Interval-Based Robust Statistical Techniques for Non-Negative Convex Functions, with Application to Timing Analysis of Computer Chips

Michael Orshansky
Wei-Shen Wang
Department of Electrical and Computer Engineering
University of Texas at Austin
Austin, TX 78712,USA
orshansky@mail.utexas.edu

Martine Ceberio
Gang Xiang
Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA
{mceberio,vladik}@cs.utep.edu

ABSTRACT

In chip design, one of the main objectives is to decrease its clock cycle. On the design stage, this time is usually estimated by using worst-case (interval) techniques, in which we only use the bounds on the parameters that lead to delays. This analysis does not take into account that the probability of the worst-case values is usually very small; thus, the resulting estimates are over-conservative, leading to unnecessary over-design and under-performance of circuits. If we knew the *exact* probability distributions of the corresponding parameters, then we could use Monte-Carlo simulations (or the corresponding analytical techniques) to get the desired estimates. In practice, however, we only have *partial* information about the corresponding distributions, and we want to produce estimates that are valid for all distributions which are consistent with this information.

In this paper, we develop general techniques that allow us, in particular, to provide such estimates for the clock time.

1. CASE STUDY

Decreasing clock cycle: a practical problem. In chip design, one of the main objectives is to decrease the chip's clock cycle. It is therefore important to estimate the clock cycle on the design stage.

The clock cycle of a chip is constrained by the maximum path delay over all the circuit paths $D \stackrel{\text{def}}{=} \max(D_1, \ldots, D_N)$, where D_i denotes the delay along the i-th path. Each path delay D_i is the sum of the delays corresponding to the gates and wires along this path. Each of these delays, in turn, depends on several factors such as the variation caused by the current design practices, environmental design characteristics (e.g., variations in temperature and in in supply voltage), etc.

SAC'06 April 23-27, 2006, Dijon, France

Traditional (interval) approach to estimating the clock cycle. Traditionally, the delay D is estimated by using the worst-case analysis, in which we assume that each of the corresponding factors takes the worst possible value (i.e., the value leading to the largest possible delays). As a result, we get the time delay that corresponds to the case when all the factors are at their worst.

It is necessary to take probabilities into account. The worst-case analysis does not take into account that different factors come from independent random processes. As a result, the probability that all these factors are at their worst is extremely small. For example, there may be slight variations of delay time from gate to gate, and this can indeed lead to gate delays. The worst-case analysis considers the case when all these random variations lead to the worst case; since these variations are independent, this combination of worst cases is highly unprobable.

As a result, the current estimates of the chip clock time are over-conservative, over up to 30% above the observed clock time. Because of this over-estimation, the clock time is set too high – i.e., the chips are usually over-designed and under-performing; see, e.g., [6, 7, 8, 22, 21, 23, 24]. To improve the performance, it is therefore desirable to take into account the probabilistic character of the factor variations.

Robust statistical methods are needed. If we knew the *exact* probability distributions of the corresponding parameters, then we could use Monte-Carlo simulations (or the corresponding analytical techniques) to get the desired estimates. In practice, however, we only have *partial* information about the corresponding distributions. For a few parameters, we know the exact distribution, but for most parameters, we only know the mean and some characteristic of the deviation from the mean – e.g., the interval that is guaranteed to contain possible values of this parameter.

In principle, we could pick up some distributions which are consistent with this partial information – e.g., truncated normal distributions. However, the resulting estimates depend on which distributions we pick; so, if we simply pick *some* distributions and it turns our that the actual distributions are different, we may be underestimating the clock time. It is therefore desirable to provide bounds that work for all the distributions which are consistent with the given information.

In statistics, estimates which are guaranteed for all distributions from some non-parametric class are called *robust* (see, e.g., [13]). In these terms, our objective is to provide robust statistical estimates for the clock time.

What we do in this paper. In this paper, we develop general techniques that allow us, in particular, to provide robust estimates for the clock time.

In deriving these estimates, we will use the extensions of interval methods to cases with partial information about probabilities described, e.g., in [11, 17, 18, 19]; see also [1, 2, 3, 4, 5, 20].

