# Multi-Linearity Self-Testing with Relative Error

Frédéric Magniez*

Université Paris–Sud, 91405 Orsay, France
magniez@lri.fr, http://www.lri.fr/~magniez

**Abstract.** We investigate self-testing programs with relative error by allowing error terms proportional to the function to be computed. Until now, in numerical computation, error terms were assumed to be either constant or proportional to the $p$-th power of the magnitude of the input, for $p \in [0, 1)$. We construct new self-testers with relative error for real-valued multi-linear functions defined over finite rational domains. The existence of such self-testers positively solves an open question in [KMS99]. Moreover, our self-testers are very efficient: they use few queries and simple operations.

**Keywords** — Program verification, approximation error, self–testing programs, robustness and stability of functional equations.

## 1 Introduction

It is not easy to write a program $P$ to compute a real-valued function $f$. By definition of floating point computations, a program $P$ can only compute an approximation of $f$. The succession of inaccuracies in computational operations could be significant. Moreover once $P$ is implemented it is more difficult to verify its correctness, *i.e.* that $P(x)$ is a good approximation of $f(x)$ for all valid inputs $x$. In a good approximation one would like the significant figures to be correct. This leads us to the notion of relative error. If $a$ is a real number and $\hat{a}$ is its approximation, then the quantity $\theta \stackrel{\text{def}}{=} |\hat{a} - a|/a$ is called the *relative error* of the approximation.

In recent years, several notions were developed to address the software correctness problem. Here we focus on the following scenario. First, the program to be tested is viewed as a black box, *i.e.* we can only query it on some inputs. Second, we want a very efficient testing procedure. In particular, a test should be more efficient than any known correct implementation. For exact computation, program checking [Blu88, BK95], self-testing programs [BLR93], and self-correcting programs [BLR93, Lip91] were developed in the early 90's. A *program checker* for $f$ verifies whether the program $P$ computes $f$ on a particular input $x$; a *self-tester* for $f$ verifies whether the program $P$ is correct on most inputs; and a *self-correcting program* for $f$ uses a program $P$, which is correct on most

* Partially supported by a CNRS–Conicyt'98 Project, ESPRIT Working Group RAND2 No. 21726, and Franco-Hungarian bilateral project Balaton No. 99013.

inputs, to compute $f$ correctly everywhere with high probability. Let us insist that checkers, self-testers and self-correctors can only use the program $P$ as a black box, and are required to be different and simpler than any known implementation of $f$ (see [BK95] for a formal definition). In this context, results on testing linear functions and polynomials have theoretical implications for probabilistically checkable proofs [ALM+92, AS92] and in approximation theory. For a survey see [Bab93].

Let us recall the problem of linearity testing which has been fundamental in the development of testers [BLR93]. Given a program $P$ which computes a function from one Abelian group $G$ into another group, we want to verify that $P$ computes a homomorphism on most inputs in $G$. The Blum-Luby-Rubinfeld linearity test is based on the linearity property $f(x+y) = f(x) + f(y)$, for all $x, y \in G$, which is satisfied when $f$ is a homomorphism. The test consists in verifying the previous linearity equation on random instances. More precisely, it checks for random inputs $x, y \in G$ that $P(x+y) = P(x)+P(y)$. If the probability of failing the linearity test is small, then $P$ computes a homomorphism except on a small fraction of inputs. This property of the linearity equation is usually called the *robustness* of the linearity equation. This term was defined in [RS96] and studied in [Rub99]. The analysis of the test is due to Coppersmith [Cop89]. It consists in correcting $P$ by querying it on few queries. Let $g$ be the function which takes at $x$ the majority of the votes $(P(x+y) - P(y))$, for all $y \in G$. When the failure probability in the linearity test is small, majority turns out to quasi-unanimity, $g$ equals $P$ on a large fraction of inputs, and $g$ is linear. This idea of *property testing* has been recently formalized and extended to testing graph properties in [GGR96, GR97].

These notions of testing were extended to approximate computation with absolute error for self-testers/correctors [GLR+91] and for checkers [ABCG93]. In [GLR+91] Gemmel et al. studied only functions defined over algebraically closed domains. Ergün, Ravi Kumar, and Rubinfeld [EKR96] initiated and solved the problem of self-testing with absolute error for linear functions, polynomials, and additive functions defined over rational domains. Rational domains were first considered by Lipton [Lip91] and these are the sets $\mathcal{D}_{n,s} \stackrel{\text{def}}{=} \{i/s : |i| \leq n, i \in \mathbb{Z}\}$, for some integer $n \geq 1$ and real $s > 0$. In these past works the absolute error of the approximation $\hat{a}$ of $a$ is defined by $\varepsilon \stackrel{\text{def}}{=} |\hat{a} - a|$. In this approximate context the linearity testing problem consists now in verifying that a given program $P$ computes approximately a real linear function over $\mathcal{D}_{n,s}$. To allow absolute error in the computation of $P$, the approximate linearity test consists in verifying that $|P(x+y) - P(x) - P(y)| \leq \varepsilon$, for random $x, y \in \mathcal{D}_{n,s}$ and some fixed $\varepsilon > 0$. Then the analysis is very similar to that of the exact case. Since the majority is not adapted to approximate computation, it is replaced by the median. Moreover both the closeness of $g$ to $P$ and the linearity of $g$ are approximated. Therefore we need a second stage which consists in proving the *local stability* of the linearity equation for absolute error, that is, every function satisfying $|f(x+y) - f(x) - f(y)| \leq \varepsilon$, for all $x, y \in \mathcal{D}_{n,s}$, is close to a perfectly linear function. This part is a well-studied problem in mathematics for several kinds of error terms when $x$ and

