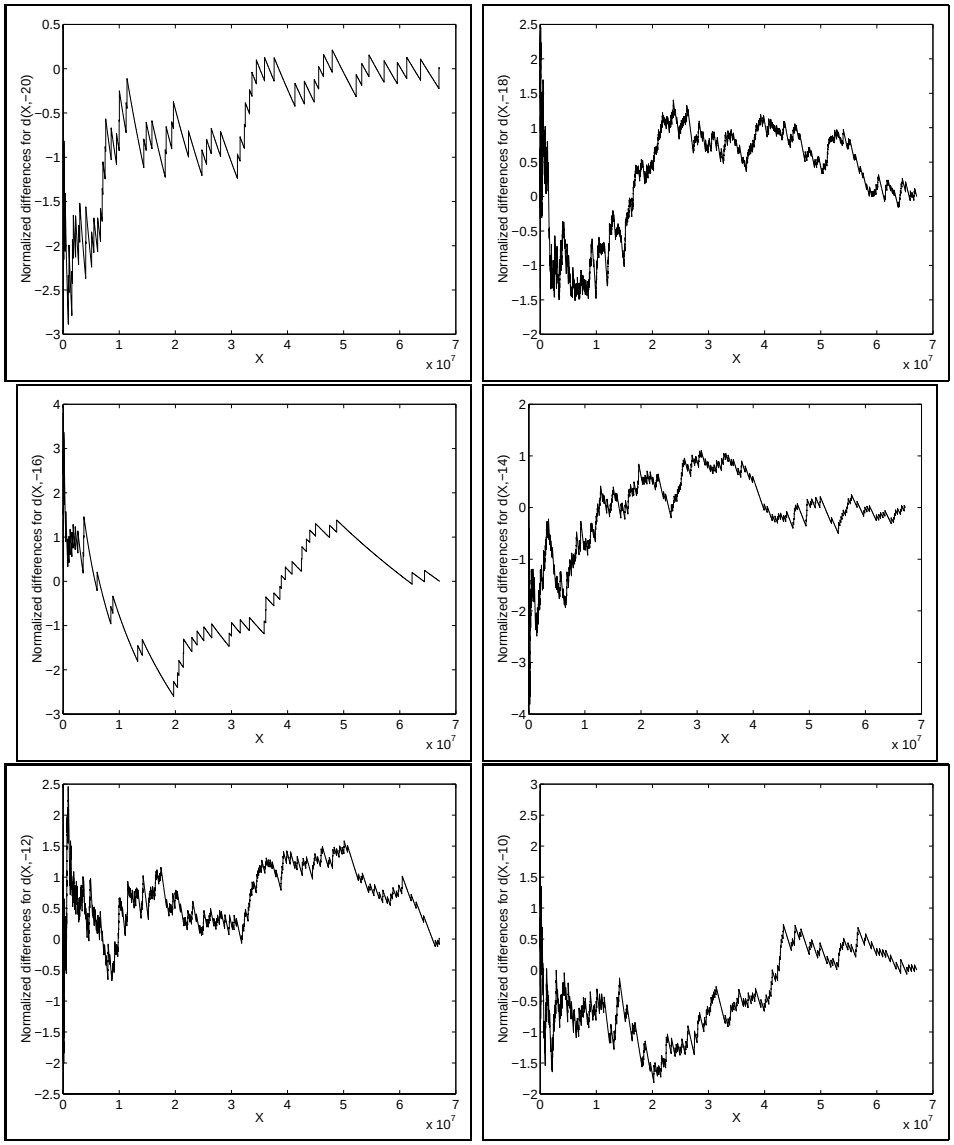

Normalized differences for c(X,14)
X
x 10^7

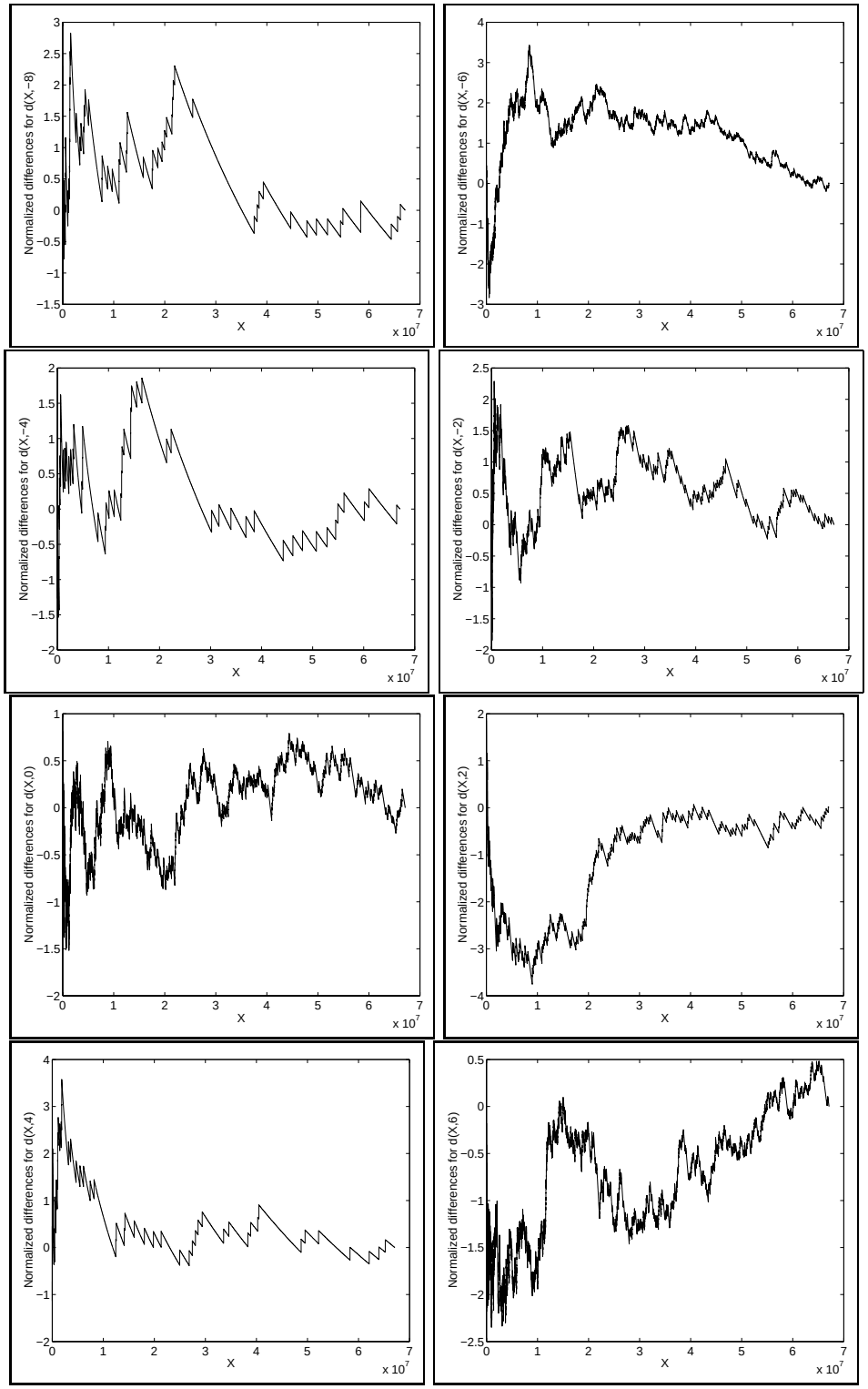

Normalized differences for c(X,16)
X
x 10^7

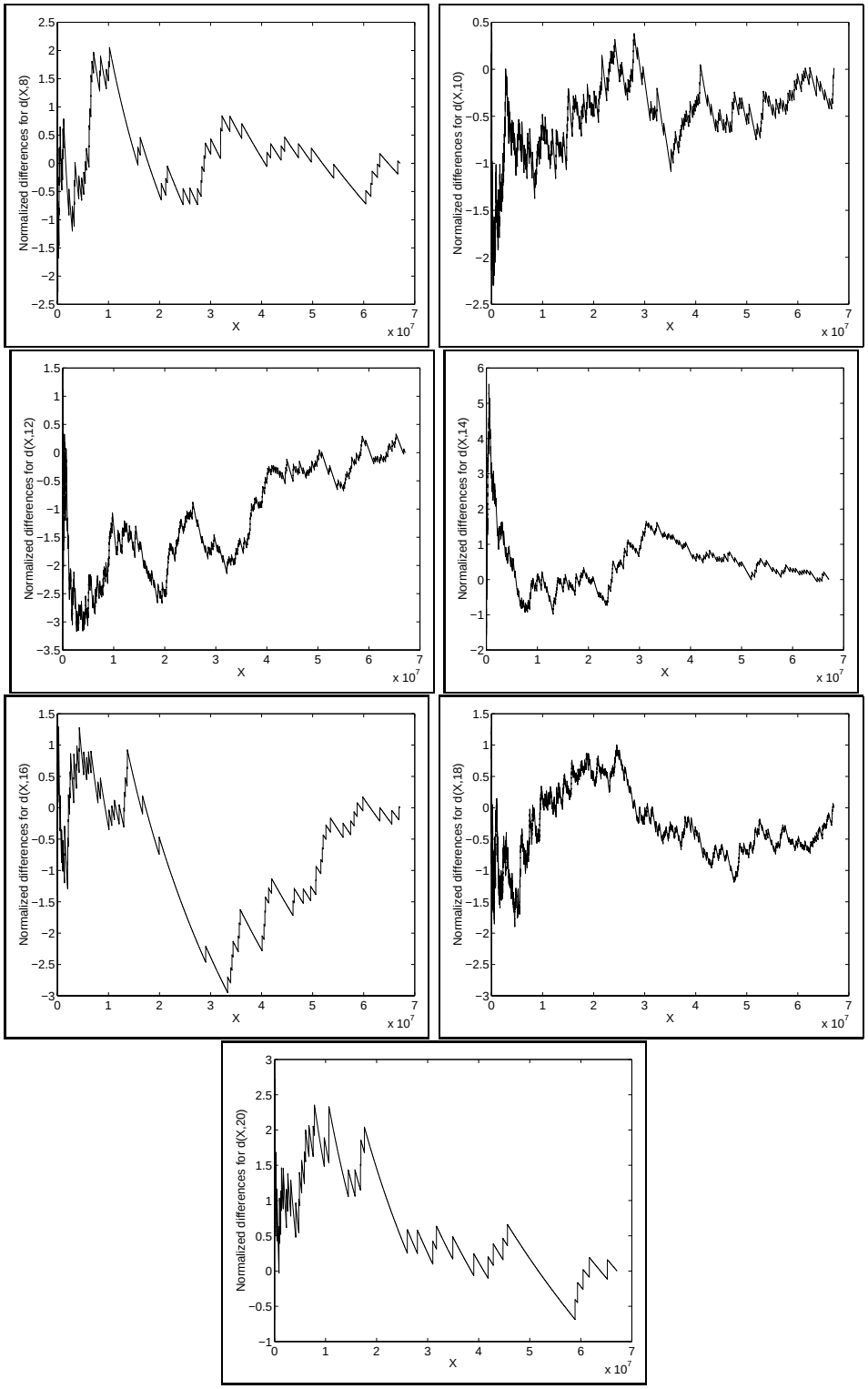