2. TOWARDS A MATHEMATICAL FORMULATION OF THE PROBLEM

Case study: how the desired delay D depends on the parameters. The variations in the each gate delay d are caused by the difference between the actual and the nominal values of the corresponding parameters. It is therefore desirable to describe the resulting delay d as a function of these differences $x_1, \ldots, x_n$. Since these differences are usually small, we can safely ignore quadratic (and higher order) terms in the Taylor expansion of the dependence of d on x_i and assume that the dependence of each delay d on these differences can be described by a linear function.

As a result, each path delay D_i – which, as we have mentioned, is the sum of delays at different gates and wires – can also be described as a linear function of these differences, i.e., as $D_i = a_i + \sum_{j=1}^{n} a_{ij} \cdot x_j$.

Thus, the desired maximum delay $D = \max_i D_i$ has the form

$$D = \max_i \left(a_i + \sum_{j=1}^{n} a_{ij} \cdot x_j \right). \qquad (1)$$

How we can describe such functions in general terms. In this paper, we will use two properties of the time delay. First, we will use the fact that the time delay is always non-negative; second, we will use the fact that the dependence (1) is convex.

Let us recall that a function $f : R^m \to R$ is called *convex* if

$$f(\alpha \cdot x + (1-\alpha) \cdot y) \le \alpha \cdot f(x) + (1-\alpha) \cdot f(y)$$

for every $x, y \in R^m$ and for every $\alpha \in (0,1)$. It is known that the maximum of several linear functions is convex, so the function (1) is convex. Vice versa, every convex function can be approximated, with an arbitrary accuracy, by maxima of linear functions – i.e., by expressions of type (1).

So, in general terms, we can say that we are interested in the robust statistical properties of the value $y = F(x_1, \ldots, x_n)$, where F is a non-negative convex function of the variables x_j.

In which characteristics of $y = F(x_1, \ldots, x_n)$ we are interested. We would like to get as much information as possible about the probability distribution of y. In engineering, statistical analysis usually starts with estimating the first and the second moments of the distribution. Let us therefore find estimates for the first moment $M_1 \stackrel{\text{def}}{=} E[y]$ and for the second moment $M_2 \stackrel{\text{def}}{=} E[y^2]$.

It is often also useful to find the values of the higher moments $M_v \stackrel{\text{def}}{=} E[y^v]$ for $v > 2$.

In many practical situations, e.g., for the clock timing, one of the possible objectives is to find a value y_0 such that $y \le y_0$ with the probability $\ge 1 - \varepsilon$ (where $\varepsilon > 0$ is a given small probability).

Once we know M_1 and M_2, how can we estimate y_0: general case. If we have no additional information about the probability distribution of y, then, to estimate the desired value y_0, we can use Chebyshev inequality (see, e.g., [27]), according to which, for every $k_0 > 0$, we have

$$\text{Prob}(|y - M_1| > k_0 \cdot \sigma) \le 1/k_0^2,$$

where $\sigma \stackrel{\text{def}}{=} \sqrt{V} = \sqrt{M_2 - M_1^2}$ is the standard deviation of y. We would like this probability to be $\le \varepsilon$, so we have to take k_0 for which $1/k_0^2 = \varepsilon$, i.e., $k_0 = \varepsilon^{-1/2}$. As a result, we get $y_0 = M_1 + k_0 \cdot \sqrt{M_2 - M_1^2}$ for $k_0 = \varepsilon^{-1/2}$.

If we want to guarantee that $y \le y_0$ with a high probability, e.g., by choosing $\varepsilon = 10^{-3}$, then we must take $y_0 = E + 30\sigma$.

How good are these estimates for y_0. It is well known that as we increase the number of terms in a linear combination of several small random variables, the resulting distribution of a sum tends to Gaussian – this Central Limit Theorem is one of the main reasons why Gaussian distribution is so frequent in practice; see, e.g., [27]. So, it is reasonable to assume that the distribution of each path delay D_i is close to Gaussian. It therefore makes sense to also assume that the distribution for $y = \max D_i$ is Gaussian as well. Under this assumption, we get much better estimates for y_0; for example:

- with 90% probability, we have $y \le y_0 = M_1 + 2\sigma = M_1 + 2 \cdot \sqrt{M_2 - M_1^2}$; for the Chebyshev inequality, a similar bound with probability $\varepsilon = 0.1$ would require $M_1 + 3\sigma$;
- with 99.9% probability, we have $D \le y_0 = M_1 + 3\sigma = M_1 + 3 \cdot \sqrt{M_2 - M_1^2}$; for the Chebyshev inequality, a similar bound with probability $\varepsilon = 0.001$ would require $M_1 + 30\sigma$.