$y$ describe a group like $\mathbb{Z}$. It corresponds to the study of Hyers-Ulam stability. The stability problem is due to Ulam and was first solved in the absolute error case in 1941 by Hyers [Hye41]. For a survey of Hyers-Ulam stability see [For95, HR92].

Using elegant techniques of Hyers-Ulam stability theory, Kiwi et al. extended a part of [EKR96]'s work for non-constant error terms [KMS99]. They considered error terms proportional in every input $x$ to $|x|^p$, for any $0 \leq p < 1$, that is, they considered computations where inaccuracies depend on the size of the values involved in the calculations. This model corresponds to many practical situations. Among other things, they show how self-testing whether a program approximately computes a linear function for these errors terms. For this they proved the local stability of the linearity equation using its stability on the whole domain $\mathbb{Z}$ using techniques based on an argument due to Skof [Sko83]. The robustness part is similar to absolute error case, but the set of voters in the median defining $g(x)$ depends on $x$ since big voters may induce big errors for small $x$. Since the linearity equation is unstable for the case $p = 1$ [HŠ92], their work did not lead to self-testers either for the case $p = 1$, which corresponds to linear error terms, or for relative error terms (*i.e.* proportional to the function to be computed) [KMS99, Sect. 5].

In this paper, we investigate the study of approximate self-testing with relative error. Relative error is one of the most important notions in numerical computation. Proving that a program is relatively close to its correct implementation is the challenge of many numerical analysts. We hope to contribute to make self-testers more adapted to numerical computation. In this setting self-testing consists in the following task:

**Problem.** *Given a class of real-valued functions $\mathcal{F}$ defined over a finite domain $D$, and some positive constants $c_1, c_2, \delta_1, \delta_2$, we want a simple and efficient probabilistic algorithm $T$ such that, for any program $P : D \to \mathbb{R}$, which is an oracle for $T$:*

- *if for some $f \in \mathcal{F}$, $\mathbf{Pr}_{x \in D}\left[|P^T(x) - f(x)| > c_1|f(x)|\right] \leq \delta_1$, then $T$ outputs* `PASS` *with high probability;*
- *if for all $f \in \mathcal{F}$, $\mathbf{Pr}_{x \in D}\left[|P^T(x) - f(x)| > c_2|f(x)|\right] > \delta_2$, then $T$ outputs* `FAIL` *with high probability.*

We give a positive answer to this problem for the set of real-valued $d$-linear functions, for any integer $d \geq 1$. This is the first positive answer to this problem in the literature. In particular, we solve some problems in [KMS99] that were mentioned previously. For the sake of brevity and clarity we will consider functions defined over positive integer domains $\mathcal{D}_n^+ = \{i \in \mathbb{N} : 1 \leq i \leq n\}$, for some even integer $n \geq 1$. But all of our results remain valid for more general rational domains.

First we define in Theorem 2 a new probabilistic test for linear functions. It is constructed from a new functional equation for linearity which is robust (Theorem 3) and stable for linear error terms (Theorem 4). We use it to build an approximate self-tester for linear functions which allows linear error terms (Theorem 5). From it we are able to construct the first approximate self-tester

with relative error in the sense of the stated problem (Theorem 6). This self-tester is generalized for multi-linear functions in Theorem 7 using an argument similar to that in [FHS94]. These self-testers are quite surprising since they only use comparisons, additions, and multiplications by powers of 2 (*i.e.* left or right shifts in binary representation). Moreover the number of queries and operations does not depend on $n$.

## 2 Linearity

The linearity test of [KMS99] is based on the linearity equation $f(x+y) = f(x) + f(y)$ which is robust and stable for error terms proportional to $|x|^p$, where $0 \leq p < 1$, but unstable when $p = 1$. More precisely they showed:

**Theorem 1 ([KMS99, Theorem 2])** *Let $0 \leq \delta \leq 1$, $\theta \geq 0$, and $0 \leq p < 1$. If $P : \mathcal{D}_{8n}^{+} \to \mathbb{R}$ is such that*

$$\mathop{\mathbf{Pr}}_{x,y \in \mathcal{D}_{4n}^{+}} [|P(x+y) - P(x) - P(y)| > \theta \, \mathbf{Max}\{x^p, y^p\}] \leq \delta,$$

*then there exists a linear function $l : \mathcal{D}_{n}^{+} \to \mathbb{R}$ such that for $C_p = (1+2^p)/(2-2^p)$,*

$$\mathop{\mathbf{Pr}}_{x \in \mathcal{D}_{n}^{+}} [|P(x) - l(x)| > 17 C_p \theta x^p] \leq O(\sqrt{\delta}).$$