Normalized differences for c(X,18)
X
x 10^7


Normalized differences for c(X,20)
X
x 10^7

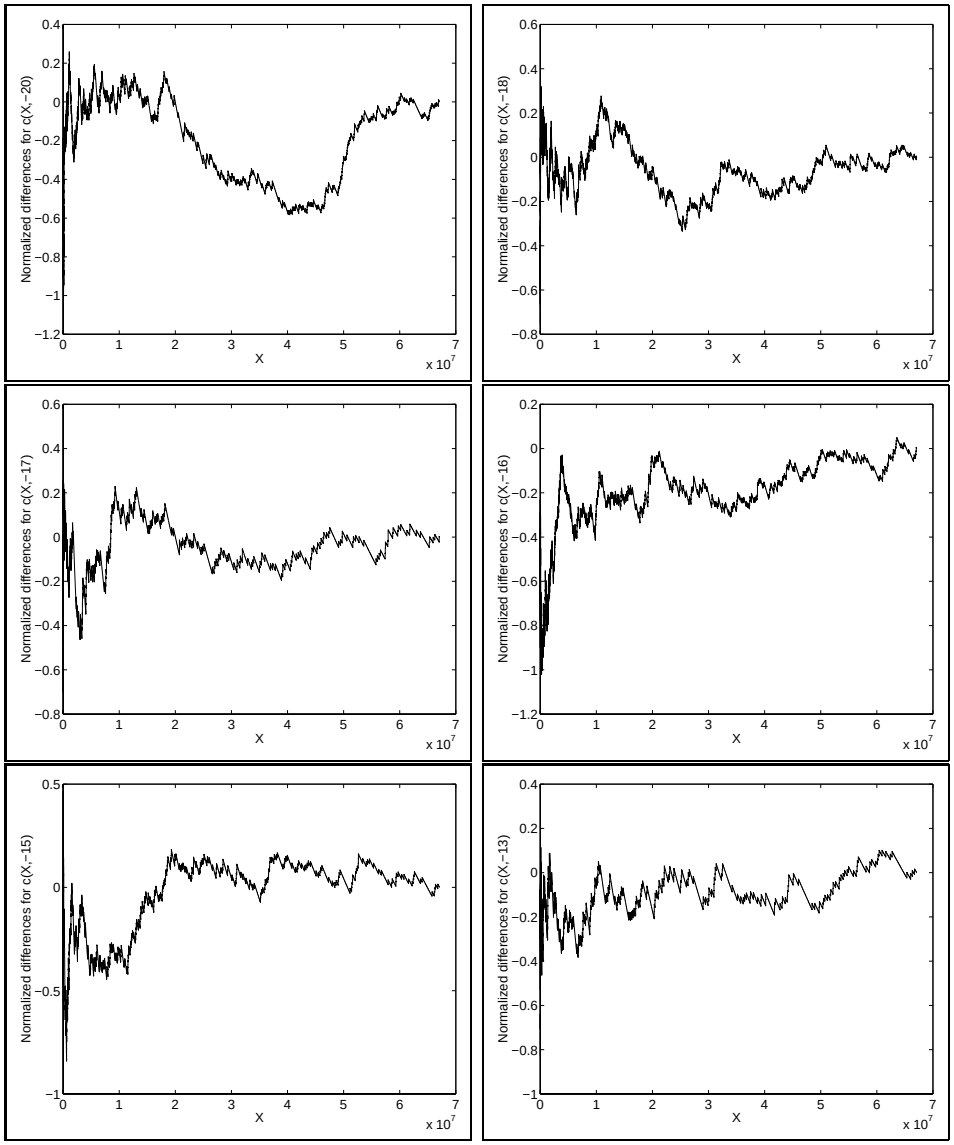
Normalized Differences for $\theta_f(x, a)$

