Breakthrough in Interval Data Fitting I. The Role of Hausdorff Distance

Marek W. Gutowski

Institute of Physics, Polish Academy of Sciences, 02–668 Warszawa, Poland,
email: gutow@ifpan.edu.pl

Abstract — This is the first of two papers describing the process of fitting experimental data under interval uncertainty. Probably the most often encountered application of **global optimization methods** is finding the so called *best fitted* values of various parameters, as well as their uncertainties, based on experimental data. Here I present the methodology, designed from the very beginning as an interval-oriented tool, meant to replace to the large extent the famous **Least Squares (LSQ)** and other slightly less popular methods. Contrary to its classical counterparts, the presented method does not require any poorly justified prior assumptions, like *smallness* of experimental uncertainties or their normal (Gaussian) distribution. Using interval approach, we are able to fit rigorously and reliably not only the simple functional dependencies, with no extra effort when both variables are uncertain, but also the cases when the constitutive equation exists in **implicit** rather than **explicit functional form**. The magic word and a key to success of **interval approach** appears the **Hausdorff distance**.

1 Introduction

Handling experimental uncertainties lies at the heart of physics and other sciences. The situation is rather clear during simple measurements, for example when determining the mass, m, of an object under investigations. In this case the uncertainty of a single measurement is equal to one half of the smallest division (Δ) of the measuring device, here the balance. Consequently, we can be sure that the unknown mass certainly belongs to the interval $[m - \Delta/2,\, m + \Delta/2]$, where m is a direct readout from the instrument's scale. We can try to increase our confidence into the result by repeating the measurement several times or by using several measuring devices, maybe delivering the results with different precision Δ. By doing so, we necessarily switch to the probabilistic way of presenting measurement's results, namely as a pair of two numbers: the most likely value, m, and its standard deviation, $\sigma(m)$, expressing the uncertainty of a measurement.

Nowadays more and more measurements belong to a class called *indirect measurements*, where we deduce the numerical values of various interesting parameters without comparing them directly to the appropriate standard. When determining the resistance of a sample we usually take a series of measurements, obtaining many pairs (voltage, current). Then, applying the Ohm's law relating those two physical quantities, we find the single proportionality constant, called resistance, R = voltage/current. Using many

measurements, not just one, we switch naturally to the probabilistic form of result's presentation: as the most likely value and its standard deviation (here: R and $\sigma(R)$).

This situation is highly unsatisfactory. First, we can never be 100% sure whether the partial uncertainties are indeed *small*, thus justifying the application of various 'error propagation laws'. Secondly, neither the probability density function of measurements nor of the result need not to be normal (Gaussian), thus invalidating the frequent claim that the probability of the true value belonging to the interval $[m - \sigma,\ m + \sigma]$ is approximately equal to 67%. In reality, the overwhelming majority of contemporary measuring devices deliver discrete, digital results. Such results obviously cannot be normally distributed.

In this paper I present other ways of experimental data processing, using interval methods. Here we operate on guaranteed rather than probabilistic quantities, thus the results obtained on this way should also be guaranteed. In practice, the results produced by straightforward interval implementation of the classical methods often appear disappointing. In this paper I point to the possible reasons and show a remedy.

2 Basics of interval calculus

From now on we will use *intervals* every time our numbers are uncertain. Specifically, by interval $\mathbf{x}$ we will understand a subset of real numbers: $\mathbf{x} = [\underline{x},\ \overline{x}]$, where both $\underline{x} \leqslant \overline{x}$ are real numbers. The set of all intervals is usually written as $\mathbb{IR}$. We can identify real numbers with intervals of type $\mathbf{x} = [x, x]$, i.e. having their *lower* and *upper bound* equal to each other, and called *degenerate intervals*, *thin intervals*, or *singletons*. This way the set of intervals may be regarded to be an extension of real line: $\mathbb{R}^1 \subset \mathbb{IR}$.

Arithmetic operations on intervals are defined in such a way that the resulting interval always contains all the possible outcomes of the operation in question (and only those outcomes) when the real operands are arbitrarily drawn from their interval representations. One has to keep in mind that – in case of more complicated arithmetic expressions – the final result may be *overestimated*.