*(If $p = 0$ then the latter inequality holds with $O(\delta)$ in its RHS.)*

**Remark.** In this theorem and in the rest of the paper we only consider uniform probabilities.

For $p = 1$ the statement of this theorem does not hold anymore. Let $\theta > 0$ and $f(x) \stackrel{\text{def}}{=} \theta x \log_2(x+1)$, for all $x > 0$. In [HŠ92] it is shown that $f$ satisfies $|f(x+y) - f(x) - f(y)| \leq 2\theta \, \mathbf{Max}\{x, y\}$, for all $x, y > 0$, but $f$ is not close to any linear function. Hence either the test or the error term has to be modified, but both can not be kept. In [KMS99] the linearity test was unchanged and error terms proportional to $x^p$ were considered, for some $0 \leq p < 1$. In this paper we change the test but keep a linear error term.

All results of this paper are based on the following theorem. It defines a probabilistic test such that the distance of any program to linear functions is upper bounded by a constant times its failure probability on it. Here the distance is not yet relative but it is defined for a linear error term. Let $\mathbf{Med}_{x \in X}(f(x))$ denote the median value of $f : X \to R$ when $x$ ranges over $X$:

$$\mathop{\mathbf{Med}}_{x \in X}(f(x)) \stackrel{\text{def}}{=} \mathbf{Inf} \left\{ a \in \mathbb{R} \, : \, \mathop{\mathbf{Pr}}_{x \in X} [f(x) \geq a] \leq 1/2 \right\}.$$

For every integer $x \geq 1$, let $k_x$ define the number:

$$k_x \stackrel{\text{def}}{=} \mathbf{Min} \left\{ k \in \mathbb{N} : 2^k x \geq \frac{n}{2} \right\}.$$

**Theorem 2** *Let $0 \leq \delta < 1/96$ and $\theta \geq 0$. If $P : \mathcal{D}_{8n}^+ \to \mathbb{R}$ is such that*

$$\mathop{\mathbf{Pr}}_{x,y \in \mathcal{D}_{4n}^+} \left[ |P(2^{k_x}x + y) - 2^{k_x}P(x) - P(y)| > \theta n \right] \leq \delta,$$

*then the linear function $l : \mathcal{D}_n^+ \to \mathbb{R}$, which is defined by*

$$l(n) \stackrel{\text{def}}{=} \mathop{\mathbf{Med}}_{y \in \mathcal{D}_{2n}^+} (P(n + y) - P(y)),$$

*satisfies*

$$\mathop{\mathbf{Pr}}_{x \in \mathcal{D}_n^+} [|P(x) - l(x)| > 32\theta x] \leq 16\delta.$$

The proof of this theorem goes in two parts: the robustness (Theorem 3), and the stability (Theorem 4). Let us give the intuition for this test. When $x \geq n/2$, *i.e.* $x$ is *large*, the test looks like the standard linearity test. But when $x < n/2$, *i.e.* $x$ is *small*, we add a dilation term which amplifies small errors.

### 2.1 Robustness

This part consists in constructing, using $P$, a function $g$ which is not linear, but approximately linear for large inputs, and perfectly homothetic for small inputs. In a sense $g$ approximately corrects the program $P$.

The following theorem sates the existence of such a function $g$. The definition of $g$ is based on the probability test and it consists in performing, for some $x \in \mathcal{D}_{2n}^+$, the median of votes $(P(2^{k_x}x + y) - P(y))/2^{k_x}$ for all $y \in \mathcal{D}_{2n}^+$. If the probability that $P$ fails the test is small, then $g$ satisfies the following theorem.

**Theorem 3 (Robustness)** *Let $0 \leq \delta < 1/96$ and $\theta \geq 0$. If $P : \mathcal{D}_{8n}^+ \to \mathbb{R}$ is such that*

$$\mathop{\mathbf{Pr}}_{x,y \in \mathcal{D}_{4n}^+} \left[ |P(2^{k_x}x + y) - 2^{k_x}P(x) - P(y)| > \theta n \right] \leq \delta,$$

*then the function $g : \mathcal{D}_{2n}^+ \to \mathbb{R}$ which is defined by*

$$g(x) \stackrel{\text{def}}{=} \mathop{\mathbf{Med}}_{y \in \mathcal{D}_{2n}^+} (P(2^{k_x}x + y) - P(y))/2^{k_x},$$

*satisfies*

$$\mathop{\mathbf{Pr}}_{x \in \mathcal{D}_n^+} [|P(x) - g(x)| > 2\theta x] \leq 16\delta, \tag{1}$$

$$\forall x, y \in \{n/2, \ldots, n\}, \quad |g(x + y) - g(x) - g(y)| \leq 6\theta n, \tag{2}$$

$$\forall x \in \mathcal{D}_n^+, \quad g(x) = g(2^{k_x}x)/2^{k_x}. \tag{3}$$

*Proof.* The proof uses standard techniques developed in [BLR93, EKR96, KMS99]. Let us observe first that the function $g$ satisfies

$$g(x) = \begin{cases} \mathop{\mathbf{Med}}_{y \in \mathcal{D}_{2n}^+} (P(x + y) - P(y)), & \text{if } x \geq n/2, \\ g(2^{k_x}x)/2^{k_x}, & \text{otherwise,} \end{cases}$$

and therefore $g$ satisfies (3). Before proving that $g$ also satisfies (1) and (2), we state a useful fact called the *halving principle* [KMS99].