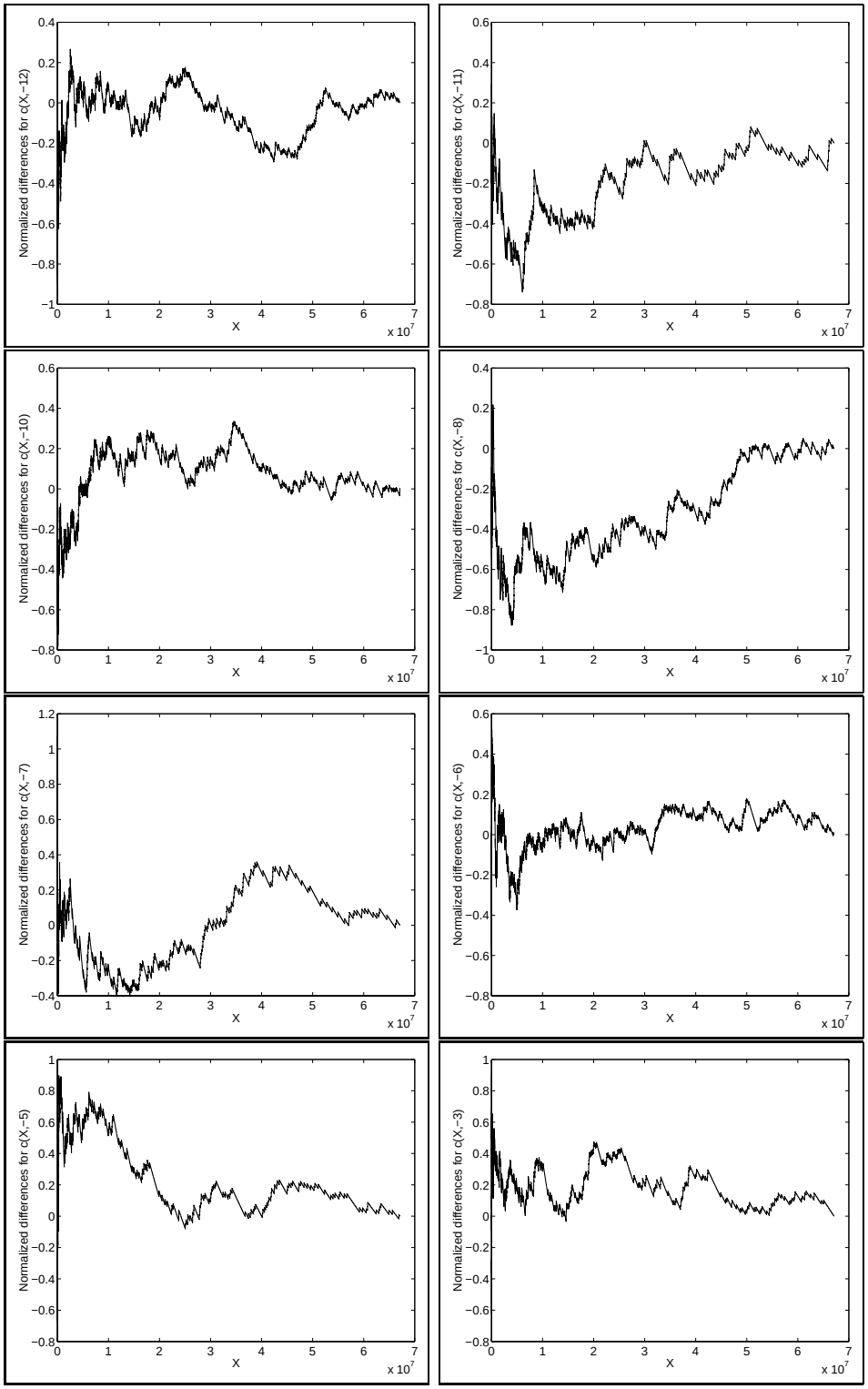


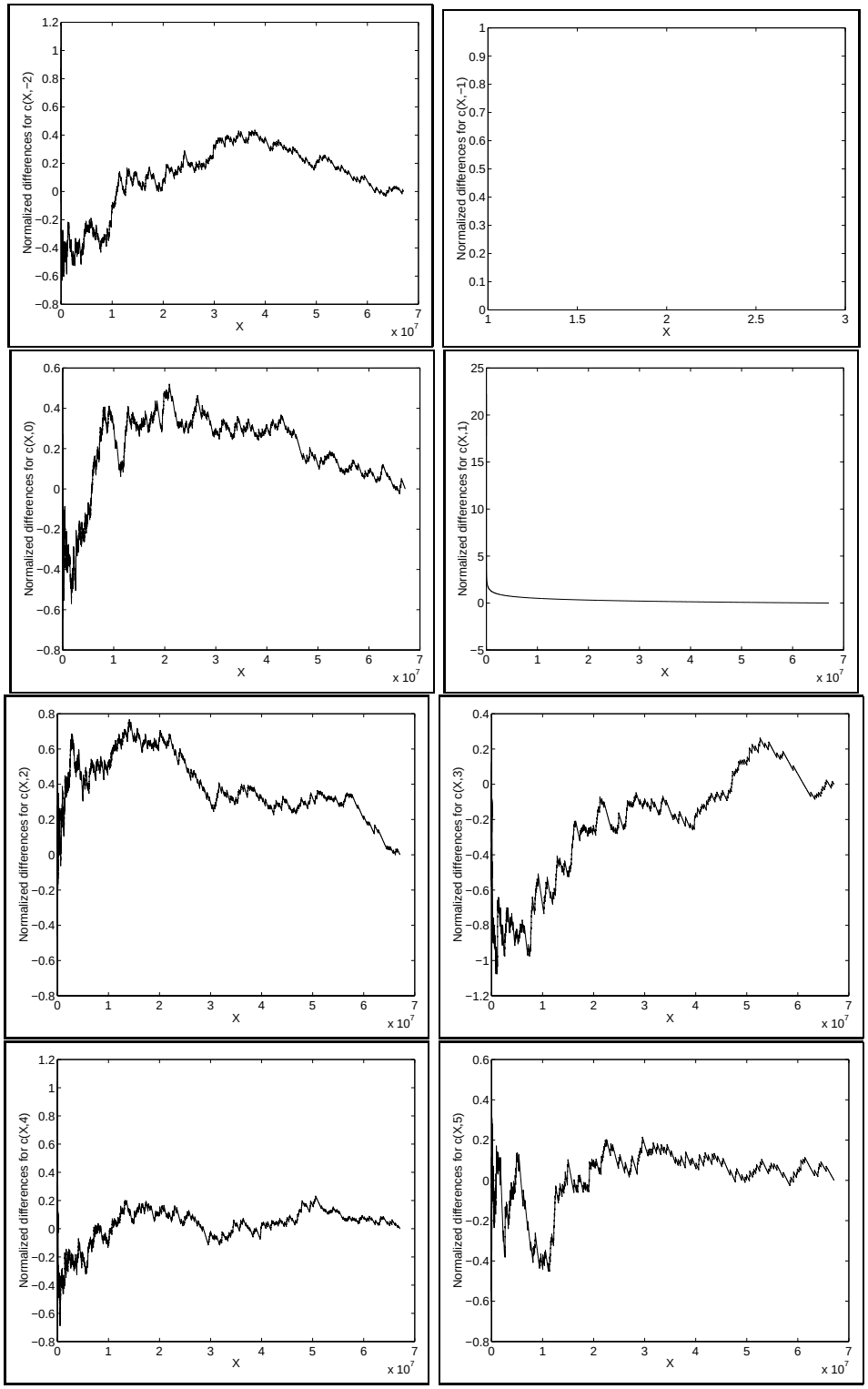

Normalized differences for d(X,−8)
Normalized differences for d(X,−6)
Normalized differences for d(X,−4)
Normalized differences for d(X,−2)
Normalized differences for d(X,0)
Normalized differences for d(X,2)
Normalized differences for d(X,4)
Normalized differences for d(X,6)
X
x 10^7