We know that the distribution of y is sometimes close to normal, and for the normal distribution, the actual bound is much smaller: $y_0 = M_1 + 3\sigma$. Thus, the bound based on the first moments is, most probably, an overestimation – hence, the first two moments may not be sufficient.

How to use higher moments in estimating y_0. Since the first two moments $M_1 = E[y]$ and $M_2 = E[y^2]$ are not sufficient, a natural idea is to use higher moments $M_{2q} \stackrel{\text{def}}{=} E[y^{2q}]$ and $M_{2q+1} \stackrel{\text{def}}{=} E[y^{2q+1}]$.

The idea of using the higher moments to estimate y_0 is similar to Chebyshev's inequality. Indeed, if we know the central moment

$$C_{2q} \stackrel{\text{def}}{=} E[(y - M_1)^{2q}] = \int \rho(y) \cdot (y - M_1)^{2q}\, dy,$$

then for $\sigma_{2q} \stackrel{\text{def}}{=} C_{2q}^{1/(2q)}$, we can conclude that the probability of $y > M_1 + k_0 \cdot \sigma_{2q}$ cannot exceed $1/k_0^{2q}$: otherwise, for all $y > M_1 + k_0 \cdot \sigma_{2q}$, we have $(y - M_1)^{2q} > (k_0 \cdot \sigma_{2q})^{2q}$, so the value C_{2q} of the above integral will be higher than

$\text{Prob}(y - M_1 > k_0 \cdot \sigma_{2q}) \cdot (k_0 \cdot \sigma_{2q})^{2q} \ge (1/k_0^{2q}) \cdot (k_0 \cdot \sigma_{2q})^{2q} =$

$$\sigma_{2q}^{2q} = C_{2q}.$$

Thus, to guarantee that $y \le y_0$ with probability $\ge 1-\varepsilon$, we can take $y_0 = M_1 + k_0 \cdot \sigma_{2q}$ with $k_0 = \varepsilon^{-1/(2q)}$.

The larger q, the smaller k_0. For example, for $\varepsilon = 10^{-3}$:

- for $q = 1$, we needed $k_0 = (10^{-3})^{-1/2} \approx 30$;
- for $q = 2$, we need $k_0 = (10^{-3})^{-1/4} \approx 5.5$;
- for $q = 3$, we need $k_0 = (10^{-3})^{-1/6} \approx 3$.

So, to estimate y_0, we must find the central moment C_{2q}. This can be done in a straightforward way. Let us show, e.g., how this can be done for $q = 2$. Since

$$(y - M_1)^4 = y^4 - 4 \cdot y^3 \cdot M_1 + 6 \cdot y^2 \cdot M_1^2 - 4 \cdot y \cdot M_1^3 + M_1^4,$$

we conclude that

$$C_4 = E[(y - M_1)^4] =$$

$$E[y^4] - 4 \cdot E[y^3] \cdot E + 6 \cdot E[y^2] \cdot M_1^2 - 4 \cdot E[y] \cdot M_1^3 + M_1^4 =$$

$$M_4 - 4 \cdot M_3 \cdot M_1 + 6 \cdot M_2 \cdot M_1^2 - 3 \cdot M_1^4.$$

So, to estimate y_0, we must estimate the values of the moments M_v.

What information we can use. What information can we use for these estimations? We can safely assume that different factors x_j are statistically independent. About some of the variables x_j, we know their exact statistical characteristics; about some other variables x_j, we only know their interval ranges $[\underline{x}_j, \overline{x}_j]$ and their means E_j.

We are interested in the ranges of possible values of M_v. For each j for which we do not know the exact probability distribution, there exist many different probability distributions that are consistent with this information. For different distributions, in general, we get different values of M_v.

Our objective is thus to find the ranges of possible values of M_v.