Since intervals are sets of numbers, the set operations apply to them as well. Consequently, we can obtain the intersection of two intervals or their union. This last operation may produce a disconnected set, not an interval. Therefore the union of two intervals is often replaced by an *interval hull*, that is the smallest interval containing them both: $\mathrm{hull}\,(\mathbf{x}, \mathbf{y}) \equiv \square(\mathbf{x},\ \mathbf{y}) = \big[\min(\underline{x}, \underline{y}),\ \max(\overline{x},\ \overline{y})\big]$.

The nice feature of interval computations is its straightforward extensibility to operate on vectors, matrices, or real-valued functions as well. The vectors with interval components are usually called *boxes* and it is easy to see why. For more information on interval computations the reader is referred to other sources [1]. In what follows we will need two real-valued functions defined for interval arguments: $\mathrm{w}(\mathbf{x}) = \overline{x} - \underline{x} \geqslant 0$ (width of interval $\mathbf{x}$), and $\mathrm{c}(\mathbf{x}) = (\underline{x} + \overline{x})/2$ – the center of interval $\mathbf{x}$.

2.1 Common features of interval algorithms

Interval-oriented algorithms usually operate on lists of boxes and treat them accordingly to their current status. *Good* boxes are retained for further reference, while *bad* boxes are discarded as soon as possible. The third category, called *pending*, *unresolved* or *undetermined*, is the most interesting and as such is the main object of processing. The general idea is to start from a single big box, suspected to contain the solution (or

solutions) and therefore initially labelled as pending. Pending boxes, one by one, are divided into smaller parts (usually two) and disappear from the list. Offspring boxes are subjected to one or more tests aimed to determine their current status. Some of them are recycled back onto pending list, while *bad* ones are immediately discarded, and *good* are collected separately as a part of the final result. The procedure terminates when either the list of pending boxes is empty or contains only 'small' boxes. On exit, the result is a union of good boxes, possibly appended with pending ones. Empty output is a **proof** that the solution(s) of our problem, if any, are located outside the initial box.

Of course, the exact meaning of *bad* and *good* depends on context.

3 The problem of experimental data fitting and its solution

Suppose we have a theory $\mathcal{T}$, characterized by k unknown parameters $\mathbf{p}_1, \ldots, \mathbf{p}_k$, and N ($N \geqslant k$) results of measurements $\mathbf{m}_1, \ldots, \mathbf{m}_N$, each taken at different values of some well controlled variables, for example the varying temperature or magnetic field. Let $\mathbf{x}_j$ denotes the value(s) of controlled variable(s) (environment) during the measurement $\mathbf{m}_j$. In what follows we will use the simplified notation:

$$\begin{array}{lcll} \mathcal{T}(\mathbf{p}_1, \ldots, \mathbf{p}_k; \mathbf{x}_j) & = & \mathbf{t}_j(\mathbf{p}), & j = 1, \ldots, N \\ \mathbf{m}_j(\mathbf{x}_j) & \equiv & \mathbf{m}_j, & \end{array} \tag{1}$$

where the upper line describes the theoretical outcomes of the experiment and the lower one – actual experimental results. Usually each theoretical outcome $\mathbf{t}_j$ (we will often skip the explicit dependence of $\mathbf{t}_j$ on the set of unknown parameters $\mathbf{p}$) is crisp[1] when all the arguments of $\mathcal{T}$ (including $\mathbf{x}_j$) are crisp. Contrary, the measurements $\mathbf{m}_j$ are always uncertain and will be considered from now on to be intervals (or more generally: interval vectors).

The task for an experimentalist is to adjust the values of all unknown parameters $\mathbf{p}_1, \ldots, \mathbf{p}_k$, in such a way that every measurement $\mathbf{m}_j$ differs as little as possible from the corresponding theoretical prediction $\mathbf{t}_j$. In fact, we would be happy to finish the procedure observing

$$\mathbf{t}_j \cap \mathbf{m}_j \neq \varnothing, \quad \forall j \in \{1, 2, \ldots, N\} \tag{2}$$

(in simple words: the theoretical curve passes through all the experimental points) and with unknown parameters, $\mathbf{p}$'s, as narrow as possible.