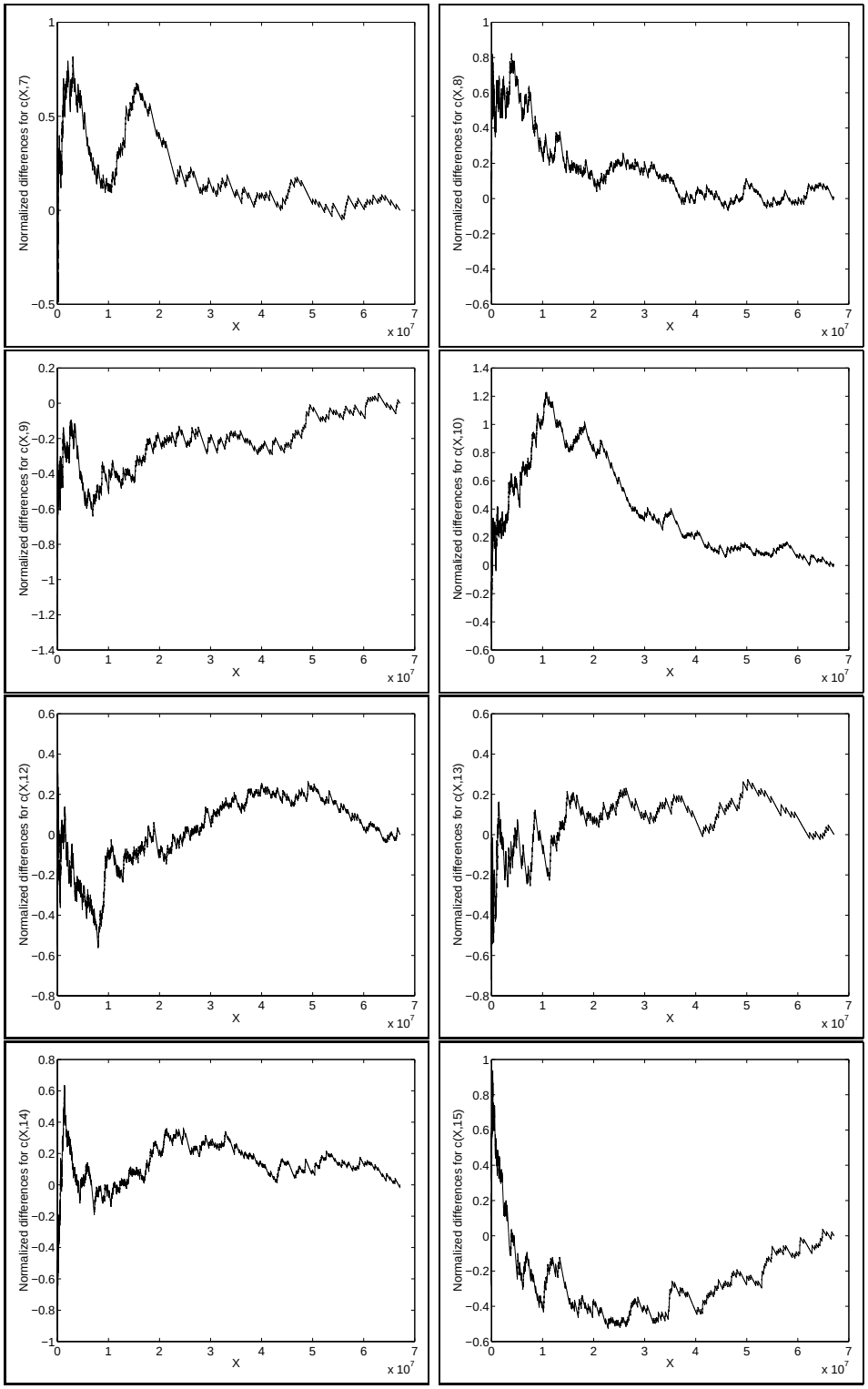

Normalized differences for d(X,8)
Normalized differences for d(X,10)
Normalized differences for d(X,12)
Normalized differences for d(X,14)
Normalized differences for d(X,16)
Normalized differences for d(X,18)
Normalized differences for d(X,20)
X
x 10^7

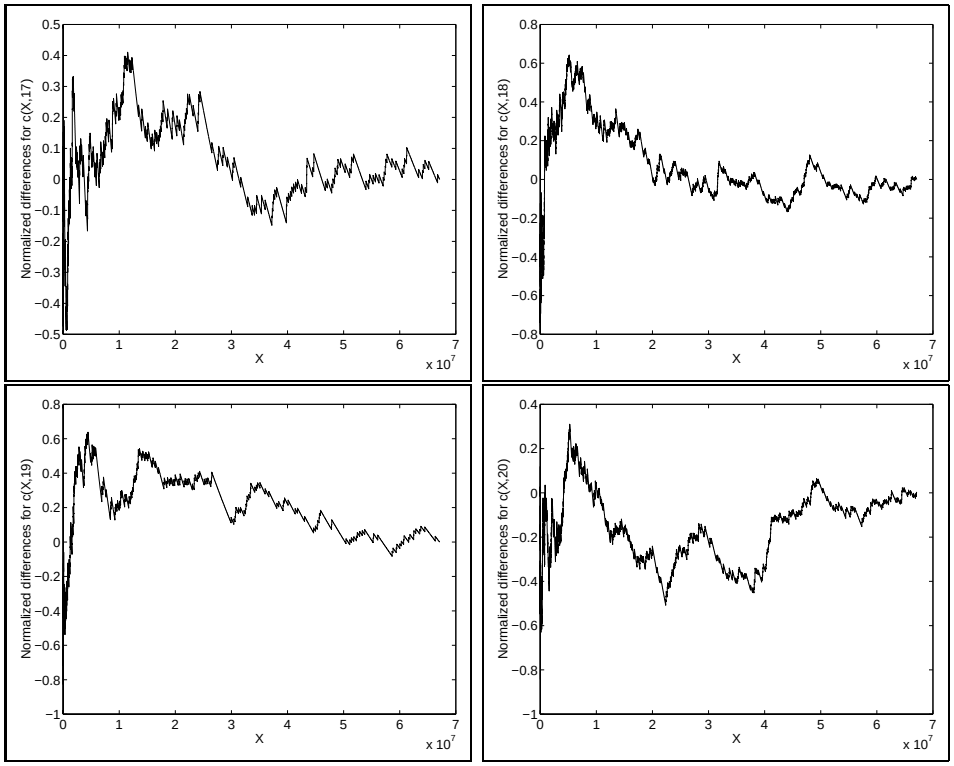
Normalized Differences for $\pi_g(x, a)$