How to estimate the desired value y_0 based on the bounds for M_v: general case. We have already mentioned that if we knew the exact values of the moments, then we could take $y_0 = M_1 + k_0 \cdot \sigma_{2q}$, where $\sigma_{2q} = C_{2q}^{1/(2q)}$ and $k_0 = \varepsilon^{-1/(2q)}$.

Since we do not know the exact distribution, we can only find the bounds $[\underline{M}_1, \overline{M}_1]$ and $[\underline{C}_{2q}, \overline{C}_{2q}]$ for the corresponding moments. Thus, to guarantee that $y \le y_0$ with the probability $\ge 1-\varepsilon$, we must take, as y_0, the largest possible value of $y_0 = M_1 + k_0 \cdot \sigma \cdot \sigma_{2q}$, i.e., we must take

$$y_0 = \overline{M}_1 + k_0 \cdot \overline{\sigma}_{2q},$$

where $\overline{\sigma}_{2q} \stackrel{\text{def}}{=} (\overline{C}_{2q})^{1/(2q)}$.

So, to estimate y_0, we must find the upper bound $\overline{C}_{2q}$ on the central moment C_{2q}. This can be done in a straightforward way. Let us show, e.g., how this can be done for $q = 2$. We have already mentioned that

$$C_4 = M_4 - 4 \cdot M_3 \cdot M_1 + 6 \cdot M_2 \cdot M_1^2 - 3 \cdot M_1^4.$$

Hence, as an upper bound $\overline{C}_4$ for C_4, we can take

$$\overline{C}_4 = \overline{M}_4 - 4 \cdot \underline{M}_3 \cdot \underline{M}_1 + 6 \cdot \overline{M}_2 \cdot \overline{M}_1^2 - 3 \cdot \underline{M}_1^4.$$

Similar formulas can be produced for an arbitrary q.

Case of second moment: motivations. For the case $q = 1$, we can get better estimates for y_0. Indeed, when we know the exact values of M_1 and $\sigma = \sqrt{M_2 - M_1^2}$, then the corresponding value y_0 is equal to $M_1 + k_0 \cdot \sigma$ for some constant k_0. Thus, to guarantee the desired inequality for all possible values $M_1 \in [\underline{M}_1, \overline{M}_1]$ and $M_2 \in [\underline{M}_2, \overline{M}_2]$, we should take, as y_0, the largest possible value of $M_1 + k_0 \cdot \sqrt{M_2 - M_1^2}$ when M_1 and M_2 are within the corresponding intervals.

The desired expression is increasing w.r.t. M_2, so its maximum is attained when M_2 takes the largest possible value $\overline{M}_2$. With respect to M_1, this expression is not always monotonic, its derivative is equal to 0 when $1 + \dfrac{k_0 \cdot (-2M_1)}{2\sqrt{\overline{M}_2 - M_1^2}} = 0$, i.e., when $M_1 = \dfrac{\sqrt{\overline{M}_2}}{\sqrt{k_0^2 + 1}}$. Once can easily see that this value is the maximum of our expression. Thus, we arrive at the following algorithm.

Algorithm for $q = 1$: description. First, we compute the value $E_0 \stackrel{\text{def}}{=} \dfrac{\sqrt{\overline{M}_2}}{\sqrt{k_0^2 + 1}}$. Then:

- If $\overline{M}_1 \le E_0$, we take $\overline{y}_0 = \overline{M}_1 + k_0 \cdot \sqrt{\overline{M}_2 - (\overline{M}_1)^2}$.
- If $\underline{M}_1 \le E_0 \le \overline{M}_1$, we take $\overline{y}_0 = E_0 + k_0 \cdot \sqrt{\overline{M}_2 - E_0^2}$.
- If $E_0 \le \underline{M}_1$, we take $\overline{y}_0 = \underline{M}_1 + k_0 \cdot \sqrt{\overline{M}_2 - (\underline{M}_1)^2}$.

Let us now describe how to estimate the bounds for the moments.