**Fact 1 (Halving principle)** *Let $\Omega$ and $S$ denote finite sets such that $S \subseteq \Omega$, and let $\psi$ be a boolean function defined over $\Omega$. Then for uniform probabilities,*

$$\mathop{\mathbf{Pr}}_{x\in S}[\psi(x)] \leq \frac{|\Omega|}{|S|} \mathop{\mathbf{Pr}}_{x\in\Omega}[\psi(x)].$$

First we show that $g$ is close to $P$ as defined in (1). To simplify notation, let $P_{x,y} = P(2^{k_x}x + y) - P(y) - 2^{k_x}P(x)$. By definition of $g$ we get

$$\mathop{\mathbf{Pr}}_{x\in\mathcal{D}_n^+}[|g(x) - P(x)| > 2\theta x] = \mathop{\mathbf{Pr}}_{x\in\mathcal{D}_n^+}\left[|\mathop{\mathbf{Med}}_{y\in\mathcal{D}_{2n}^+}(P_{x,y})| > 2\theta 2^{k_x}x\right].$$

Notice that Markov's inequality gives a bound on the RHS of this equality:

$$\mathop{\mathbf{Pr}}_{x\in\mathcal{D}_n^+}\left[|\mathop{\mathbf{Med}}_{y\in\mathcal{D}_{2n}^+}(P_{x,y})| > 2\theta 2^{k_x}x\right] \leq 2 \mathop{\mathbf{Pr}}_{x\in\mathcal{D}_n^+, y\in\mathcal{D}_{2n}^+}\left[|P_{x,y}| > 2\theta 2^{k_x}x\right].$$

But $2^{k_x}x \geq n/2$, for all $x \in \mathcal{D}_n^+$, then using the halving principle we get

$$\mathop{\mathbf{Pr}}_{x\in\mathcal{D}_n^+, y\in\mathcal{D}_{2n}^+}\left[|P_{x,y}| > 2\theta 2^{k_x}x\right] \leq \frac{|\mathcal{D}_{4n}^+|^2}{|\mathcal{D}_n^+| \cdot |\mathcal{D}_{2n}^+|} \mathop{\mathbf{Pr}}_{x,y\in\mathcal{D}_{4n}^+}[|P_{x,y}| > \theta n].$$

Therefore $g$ satisfies (1).

Now we prove that $g$ satisfies (2). First we show that for all $c \in \{n/2, \ldots, 2n\}$ the median value $g(c)$ is close to any vote $(P(c+y)-P(y))$ with high probability:

$$\mathop{\mathbf{Pr}}_{y\in\mathcal{D}_{2n}^+}[|g(c) - (P(c+y) - P(y))| > 2\theta n] \leq 16\delta. \tag{4}$$

Note that $k_c = 0$, therefore Markov's inequality implies

$$\mathop{\mathbf{Pr}}_{y\in\mathcal{D}_{2n}^+}[|g(c) - (P(c+y) - P(y))| > 2\theta n] \leq 2 \mathop{\mathbf{Pr}}_{y,z\in\mathcal{D}_{2n}^+}[|P_{c+z,y} - P_{c+y,z}| > 2\theta n].$$

Then one can get inequality (4) using the union bound and the halving principle.

Now let $a$ and $b$ be two integers such that $\frac{n}{2} \leq a, b \leq n$. Let $c$ take on the values $a$, $b$ and $a+b$ in (4), and apply the halving principle to obtain:

$$\left.\begin{array}{r} \mathop{\mathbf{Pr}}_{y\in\mathcal{D}_n^+}[|g(a+b) - (P(a+b+y) - P(y))| > 2\theta n] \\ \mathop{\mathbf{Pr}}_{y\in\mathcal{D}_n^+}[|g(a) - (P(a+y) - P(y))| > 2\theta n] \\ \mathop{\mathbf{Pr}}_{y\in\mathcal{D}_n^+}[|g(b) - (P(b+(a+y)) - P(a+y))| > 2\theta n] \end{array}\right\} \leq 32\delta.$$

Therefore with probability at least $1 - 96\delta > 0$ there exists $y \in \mathcal{D}_n^+$ for which none of these inequalities are satisfied. Pick such a $y$ to obtain inequality (2). □

### 2.2 Stability

In this section we prove that every function $g$ satisfying the conditions of Theorem 3 is close to a perfectly linear one.