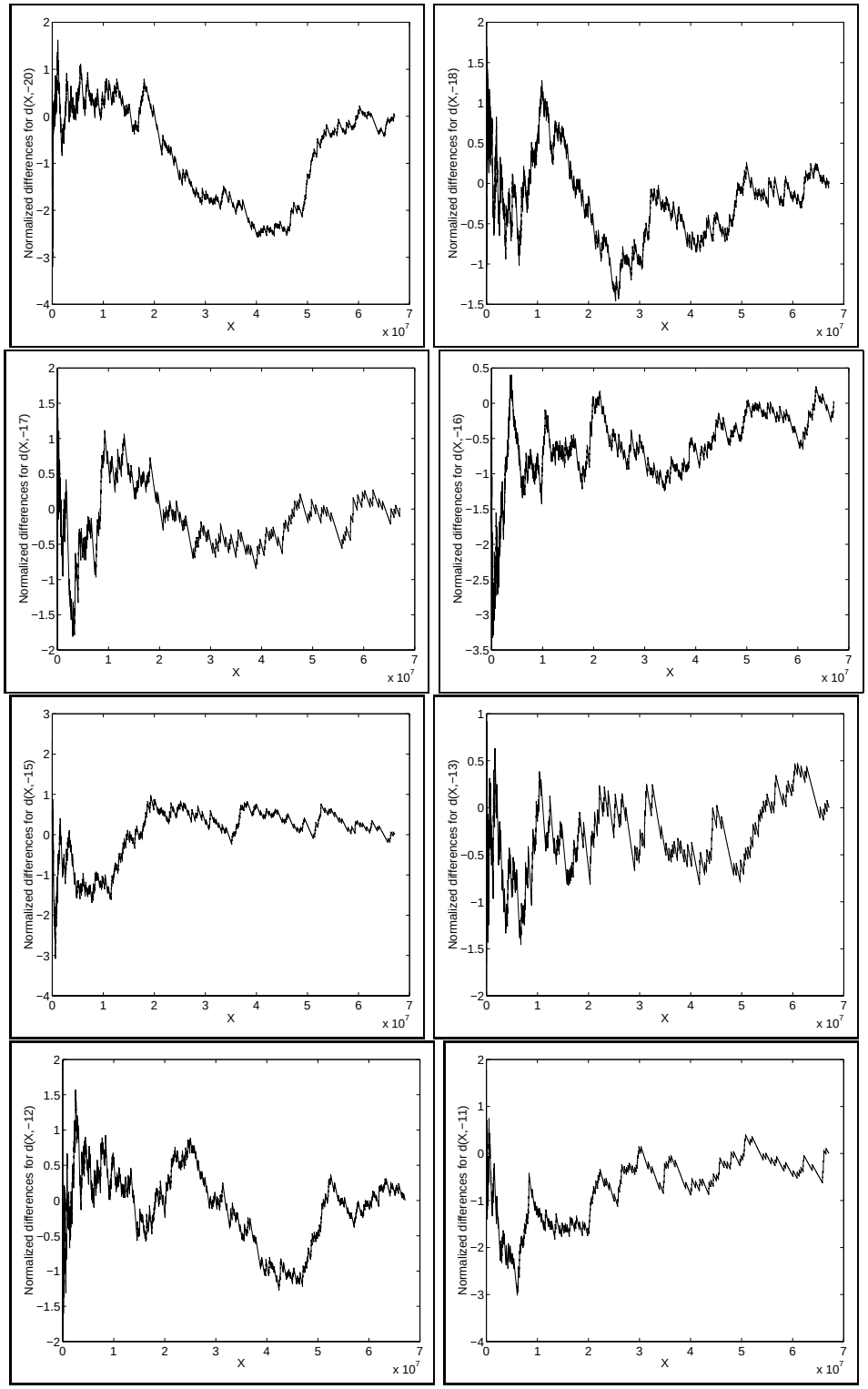

Normalized differences for c(X, −12)
Normalized differences for c(X, −11)
Normalized differences for c(X, −10)
Normalized differences for c(X, −8)
Normalized differences for c(X, −7)
Normalized differences for c(X, −6)
Normalized differences for c(X, −5)
Normalized differences for c(X, −3)
X
x 10^7

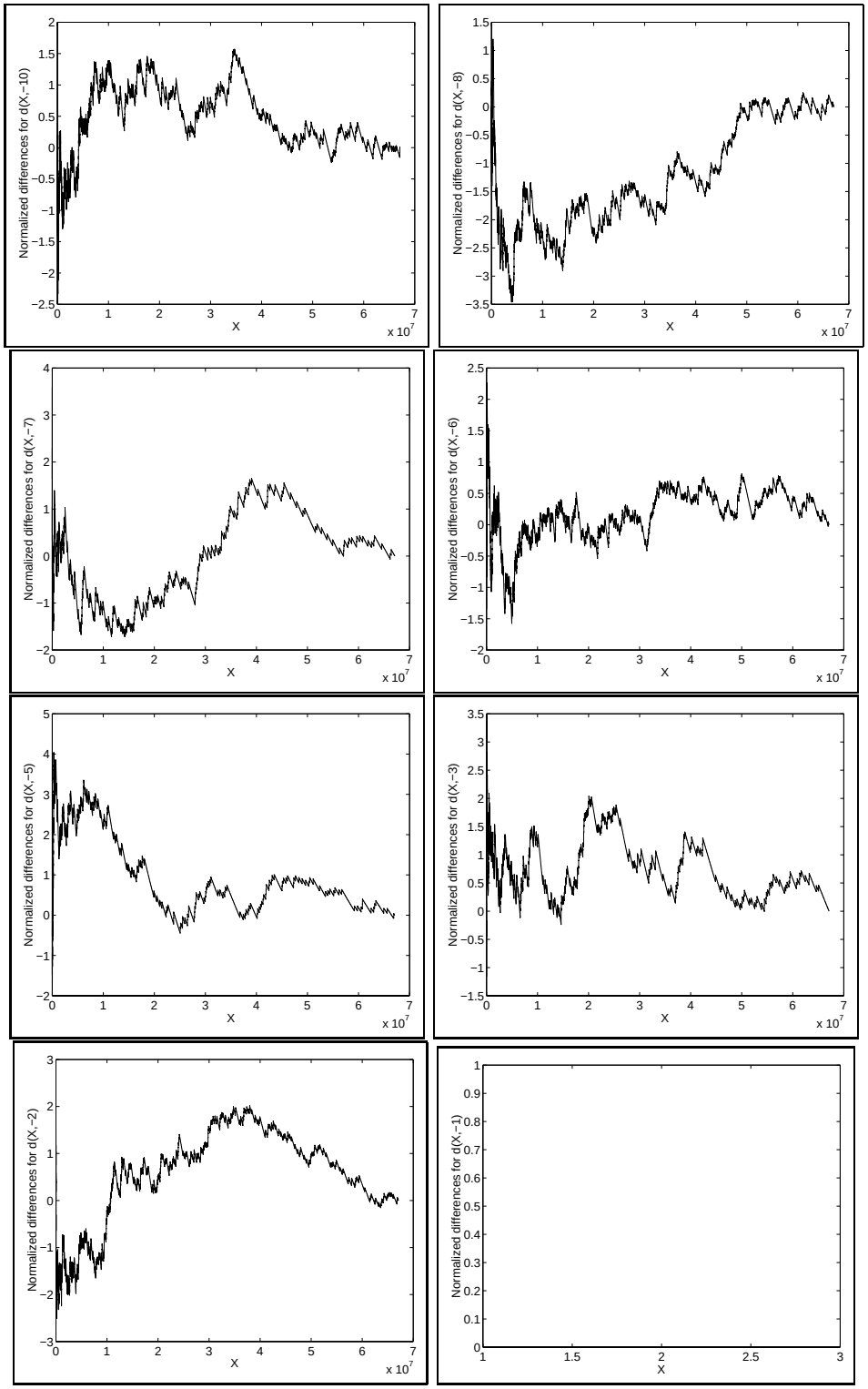

Normalized differences for c(X,−2)
Normalized differences for c(X,−1)
Normalized differences for c(X,0)
Normalized differences for c(X,1)
Normalized differences for c(X,2)
Normalized differences for c(X,3)
Normalized differences for c(X,4)
Normalized differences for c(X,5)
X
x 10^7