3. FORMULATION OF THE PROBLEM AND THE MAIN RESULT

GIVEN:

- natural numbers n, $k \le n$, and $v \ge 1$;
- a function $y = F(x_1, \ldots, x_n)$ (algorithmically defined) such that for every combination of values $x_{k+1}, \ldots, x_n$, the dependence of y on $x_1, \ldots, x_k$ is convex;
- $n-k$ probability distributions $x_{k+1}, \ldots, x_n$ – e.g., given in the form of cumulative distribution function (cdf) $F_j(x)$, $k+1 \le j \le n$;
- k intervals $\mathbf{x}_1, \ldots, \mathbf{x}_k$, and
- k values $E_1, \ldots, E_k$.

such that for every $x_1 \in [\underline{x}_1, \overline{x}_1], \ldots, x_k \in [\underline{x}_k, \overline{x}_k]$, we have $F(x_1, \ldots, x_n) \ge 0$ with probability 1.

TAKE: all possible joint probability distributions on R^n for which:

- all n random variables are independent;
- for each j from 1 to k, $x_j \in \mathbf{x}_j$ with probability 1 and the mean value of x_j is equal E_j;
- for $j > k$, the variable x_j has a given distribution $F_j(x)$.

FIND: for the variable $y = F(x_1, \ldots, x_n)$, find the set $\mathbf{M}_v = [\underline{M}_v, \overline{M}_v]$ of all possible values of $M_v \stackrel{\text{def}}{=} E[y^v]$ for all such distributions.

Comment: how this problem is related to interval computations and its known extensions. When the only information we have is intervals of possible values of x_j, then we can use *interval computations* to estimate the range of an expression $y = F(x_1, \ldots, x_n)$.

The main idea behind interval computations is as follows. When a computer computes an expression, it *parses* it, i.e., represents this expression as a sequence of elementary operations $a \otimes b$ such as $+$, $\cdot$, and max. For each elementary operation, we know how to transform the intervals of **a** and **b** of possible values of a and b into the interval **c** of possible values of $c = a \otimes b$; the corresponding interval operations are called *interval arithmetic*. It is therefore reasonable to replace, in the sequence of elementary operations that form the computation of D, each operation with real numbers by the corresponding interval operation. The resulting interval is guaranteed to *enclose* the desired range – and sometimes, it is equal (or close) to this range; see, e.g., [14].

In [11, 17, 18, 19], interval arithmetic has been extended to the case when, in addition to the interval of possible values, we also have an additional information about the probabilities of different values within these intervals. In principle, we can similarly replace, in the computation of D, each operation with real numbers by the corresponding operation from [11, 17, 18, 19], and, e.g., get an enclosure for the desired interval $\mathbf{E}$. The problem with this approach is that, similarly to the case of interval computations, in general, we only get an *enclosure* which may be much wider than the actual interval $\mathbf{E}$.

The objective of this paper is to produce the *exact* intervals $\mathbf{M}_v$ (or at least approximations within a given accuracy). The following result explains how we can compute these intervals.

PROPOSITION 1.

- *The smallest possible value $\underline{M}_v$ is attained when for each j from 1 to k, we use a 1-point distribution in which $x_j = E_j$ with probability 1.*
- *The largest possible values $\overline{M}_v$ is attained when for each j from 1 to k, we use a 2-point distribution for x_j, in which:*
  - $x_j = \underline{x}_j$ *with probability* $\underline{p}_j \stackrel{\text{def}}{=} \dfrac{\overline{x}_j - E_j}{\overline{x}_j - \underline{x}_j}$.
  - $x_j = \overline{x}_j$ *with probability* $\overline{p}_j \stackrel{\text{def}}{=} \dfrac{E_j - \underline{x}_j}{\overline{x}_j - \underline{x}_j}$.

Resulting algorithm for computing exact bounds on M_v. Because of Proposition 1, we can compute the bounds $\underline{M}_v$ and $\overline{M}_v$ by using the following Monte-Carlo simulations:

- To estimate $\underline{M}_v$, we:
  - set the values x_j, $1 \le j \le k$, to be equal to E_j, and
  - simulate the values x_j, $k < j \le n$, as random variables distributed according to the distributions $F_j(x)$.

  For each simulation, we get a value $y = F(x_1, \ldots, x_n)$; the average of the v-th powers y^v of resulting values y is the estimate for $\underline{M}_v$.
- To estimate $\overline{M}_v$, we:
  - set each value x_j, $1 \le j \le k$, to be equal to $\overline{x}_j$ with probability $\overline{p}_j$ and to the value $\underline{x}_j$ with the probability $\underline{x}_j$;
  - simulate the values x_j, $k < j \le n$, as random variables distributed according to the distributions $F_j(x)$;

  for each simulation, we get a value $y = F(x_1, \ldots, x_n)$; the average of the v-th powers y^v of resulting values y is the estimate for $\overline{M}_v$.