**Theorem 4 (Stability)** *Let $\theta' \geq 0$. If $g : \mathcal{D}_{2n}^{+} \to \mathbb{R}$ is such that*

$$\forall x, y \in \{n/2, \ldots, n\}, \quad |g(x+y) - g(x) - g(y)| \leq \theta' n,$$

$$\text{and} \quad \forall x \in \mathcal{D}_n^+, \quad g(x) = g(2^{k_x} x)/2^{k_x},$$

*then the linear function $l : \mathcal{D}_n^+ \to \mathbb{R}$, which is defined by $l(n) \stackrel{\text{def}}{=} g(n)$, satisfies, for all $x \in \mathcal{D}_n^+$,*

$$|g(x) - l(x)| \leq 5\theta' x.$$

*Proof.* Here we borrow a technique developed in [KMS99] that we apply to the function $g$ where it is approximately linear. First we extend $g$ restricted to $\{n/2, \ldots, n\}$ to a function $h$ defined over the whole semi-group $\{x \in \mathbb{N} : x \geq n/2\}$. The extension $h$ is defined for all $x \geq n/2$ by

$$h(x) \stackrel{\text{def}}{=} \begin{cases} g(x) & \text{if } n/2 \leq x \leq n, \\ h(x - n/2) + g(n)/2 & \text{otherwise.} \end{cases}$$

One can verify that $h$ satisfies the following doubling property, for all $x \geq \frac{n}{2}$,

$$|h(2x) - 2h(x)| \leq 5\theta' n/2.$$

Then we apply a result due to [KMS99] which is based on some techniques developed in [Hye41].

**Lemma 1 ([KMS99, Lemma 3])** *Let $E_1$ be a semi-group and $E_2$ a Banach space. Let $\varepsilon \geq 0$ and $h : E_1 \to E_2$ be a mapping such that for all $x \in E_1$*

$$\|h(2x) - 2h(x)\| \leq \varepsilon.$$

*If $f : E_1 \to E_2$ is such that $f(x) = \lim_{m \to \infty} h(2^m x)/2^m$ is a well defined mapping, then for all $x \in E_1$*

$$\|h(x) - f(x)\| \leq \varepsilon.$$

Let $f$ be this function. Then by definition of $h$ we get that $f(x) = xg(n)/n$, for all $x \geq n/2$, therefore $f$ is linear and $f = l$. We conclude the proof by recalling that $g$ equals $h$ on $\{n/2, \ldots, n\}$, and when $1 \leq x < n/2$, $g(x) = g(2^{k_x} x)/2^{k_x}$. □

## 3 Testing with relative error

In this section we show how our results lead to approximate self-testers with linear error terms and with relative error. First let us define the relative distance. Let $\mathcal{F}$ be a collection of real functions defined over a finite domain $D$. For a real

$\theta \geq 0$ and functions $P, f : D \to \mathbb{R}$, we will define the *$\theta$-relative distance* between $P$ and $f$ on $D$ by

$$\theta\text{-}\mathbf{rdist}_D(P, f) \stackrel{\text{def}}{=} \mathop{\mathbf{Pr}}_{x \in D} \left[ |P(x) - f(x)| > \theta |f(x)| \right],$$

and the $\theta$-relative distance between $P$ and $\mathcal{F}$ on $D$ by

$$\theta\text{-}\mathbf{rdist}_D(P, \mathcal{F}) \stackrel{\text{def}}{=} \mathop{\mathbf{Inf}}_{f \in \mathcal{F}} \theta\text{-}\mathbf{rdist}_D(P, f).$$

Note that the relative distance between $P$ and $f$ is not symmetric in general. For example if $\theta > 1$, $P(x) = 0$, and $f(x) = \theta$, for all $x \in \mathcal{D}_n^+$, then $\theta\text{-}\mathbf{rdist}_{\mathcal{D}_n^+}(P, f) = 0$ and $\theta\text{-}\mathbf{rdist}_{\mathcal{D}_n^+}(f, P) = 1$. We will also need another distance which is symmetric but not relative. It is defined for any non negative error term $\beta : D \to \mathbb{R}_+$. The *$\beta$-distance* between $P$ and $f$ on $D$ is

$$\beta\text{-}\mathbf{dist}_D(P, f) \stackrel{\text{def}}{=} \mathop{\mathbf{Pr}}_{x \in D} \left[ |P(x) - f(x)| > \beta(x) \right],$$

and the $\beta$-distance between $P$ and $\mathcal{F}$ on $D$ is

$$\beta\text{-}\mathbf{dist}_D(P, \mathcal{F}) \stackrel{\text{def}}{=} \mathop{\mathbf{Inf}}_{f \in \mathcal{F}} \beta\text{-}\mathbf{dist}_D(P, f).$$

First we define the approximate self-tester for $\beta\text{-}\mathbf{dist}_D$ using the definition of [KMS99] which generalizes that of [EKR96, GLR$^+$91].

**Definition 1** *Let $\delta_1, \delta_2 \in [0, 1]$, $D_2 \subseteq D_1$, and $\mathcal{F}$ be a collection of real-valued functions defined over $D_1$. Let $\beta_1$ and $\beta_2$ be non negative real-valued functions also defined over $D_1$. A* $(D_1, \beta_1, \delta_1; D_2, \beta_2, \delta_2)$-self-tester *for $\mathcal{F}$ is a probabilistic oracle program $T$ such that for any program $P : D_1 \to \mathbb{R}$:*

- *If $\beta_1\text{-}\mathbf{dist}_{D_1}(P, \mathcal{F}) \leq \delta_1$ then $T^P$ outputs* `PASS` *with probability at least* $2/3$.[1]
- *If $\beta_2\text{-}\mathbf{dist}_{D_2}(P, \mathcal{F}) > \delta_2$ then $T^P$ outputs* `FAIL` *with probability at least* $2/3$.[2]

Now we extend this definition for relative distance.