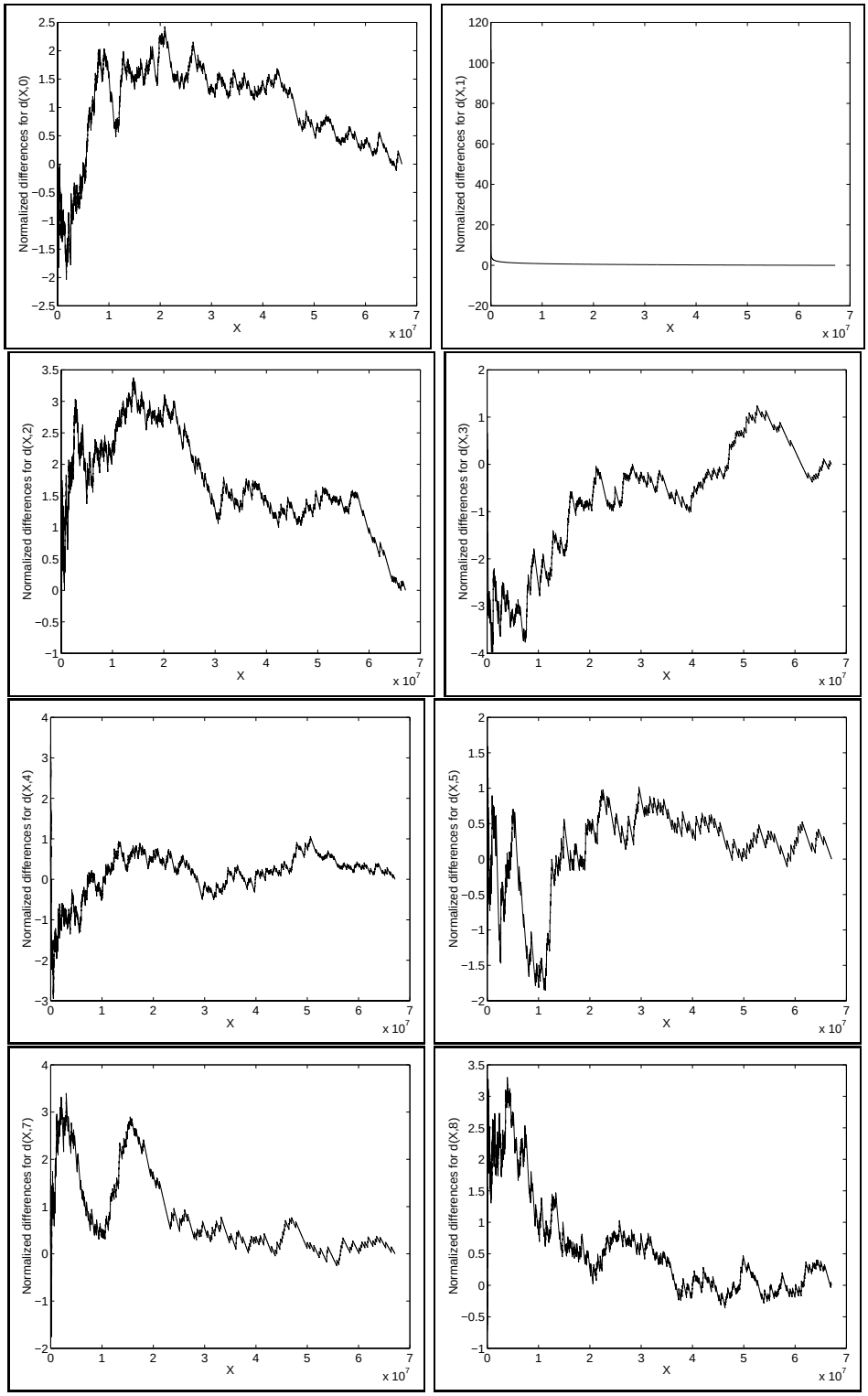

Normalized differences for c(X,7)
Normalized differences for c(X,8)
Normalized differences for c(X,9)
Normalized differences for c(X,10)
Normalized differences for c(X,12)
Normalized differences for c(X,13)
Normalized differences for c(X,14)
Normalized differences for c(X,15)
X
x 10^7

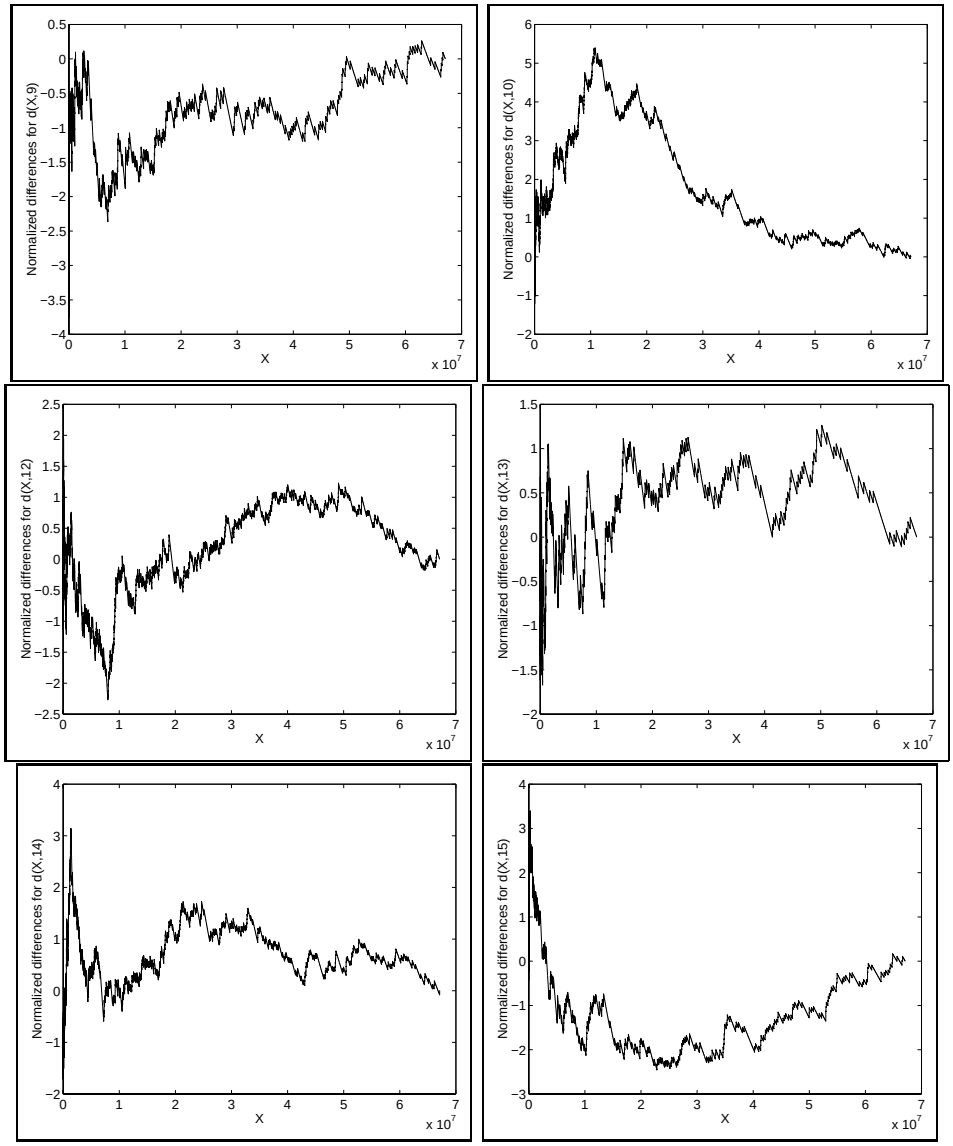


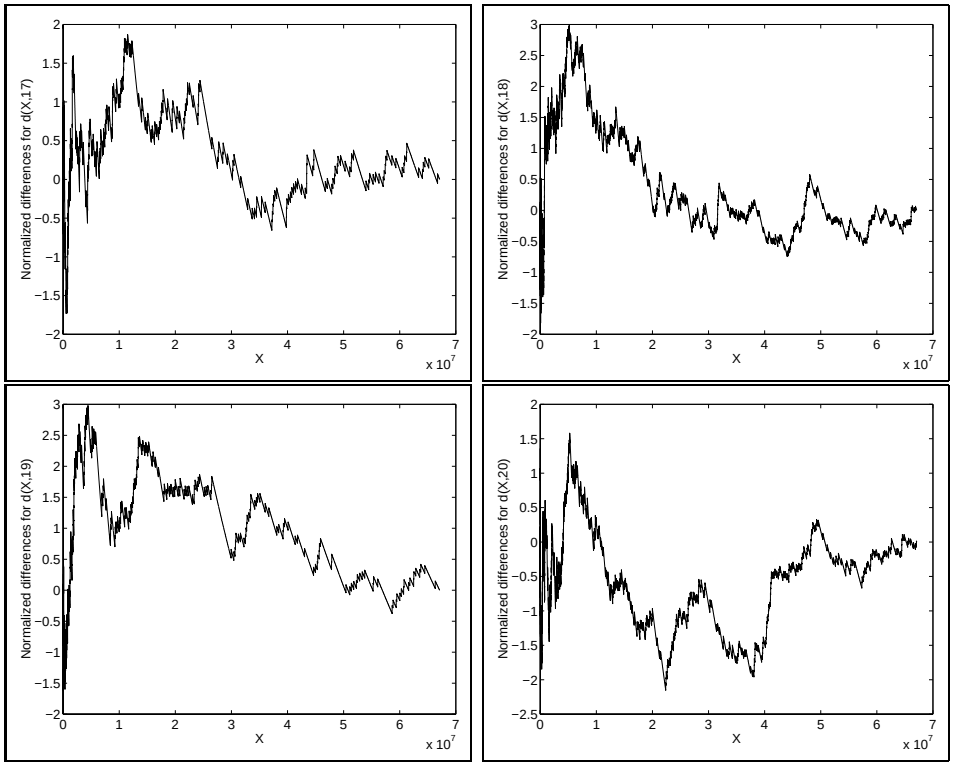
Normalized Differences for $\theta_g(x, a)$