Comment about Monte-Carlo techniques. Before presenting the algorithm for computing the upper bound on y_0, let us remark that some readers may feel uncomfortable with the use of Monte-Carlo techniques. This discomfort comes from the fact that in the *traditional* statistical approach, when we know the exact probability distributions of all the variables, Monte-Carlo methods – that simply simulate the corresponding distributions – are inferior to analytical methods. This inferiority is due to two reasons:

- First, by design, Monte-Carlo methods are approximate, while analytical methods are usually exact.
- Second, the accuracy provided by a Monte-Carlo method is, in general, proportional to $\sim 1/\sqrt{N_i}$, where N_i is the total number of simulations. Thus, to achieve reasonable quality, we often need to make a lot of simulations – as a result, the computation time required for a Monte-Carlo method becomes much longer than for an analytical method.

In *robust* statistic, there is often an additional reason to be uncomfortable about using Monte-Carlo methods:

- Practitioners use these methods by selecting a finite set of distributions from the infinite class of all possible distributions, and running simulations for the selected distributions.
- Since we do not test all the distributions, this practical heuristic approach sometimes misses the distributions on which the minimum or maximum of the corresponding distribution is actually attained.

In our case, we also select a finite collection of distributions from the infinite set. However, in contrast to the heuristic (un-justified) selection – which is prone to the above criticism, our selection is *justified*. Proposition 1 *guarantees* that the values corresponding to the selected distributions indeed provide the smallest and the largest values of the characteristics M_v.

In such situations, where a justified selection of Monte-Carlo methods is used to solve a problem of robust statistics, such Monte-Carlo methods often lead to *faster* computations than known analytical techniques. The speed-up caused by using such Monte-Carlo techniques is one of the main reasons why they were invented in the first place – to provide fast estimates of the values of multi-dimensional integrals. Many examples of efficiency of these techniques are given, e.g., in [25]; in particular, examples related to estimating how the uncertainty of inputs leads to uncertainty of the results of data processing are given in [26].

Comment about non-linear terms. In the formula (1), we ignored quadratic and higher order terms in the dependence of each path time D_i on the parameters x_j. It is known that the maximum $D = \max D_i$ of convex functions D_i is always convex. So, according to Proposition 1, the above algorithm will work if we take quadratic terms into consideration – provided that each dependence $D_i(x_1, \ldots, x_k, \ldots)$ is still convex.

4. ACKNOWLEDGMENTS

This work was supported in part by NASA under cooperative agreement NCC5-209, NSF grant EAR-0225670, NIH grant 3T34GM008048-20S1, and Army Research Lab grant DATM-05-02-C-0046

5. REFERENCES

[1] D. Berleant, M.-P. Cheong, C. Chu, Y. Guan, A. Kamal, G. Sheblé, S. Ferson, and J. F. Peters, Dependable handling of uncertainty, *Reliable Computing* 9(6) (2003), pp. 407–418.

[2] D. Berleant, L. Xie, and J. Zhang, Statool: a tool for Distribution Envelope Determination (DEnv), an interval-based algorithm for arithmetic on random variables, *Reliable Computing* 9(2) (2003), pp. 91–108.

[3] D. Berleant and J. Zhang, Using Pearson correlation to improve envelopes around the distributions of functions, *Reliable Computing*, 10(2) (2004), pp. 139–161.

[4] D. Berleant and J. Zhang, Representation and Problem Solving with the Distribution Envelope Determination (DEnv) Method, *Reliability Engineering and System Safety*, 85 (1–3) (July-Sept. 2004).

[5] D. Berleant and J. Zhang, Using Pearson correlation to improve envelopes around the distributions of functions, *Reliable Computing*, 10(2) (2004), pp. 139–161.

[6] D. Boning and S. Nassif, "Models of process variations in device and interconnect", in: A. Chandrakasan (ed.), *Design of High-Performance Microcomputer Ciucruits*, 2000.