**Definition 2** *Let $\delta_1, \delta_2 \in [0, 1]$, $D_2 \subseteq D_1$, and $\mathcal{F}$ be a collection of real-valued functions defined over $D_1$. Let $\theta_1$ and $\theta_2$ be non negative reals. A* $(D_1, \theta_1, \delta_1; D_2, \theta_2, \delta_2)$-self-tester with relative error *for $\mathcal{F}$ is a probabilistic oracle program $T$ such that for any program $P : D_1 \to \mathbb{R}$:*

- *If $\theta_1\text{-}\mathbf{rdist}_{D_1}(P, \mathcal{F}) \leq \delta_1$ then $T^P$ outputs* `PASS` *with probability at least* $2/3$.
- *If $\theta_2\text{-}\mathbf{rdist}_{D_2}(P, \mathcal{F}) > \delta_2$ then $T^P$ outputs* `FAIL` *with probability at least* $2/3$.

[1] One can also want this probability to be greater than any confidence parameter $\gamma \in (0, 1)$. Here we simplify our discussion by fixing this parameter to $2/3$.

[2] Same remark.

Usually one would like a self-tester to be different and simpler than any correct program. For example we can ask the self-tester to satisfy the *little-oh property* [BK95], *i.e.* its running time have to be asymptotically less than that of any known correct program. This property could be too restrictive for family testing. Here we simplify this condition. If $T$ is a self-tester for $d$-linearity over $\mathcal{D}_n^+$ then $T$ is required to use only comparisons, additions, and multiplications by powers of 2 (*i.e.* left or right shifts in binary representation). Moreover the number of queries and operations of $T$ has to be independent of $n$.

A direct consequence of Theorem 2 is the existence of a self-tester for the set of linear functions, denoted by $\mathcal{L}$, where the distance is defined for a linear error term.

**Theorem 5** *Let $0 < \delta < 1/144$ be a real, $\theta \geq 0$ a power of 2, and $\beta(x) = \theta x$, for all $x$. Then there exists a $(\mathcal{D}_{8n}^+, \beta/16, \delta/12; \mathcal{D}_n^+, 32\beta, 24\delta)$-self-tester for the set $\mathcal{L}$. Moreover it makes $O(1/\delta)$ queries to the program, and uses $O(1/\delta)$ comparisons, additions, and multiplications by powers of 2.*

*Proof.* Let $N \geq 1$ be an integer whose value will be fixed later. The self-tester $T$ performs $N$ independent rounds. Each round consists in performing the following experiment, where $\Theta \stackrel{\text{def}}{=} \theta n$ :

**Experiment linearity-test$(P, \Theta)$**
1. Randomly choose $x, y \in \mathcal{D}_{4n}^+$.
2. Check if $|P(2^{k_x}x + y) - 2^{k_x}P(x) - P(y)| \leq \Theta$.

A round *fails* if the inequality is not satisfied. Then $T$ outputs `FAIL` if more than a $\delta$ fraction of the rounds fail, and `PASS` otherwise.

Let us define the failure probability of $P$ in each round by

$$\mathbf{err}(P) \stackrel{\text{def}}{=} \mathop{\mathbf{Pr}}_{x,y \in \mathcal{D}_{4n}^+} \left[ |P(2^{k_x}x + y) - 2^{k_x}P(x) - P(y)| > \theta n \right].$$

First suppose $(\beta/16)$-$\mathbf{dist}_{\mathcal{D}_{8n}^+}(P, \mathcal{L}) \leq \delta/12$. The halving principle and simple manipulations lead to $\mathbf{err}(P) \leq \delta/2$. Then a standard Chernoff bound argument yields that if $N = \Omega(1/\delta)$ then $T^P$ outputs `PASS`, with probability at least $2/3$.

Now if $(32\beta)$-$\mathbf{dist}_{\mathcal{D}_n^+}(P, \mathcal{L}) > 24\delta$, then, since $3\delta/2 < 1/96$, the contraposition of Theorem 2 implies $\mathbf{err}(P) > 3\delta/2$. Again, by a Chernoff bound argument if $N = \Omega(1/\delta)$ then $T^P$ outputs `FAIL`, with probability at least $2/3$. □

The previous self-tester has two main disadvantages. First, the error term is linear but not relative. Second, it needs to test the program on a bigger domain. The following theorem gets around these two problems.

**Theorem 6** *Let $0 < \delta < 1/144$ be a real and $0 \leq \theta \leq 16$ a power of 2. Then there exists a $(\mathcal{D}_n^+, \theta/64, \delta/12; \mathcal{D}_n^+, 32\theta, 24\delta)$-self-tester with relative error for the set $\mathcal{L}$. Moreover it makes $O(1/\delta)$ queries to the program, and uses $O(1/\delta)$ comparisons, additions, and multiplications by powers of 2.*

*Proof.* Now the self-tester $T$ performs $N = O(1/\delta)$ times the following experiment:

**Experiment linearity-relative-test**$(P, \theta)$
1. Randomly choose $y \in \mathcal{D}_n^+$.
2. Compute $G = P(n-y) + P(y) \quad (\text{fix } P(0) = 0)$.
3. Compute $\Theta = \theta|G|$.
4. Do **Experiment linearity-test(extension**$(P, G)$**,**$\Theta$**)**.