Normalized differences for d(X,−20)
Normalized differences for d(X,−18)
Normalized differences for d(X,−17)
Normalized differences for d(X,−16)
Normalized differences for d(X,−15)
Normalized differences for d(X,−13)
Normalized differences for d(X,−12)
Normalized differences for d(X,−11)
X
x 10^7


Normalized differences for d(X,−10)
Normalized differences for d(X,−8)
Normalized differences for d(X,−7)
Normalized differences for d(X,−6)
Normalized differences for d(X,−5)
Normalized differences for d(X,−3)
Normalized differences for d(X,−2)
Normalized differences for d(X,−1)
X
x 10^7


Normalized differences for d(X,0)
Normalized differences for d(X,1)
Normalized differences for d(X,2)
Normalized differences for d(X,3)
Normalized differences for d(X,4)
Normalized differences for d(X,5)
Normalized differences for d(X,7)
Normalized differences for d(X,8)
X
x 10^7


Normalized differences for d(X,9)
Normalized differences for d(X,10)
Normalized differences for d(X,12)
Normalized differences for d(X,13)
Normalized differences for d(X,14)
Normalized differences for d(X,15)
X
x 10^7


Normalized differences for d(X,17)
Normalized differences for d(X,18)
Normalized differences for d(X,19)
Normalized differences for d(X,20)
X
x 10^7

7. Statistical Analysis

Do the numbers above support the hypothesis that $c_f(x,a)$ tends to a constant as $x \to \infty$? Do they support the hypothesis that

$$d_f(x,a) \;=\; c_{f,a} \;+\; \mathbf{O}(x^{-\frac{1}{2}+\epsilon})?$$

Do they support the hypothesis that

$$\limsup_{x\longrightarrow\infty} \frac{\pi_f(x,a) - c_{f,a}\pi_{1/2}(x)}{x^{1/4}(\log\log x)^{1/2}(\log x)^{-1/2}} \;=\; \sqrt{2c_f(a)}?$$

The methods of statistics suggest several ways in which this question can be formulated precisely and answered. We will discuss one such method of inference below.

For $i = 1, 2, \cdots, 128$, we set

$$c_i \;=\; \frac{\#\{(i-1)\times 2^{19} \;<\; p \;<\; i\times 2^{19} \;:\; a_f(p) = a\}}{\sum_{(i-1)\times 2^{19}<p<i\times 2^{19}} \frac{1}{2\sqrt{p}}}.$$

We compute the mean

$$m \;=\; \frac{1}{128}\sum_{i=1}^{128} c_i$$

and the variance

$$s^2 \;=\; \frac{1}{128}\sum_{i=1}^{128}(c_i - m)^2.$$

As our probability model is that of Bernoulli trials, we expect in the limit that our random variables obey the normal distribution. Here, we are defining for each prime p, the random variable

$$X_p \;=\; \begin{cases} \frac{c_f(a)}{2\sqrt{p}} & p \text{ is prime} \\ 0 & else. \end{cases}$$

It has mean

$$\mu \;=\; c_{f,a}$$

and standard deviation

$$\sigma \;=\; \sqrt{c_{f,a}}.$$

The latter is computed by considering the expectation of the random variable $(X-\mu)^2$ and taking the square root. We compare this against the mean and standard deviation obtained from the sample. In the limit, our random variable will obey the normal distribution and so we expect with probability .95 that we have (see [5], p. 224)

$$|m - \mu| \;<\; \frac{1.965}{\sqrt{128}} \approx 0.1736.$$

In the following tables we list the values of a, m, s^2, μ and $m-\mu$ where in the first table the value of μ is taken to be the constant conjectured by Lang and Trotter and the $c_i's$ are as

above. In the second table we again take μ to be the Lang-Trotter constant but we average over the values of $d_f(x,a)$ to calculate the c_i's. In the third and fourth tables we take the value of μ to be $\pi_g(2^{26},a)/\pi_{1/2}(2^{26})$ and $\theta_g(2^{26},a)/\theta_{1/2}(2^{26})$ respectively and we compute the c_i's by averaging over the values of $c_g(x,a)$ and $d_g(x,a)$ respectively. The fifth and sixth tables are the same as the third and fourth except that we replace the form g associated to $X_0(11)$ with the form $f \in S_2(512)$ with an inner twist discussed in section 5.

a	m	s^2	μ	$\lvert m-\mu \rvert$
-20	0.99627	0.40450	0.98478	0.01149
-18	1.14617	0.36992	0.98478	0.16139
-17	0.50846	0.17033	0.49920	0.00926
-16	0.98814	0.30204	0.98478	0.00336
-15	0.67419	0.22048	0.49920	0.17499
-13	0.51826	0.17902	0.49920	0.01906
-12	1.16646	0.34436	0.98478	0.18168
-11	0.47176	0.21842	0.45382	0.01795
-10	0.96303	0.36142	0.98478	0.02175
-8	0.98867	0.50298	0.98478	0.00389
-7	0.47775	0.17512	0.49920	0.02145
-6	1.15768	0.33804	0.98478	0.17290
-5	0.43565	0.15880	0.49920	0.06355
-3	0.59350	0.22616	0.49920	0.09430
-2	0.99533	0.27580	0.98478	0.01055
-1	0.52110	0.16805	0.49920	0.02190
0	1.17521	0.40080	1.31376	0.13855
1	0.00023	0.00000	0.00000	0.00023
2	0.92817	0.26719	0.98478	0.05661
3	0.61624	0.23759	0.49920	0.11705
4	1.04796	0.31049	0.98478	0.06318
5	0.52527	0.17395	0.49920	0.02608
7	0.46960	0.15900	0.49920	0.02959
8	1.04057	0.27476	0.98478	0.05578
9	0.59691	0.18926	0.49920	0.09771
10	0.82848	0.27013	0.98478	0.15630
12	1.18814	0.34932	0.98478	0.20336
13	0.43492	0.12968	0.49920	0.06427
14	0.97936	0.36755	0.98478	0.00542
15	0.61719	0.25162	0.49920	0.11799
17	0.47636	0.17773	0.49920	0.02284
18	1.08606	0.45481	0.98478	0.10128
19	0.44514	0.16162	0.49920	0.05406
20	0.95576	0.29185	0.98478	0.02902

a	m	s^2	μ	$\lvert m-\mu \rvert$
-20	0.99674	0.40449	0.98478	0.01196
-18	1.14635	0.36990	0.98478	0.16157
-17	0.50856	0.17033	0.49920	0.00937
-16	0.98833	0.30202	0.98478	0.00355
-15	0.67414	0.22049	0.49920	0.17494
-13	0.51836	0.17899	0.49920	0.01916
-12	1.16684	0.34426	0.98478	0.18206
-11	0.47195	0.21844	0.45382	0.01814
-10	0.96336	0.36140	0.98478	0.02142
-8	0.98881	0.50296	0.98478	0.00403
-7	0.47780	0.17511	0.49920	0.02139
-6	1.15781	0.33803	0.98478	0.17303
-5	0.43607	0.15898	0.49920	0.06313
-3	0.59357	0.22618	0.49920	0.09438
-2	0.99511	0.27584	0.98478	0.01033
-1	0.52111	0.16807	0.49920	0.02191
0	1.17531	0.40081	1.31376	0.13844
1	0.00004	0.00000	0.00000	0.00004
2	0.92848	0.26728	0.98478	0.05630
3	0.61613	0.23764	0.49920	0.11693
4	1.04746	0.31066	0.98478	0.06268
5	0.52533	0.17394	0.49920	0.02614
7	0.46973	0.15902	0.49920	0.02946
8	1.04076	0.27484	0.98478	0.05598
9	0.59709	0.18918	0.49920	0.09789
10	0.82877	0.27024	0.98478	0.15601
12	1.18821	0.34933	0.98478	0.20343
13	0.43508	0.12969	0.49920	0.06411
14	0.97975	0.36750	0.98478	0.00504
15	0.61727	0.25162	0.49920	0.11807
17	0.47633	0.17775	0.49920	0.02287
18	1.08624	0.45472	0.98478	0.10146
19	0.44526	0.16160	0.49920	0.05394
20	0.95609	0.29186	0.98478	0.02869