[7] D. Chinnery and K. Keutzer, "Closing the gap between ASICs and custom", *Proc. of the Design Automation Conference DAC'2000*.

[8] D. Chinnery and K. Keutzer (eds.), *Closing the Gap Between ASICs and Custom*, Kluwer, 2002.

[9] R. E. Edwards, *Functional analysis: theory and applications*, Dover Publ., N.Y., 1995.

[10] S. Ferson, *RAMAS Risk Calc 4.0: Risk Assessment with Uncertain Numbers*, CRC Press, Boca Raton, Florida, 2002.

[11] S. Ferson, L. Ginzburg, V. Kreinovich, and J. Lopez, "Absolute Bounds on the Mean of Sum, Product, etc.: A Probabilistic Extension of Interval Arithmetic", *Extended Abstracts of the 2002 SIAM Workshop on Validated Computing*, Toronto, Canada, May 23–25, 2002, pp. 70–72.

[12] J. Galambos, *The Asymptotic Theory of Extreme Order Statistics*, Wiley, New York, 1987.

[13] P. J. Huber, *Robust statistics*, Wiley, New York, 2004.

[14] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, *Applied interval analysis: with examples in parameter and state estimation, robust control and robotics*, Springer-Verlag, London, 2001.

[15] S. Kotz and S. Nadarajah, *Extreme Value Distributions: Theory and Applications*, Imperial College Press, London, UK, 2000.

[16] V. Kreinovich and S. Ferson, "Computing best-possible bounds for the distribution of a sum of several variables is NP-hard", *International Journal of Approximate Reasoning*, to appear.

[17] V. Kreinovich, S. Ferson, and L. Ginzburg, "Exact Upper Bound on the Mean of the Product of Many Random Variables With Known Expectations", *Reliable Computing*, 9(6) (2003), pp. 441–463.

[18] V. Kreinovich and L. Longpré, "Computational complexity and feasibility of data processing and interval computations, with extension to cases when we have partial information about probabilities", In: V. Brattka, M. Schroeder, K. Weihrauch, and N. Zhong, *Proceedings of the Conference on Computability and Complexity in Analysis CCA'2003*, Cincinnati, Ohio, USA, August 28–30, 2003, pp. 19–54.

[19] V. Kreinovich, G. N. Solopchenko, S. Ferson, L. Ginzburg, and R. Aló, "Probabilities, intervals, what next? Extension of interval computations to situations with partial information about probabilities", *Proceedings of the 10th IMEKO TC7 International Symposium on Advances of Measurement Science*, St. Petersburg, Russia, June 30–July 2, 2004, Vol. 1, pp. 137–142.

[20] A. T. Langewisch and F. F. Choobineh, "Mean and variance bounds and propagation for ill-specified random variables", *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, 34(4) (2004), pp. 494–506.

[21] M. Orshansky, "Increasing circuit performance through statistical design techniques", in: D. Chinnery and K. Keutzer (eds.), *Closing the Gap Between ASICs and Custom*, Kluwer, 2002.

[22] M. Orshansky and A. Bandyopadhyay, "Fast statistical timing analysis handling arbitrary delay correlations", *Proc. of the Design Automation Conference DAC'2004*, San Diego, California, June 7–11, 2004, pp. 337–342.

[23] M. Orshansky and K. Keutzer, "A general probabilistic framework for worst case timing analysis," *Proceedings of the Design Automation Conference DAC'2002*, June 2002, pp. 556–561.

[24] M. Orshansky and K. Keutzer, "From Blind Certainty to Informed Uncertainty," *Proceedings of the TAU Workshop*, December 2002.

[25] S. Rajasekaran, P. Pardalos, J. Reif, and J. Rolim (eds.), *Handbook on Randomized Computing*, Kluwer, 2001.

[26] R. Trejo and V. Kreinovich, "Error Estimations for Indirect Measurements: Randomized vs. Deterministic Algorithms For 'Black-Box' Programs", In: S. Rajasekaran, P. Pardalos, J. Reif, and J. Rolim (eds.), *Handbook on Randomized Computing*, Kluwer, 2001, pp. 673-729.

[27] H. M. Wadsworth Jr., *Handbook of statistical methods for engineers and scientists*, McGraw-Hill, N.Y., 1990.