The function **extension** is easily computable using $P$, and it is defined by:

**Function extension**$(P, G)(x)$
1. $val = 0$.
2. While $x > n$ do $x = x - n$ and $val = val + G$.
3. Return $(val + P(x))$.

Again, $T$ outputs `FAIL` if more than a $\delta$ fraction of the rounds fail, and `PASS` otherwise.

Fix $\Theta \stackrel{\text{def}}{=} \theta|G|$ and $\beta(x) \stackrel{\text{def}}{=} \theta|G|x/n$, for all $x$. Let $\widetilde{P} \stackrel{\text{def}}{=}$ **extension**$(P, G)$, and denote the failure probability of one experiment by **rerr**$(P)$.

First, suppose there exists a linear function $l$ such that $(\theta/64)$-**rdist**$_{\mathcal{D}_n^+}(P, l) \leq \delta/12$. Therefore $\mathbf{Pr}_{y \in \mathcal{D}_n^+}[|P(n-y) + P(y) - l(n)| > \theta|l(n)|/32] \leq \delta/6$. So $|G - l(n)| \leq \theta|l(n)|/32$ with probability at least $1 - \delta/6$. Suppose this last inequality is satisfied. Then one can verify that $(\theta/32)$-**rdist**$_{\mathcal{D}_{8n}^+}(\widetilde{P}, l) \leq \delta/12$. Since $\theta/32 \leq 1/2$, we also obtain $|l(n)| \leq 2|G|$. Thus the combination of the two last inequalities gives $(\beta/16)$-**dist**$_{\mathcal{D}_{8n}^+}(\widetilde{P}, l) \leq \delta/12$. In this case we previously proved that the failure probability of **linearity-test(extension**$(P, G)$**,**$\Theta$**)** is at most $\delta/2$. In conclusion, if $(\theta/64)$-**rdist**$_{\mathcal{D}_n^+}(P, l) \leq \delta/12$ then **rerr**$(P) \leq \delta/6 + \delta/2 = 2\delta/3$.

Suppose now that $(32\theta)$-**rdist**$_{\mathcal{D}_n^+}(P, \mathcal{L}) > 24\delta$. Then, for all real $G$, $(32\beta)$-**dist**$_{\mathcal{D}_n^+}(\widetilde{P}, l) > 24\delta$, where $l \in \mathcal{L}$ is defined by $l(n) \stackrel{\text{def}}{=} G$. But $\widetilde{P}(n+y) = G + \widetilde{P}(y)$, so $\mathbf{Med}_{y \in \mathcal{D}_{2n}^+}(\widetilde{P}(n+y) - \widetilde{P}(y)) = G$. Therefore the contraposition of Theorem 2 implies that, for all real $G$, **rerr**$(P) > 3\delta/2$.

To conclude the proof, apply a Chernoff bound argument. □

Now we can state our final result which extends the previous one to multi-linear functions. It is quite surprising since it does not use multiplications but only comparisons, additions, and multiplications by powers of 2.

**Theorem 7** *Let $d \geq 1$ be an integer. Let $0 < \delta \leq 1$ be a real and $0 \leq \theta \leq O(1/d^2)$ a power of 2. Then there exists a $((\mathcal{D}_n^+)^d, \theta, \delta; (\mathcal{D}_n^+)^d, O(d)\theta, O(d)\delta)$-self-tester with relative error for the set of real-valued $d$-linear functions defined on $(\mathcal{D}_n^+)^d$. Moreover it makes $O(1/\delta)$ queries to the program, and uses $O(1/\delta)$ comparisons, additions, and multiplications by powers of 2.*

*Proof (sketch).*We use some techniques from [FHS94] where a similar result for multi-variate polynomials in the context of exact computation was proven. Fact 2 and Lemma 2 lower and upper bound the distance between a $d$-variate function and $d$-linear functions by its successive distances from functions which are linear in one of their variables. Then, we estimate the latter quantity by repeating **Experiment linearity-relative-test**. More precisely, the self-tester $T$ will repeat

$O(1/\delta)$ times the following experiment. Then $T$ outputs `FAIL` if more than a $\delta$ fraction of the rounds fail, and `PASS` otherwise.

**Experiment** $d$**-linearity-relative-test**$(P, \theta)$
1. Randomly choose $\vec{z} \in (\mathcal{D}_n^+)^d$.
2. Randomly choose $i \in \{1, \ldots, d\}$.
3. Do **Experiment linearity-relative-test**$(\tilde{P}^i_{\vec{z}}, \theta)$.

The notation $\tilde{P}^i_{\vec{z}}$ denotes the function which takes at $t$ the value $P(z_1, \ldots, z_{i-1}, t, z_{i+1}, \ldots, z_d)$.

Using Fact 2 and Lemma 2, one can conclude the proof using previous methods. □

The bounds involved in the previous proof are explicitly stated in the following, where $\mathcal{L}^d$ denote the set of $d$-linear functions defined over $(\mathcal{D}_n^+)^d$, and $\mathcal{L}_i^d$ the set of functions defined over $(\mathcal{D}_n^+)^d$ which are linear in their $i$-th variable. First let us state the easy one.