a	m	s^2	μ	$\lvert m-\mu \rvert$
-20	0.99627	0.40450	0.96715	0.02912
-18	1.14617	0.36992	1.13535	0.01082
-17	0.50846	0.17033	0.50269	0.00577
-16	0.98814	0.30204	0.95377	0.03437
-15	0.67419	0.22048	0.66707	0.00712
-13	0.51826	0.17902	0.50651	0.01175
-12	1.16646	0.34436	1.15829	0.00817
-11	0.47176	0.21842	0.44917	0.02259
-10	0.96303	0.36142	0.97479	0.01177
-8	0.98867	0.50298	0.93083	0.05784
-7	0.47775	0.17512	0.47593	0.00182
-6	1.15768	0.33804	1.16211	0.00442
-5	0.43565	0.15880	0.46637	0.03073
-3	0.59350	0.22616	0.61928	0.02578
-2	0.99533	0.27580	1.00538	0.01004
-1	0.52110	0.16805	0.53518	0.01409
0	1.17521	0.40080	1.20798	0.03277
1	0.00023	0.00000	0.00382	0.00359
2	0.92817	0.26719	0.99773	0.06956
3	0.61624	0.23759	0.57914	0.03710
4	1.04796	0.31049	1.04743	0.00053
5	0.52527	0.17395	0.52562	0.00035
7	0.46960	0.15900	0.49695	0.02735
8	1.04057	0.27476	1.07610	0.03553
9	0.59691	0.18926	0.56576	0.03114
10	0.82848	0.27013	0.90025	0.07177
12	1.18814	0.34932	1.18696	0.00119
13	0.43492	0.12968	0.44152	0.00660
14	0.97936	0.36755	1.00155	0.02220
15	0.61719	0.25162	0.58870	0.02849
17	0.47636	0.17773	0.48357	0.00721
18	1.08606	0.45481	1.10094	0.01489
19	0.44514	0.16162	0.47020	0.02505
20	0.95576	0.29185	0.93083	0.02492

a	m	s^2	μ	$\lvert m-\mu \rvert$
-20	0.99674	0.40449	0.97527	0.02147
-18	1.14635	0.36990	1.13850	0.00785
-17	0.50856	0.17033	0.50496	0.00360
-16	0.98833	0.30202	0.96588	0.02245
-15	0.67414	0.22049	0.67179	0.00235
-13	0.51836	0.17899	0.51052	0.00784
-12	1.16684	0.34426	1.16532	0.00152
-11	0.47195	0.21844	0.45387	0.01808
-10	0.96336	0.36140	0.97945	0.01609
-8	0.98881	0.50296	0.94219	0.04661
-7	0.47780	0.17511	0.47636	0.00145
-6	1.15781	0.33803	1.16239	0.00459
-5	0.43607	0.15898	0.46222	0.02615
-3	0.59357	0.22618	0.61446	0.02089
-2	0.99511	0.27584	1.00813	0.01302
-1	0.52111	0.16807	0.53629	0.01518
0	1.17531	0.40081	1.20959	0.03428
1	0.00004	0.00000	0.00049	0.00045
2	0.92848	0.26728	0.99077	0.06229
3	0.61613	0.23764	0.58952	0.02661
4	1.04746	0.31066	1.04786	0.00039
5	0.52533	0.17394	0.52749	0.00216
7	0.46973	0.15902	0.49256	0.02283
8	1.04076	0.27484	1.06787	0.02711
9	0.59709	0.18918	0.57537	0.02172
10	0.82877	0.27024	0.89265	0.06387
12	1.18821	0.34933	1.19094	0.00273
13	0.43508	0.12969	0.44406	0.00898
14	0.97975	0.36750	1.00462	0.02488
15	0.61727	0.25162	0.58787	0.02940
17	0.47633	0.17775	0.48405	0.00773
18	1.08624	0.45472	1.10289	0.01665
19	0.44526	0.16160	0.46985	0.02459
20	0.95609	0.29186	0.93788	0.01821

a	m	s^2	μ	$\lvert m-\mu \rvert$
-20	0.13343	0.03575	0.12042	0.01302
-18	1.30702	0.54208	1.31502	0.00800
-16	0.10638	0.03747	0.10512	0.00126
-14	0.60083	0.19552	0.58870	0.01214
-12	0.66774	0.23311	0.68618	0.01843
-10	0.57688	0.19506	0.56385	0.01303
-8	0.11057	0.04319	0.11850	0.00794
-6	2.26261	0.92158	2.31848	0.05587
-4	0.09563	0.03100	0.09748	0.00185
-2	0.56007	0.20499	0.57723	0.01716
0	1.44901	0.46415	1.44499	0.00402
2	0.48507	0.18458	0.44917	0.03590
4	0.09224	0.02648	0.09939	0.00715
6	2.39904	0.99925	2.33569	0.06335
8	0.12521	0.04350	0.12615	0.00094
10	0.63376	0.22077	0.61164	0.02212
12	0.74263	0.27581	0.69382	0.04881
14	0.54588	0.19437	0.56385	0.01797
16	0.15833	0.06484	0.14717	0.01115
18	1.31625	0.45924	1.29399	0.02226
20	0.08229	0.02543	0.09175	0.00946

a	m	s^2	μ	$\lvert m-\mu \rvert$
-20	0.13350	0.03574	0.12482	0.00868
-18	1.30726	0.54214	1.31611	0.00884
-16	0.10636	0.03748	0.10340	0.00296
-14	0.60114	0.19546	0.59792	0.00322
-12	0.66787	0.23310	0.68652	0.01865
-10	0.57681	0.19506	0.56479	0.01202
-8	0.11058	0.04319	0.11724	0.00666
-6	2.26279	0.92154	2.32379	0.06100
-4	0.09571	0.03100	0.09801	0.00230
-2	0.56032	0.20506	0.57600	0.01568
0	1.44913	0.46412	1.44973	0.00060
2	0.48503	0.18459	0.45504	0.02999
4	0.09227	0.02648	0.09788	0.00561
6	2.39955	0.99895	2.35718	0.04236
8	0.12530	0.04351	0.12727	0.00197
10	0.63387	0.22074	0.61921	0.01466
12	0.74277	0.27576	0.70491	0.03786
14	0.54642	0.19469	0.56109	0.01466
16	0.15836	0.06483	0.14852	0.00984
18	1.31642	0.45914	1.30352	0.01290
20	0.08226	0.02543	0.08944	0.00718

References

[1] P. Bayer and J. Gonzalez, On the Hasse-Witt invariants of modular curves, Experimental Math., 6(1997), 57-76.

[2] P. Chiu, Transforms, finite fields and fast multiplication, Math. Magazine, vol. $\not\triangleleft 3$, no. $\not\trianglelefteq$ (1990).

[3] H. Cramér, Prime numbers and probability, Proc. 8th Scand. Math. Congr. Stockholm, 1934, pp. 1-9.

[4] W. Feller, An introduction to probability theory and its applications, Volume 1, 3rd ed., Wiley, New York, 1968.

[5] J. Freund, Modern elementary statistics, 3rd ed., Prentice Hall, Englewood Cliffs, 1967.

[6] S. Lang and H. Trotter, Frobenius distribution in GL_2 extensions, Lecture Notes in Mathematics 504, Springer Verlag, 1976.

[7] V. Kumar Murty, Frobenius distributions and Galois representations, in: Automorphic forms, Automorphic representations and Arithmetic, ed. R. Doran, Z.-L. Dou, G. Gilbert, pp. 193-211, Proc. Symp. Pure Math Volume 66, Part 1, AMS, Providence, 1999.

Department of Mathematics
The Pennsylvania State University
University Park, PA 16802, USA
klj@math.psu.edu

Department of Mathematics
University of Toronto
100 St. George Street
Toronto, ON M5S 3G3, CANADA
murty@math.toronto.edu