**Fact 2** *Let $\theta \geq 0$ be a real. Then for all $f : (\mathcal{D}_n^+)^d \to \mathbb{R}$*

$$\frac{1}{d} \sum_{i=1}^{d} \theta\text{-}\mathbf{rdist}_{(\mathcal{D}_n^+)^d}(f, \mathcal{L}_i^d) \leq \theta\text{-}\mathbf{rdist}_{(\mathcal{D}_n^+)^d}(f, \mathcal{L}^d).$$

The other bound is more difficult and it can be proven by induction on $d$. Due to lack of place we omit the proof.

**Lemma 2** *Let $0 \leq \theta \leq 1/(16d^2)$ be a real. Then for all $f : (\mathcal{D}_n^+)^d \to \mathbb{R}$*

$$(4d\theta)\text{-}\mathbf{rdist}_{(\mathcal{D}_n^+)^d}(f, \mathcal{L}^d) \leq 2 \sum_{i=1}^{d} \theta\text{-}\mathbf{rdist}_{(\mathcal{D}_n^+)^d}(f, \mathcal{L}_i^d).$$

## Open questions

In this paper we achieve the goal of approximate self-testing with relative error for multi-linear functions. We would like to extend this work for polynomials. More generally when we have no information *a priori* on the size of the function to be computed, constructing approximate self-testers with relative error is an interesting challenge.

## Acknowledgments

We would like to thank Stéphane Boucheron, Marcos Kiwi, Sophie Laplante and Miklos Santha for useful discussions and assistance while writing this paper.

## References


[ABCG93] S. Ar, M. Blum, B. Codenotti, and P. Gemmell. Checking approximate computations over the reals. In *Proc. 25th STOC*, pages 786–795, 1993.

[ALM+92] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and intractibility of approximation problems. In *Proc. 33rd FOCS*, pages 14–23, 1992.

[AS92] S. Arora and S. Safra. Probabilistic checkable proofs: A new characterization of NP. In *Proc. 33rd FOCS*, pages 1–13, 1992.

[Bab93] L. Babai. Transparent (holographic) proofs. In *Proc. 10th STACS*, volume 665, pages 525–534. LNCS, 1993.

[BK95] M. Blum and S. Kannan. Designing programs that check their work. *J. ACM*, 42(1):269–291, 1995.

[BLR93] M. Blum, M. Luby, and R. Rubinfeld. Self-testing/correcting with applications to numerical problems. *J. Comp. and Syst. Sci.*, pages 549–595, 1993.

[Blu88] M. Blum. Designing programs to check their work. Technical Report 88-009, ICSI, 1988.

[Cop89] D. Coppersmith, December 1989. See discussion in [BLR93].

[EKR96] F. Ergün, S. Ravi Kumar, and R. Rubinfeld. Approximate checking of polynomials and functional equations. In *Proc. 37th FOCS*, pages 592–601, 1996.

[FHS94] K. Friedl, Z. Hátsági, and A. Shen. Low-degree tests. *Proc. 5th SODA*, pages 57–64, 1994.

[For95] G. L. Forti. Hyers-Ulam stability of functional equations in several variables. *Aeq. Mathematicae*, 50:143–190, 1995.

[GGR96] O. Goldreich, S. Goldwasser, and D. Ron. Property testing and its connection to learning and approximation. In *Proc. 37th FOCS*, pages 339–348, 1996.

[GLR+91] P. Gemmell, R. Lipton, R. Rubinfeld, M. Sudan, and A. Wigderson. Self-testing/correcting for polynomials and for approximate functions. In *Proc. 23rd STOC*, pages 32–42, 1991.

[GR97] O. Goldreich and D. Ron. Property testing in bounded degree graphs. In *Proc. 37th STOC*, pages 406–415, 1997.

[HR92] D. H. Hyers and T. M. Rassias. Approximate homomorphisms. *Aeq. Mathematicae*, 44:125–153, 1992.

[HŠ92] D. H. Hyers and P. Šemrl. On the behaviour of mappings which do not satisfy Hyers-Ulam stability. *Proc. AMS*, 144(4):989–993, April 1992.

[Hye41] D. H. Hyers. On the stability of the linear functional equation. *Proc. Nat. Acad. Sci., U.S.A.*, 27:222–224, 1941.

[KMS99] M. Kiwi, F. Magniez, and M. Santha. Approximate testing with relative error. In *Proc. 31st STOC*, pages 51–60, 1999.

[Lip91] R. Lipton. New directions in testing. *Series in Discrete Mathematics and Theoretical Computer Science*, 2:191–202, 1991.

[RS96] R. Rubinfeld and M. Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM J. Comp.*, 25(2):23–32, April 1996.

[Rub99] R. Rubinfeld. On the robustness of functional equations. *SIAM J. Comp.*, 28(6):1972–1997, 1999.

[Sko83] F. Skof. Sull'approssimazione delle applicazioni localmente $\delta$-additive. *Atti Acc. Sci. Torino*, 117:377–389, 1983. in Italian.