Note that the requirement $\mathbf{t}_j \subseteq \mathbf{m}_j$ for every j, although tempting, is going too far: our results (if they exist) would be severely underestimated. Yes, we want to be accurate, but we can't afford to tolerate underestimates.

It is out of scope of this article to elaborate the details of many existing particular procedures [2]–[14] aimed to maximally contract the initial box of unknown parameters $\mathbf{p}_1 \times \cdots \times \mathbf{p}_k$ in such a way that all the relations (2) are satisfied. Some of those methods (see [9]) deliver a single box being an interval hull of solutions, while others, more time-consuming, output more boxes, covering with certainty the solution set. Technically: we call a box *bad* when at least one of inequalities (2) is violated at all its internal points and *pending* otherwise. In rare cases it happens to encounter $\mathbf{t}_j \subseteq \mathbf{m}_j$ for all j's – such box is *good* and requires no further processing.

[1] We call the object *crisp* (point, point-wise, point-like) in contrast to the one having interval character, be it a real number, vector or matrix.

3.1 Advantages and disadvantages of the rigorous solution

The practical implementations of the above described approach are still rare [15, 16] and one may wonder why. The unquestionable mathematical rigor of a procedure is certainly among its major advantages. Unfortunately, for very practical reasons, this is also its weak point. And here is why:

- the routine requires guaranteed intervals as measurements, i.e. the ones containing true outcomes of an experiment with probability equal exactly to 1. It is impossible to satisfy this requirement in laboratory practice since measurements usually have the form $m = m_0 \pm \sigma(m)$, where m_0 is a mean value and $\sigma(m)$ – its standard deviation. Nobody knows (or cares) what is the distribution of m, even its support is usually unknown (well, some physical quantities, like mass or absolute temperature, have to be non-negative). Consequently, taking intervals $[m_0 - \sigma,\, m_0 + \sigma]$ as input data, we will most likely finish with empty set of results. This is because at least one-third of N relations (2) has to fail, for arbitrary (even crisp!) set of unknown parameters $\{p_1, \ldots, p_k\}$;
- the quick and dirty fix, coming to the mind in the above situation, is to use wider intervals, like $[m_0 - 3\sigma,\, m_0 + 3\sigma]$, as 'almost guaranteed' input data. Depending on the individual luck, this trick may or may not work. If it doesn't then we are left without even a foggy idea what are the values of our unknown parameters. If it works then uncertainties of so obtained parameters are often very large, very pessimistic – at least when compared with those obtained on other ways. Besides: are we entitled to 'scale back' the obtained uncertainties dividing them all by 3? The only honest answer is *no*;
- sometimes, when the fix works, the uncertainties of unknown parameters appear unrealistically, not to say suspiciously, small. This will surely happen when among our data there is at least one element which should be labeled as a 'near outlier'. It may be due to the undetected data transmission/recording error, power line fluctuations, or whatever.
- another approach is to require fitted curve to pass through at least $N_1 < N$ experimental points, without any prior indication which points are to be preferred. The sensible choice for N_1 is $N_1 \approx 2N/3$. We will not discuss further this idea.

On the other hand, when all our data are credible but the solution set is empty – our theory $\mathcal{T}$ must be flawed. This could be a very strong statement, but is not. There is one more possibility: the experimental uncertainties are seriously underestimated, deliberately or otherwise. So, experimenters, be warned: cheating will be severely punished by interval methods and will bring you nowhere!

Concluding: while the rigorous approach has many advantages, we definitely need something else.

3.2 Back to the basics

The situation is especially upsetting when experimental data look fine, and we are sure the theory is correct, yet the interval routine returns no answer at all. What can we do?

In classical data analysis, the LSQ method (**L**east **Sq**uares) is most commonly used to find unknown parameters. In short, its essence is to minimize the quantity called *chi*

squared:

$$\chi^2(\mathbf{p}) = \sum_{j=1}^{N} \frac{(m_j - t_j(\mathbf{p}))^2}{\sigma_j^2}, \tag{3}$$

where $\mathbf{p}$ is a (crisp) vector of k unknown parameters – arguments of the function χ^2. We are looking for such a vector $\mathbf{p}^\star$ as to have $\chi^2(\mathbf{p}^\star) = \min$. It is well known that LSQ method never fails and always produces some results, even for completely wrong theory.

The LSQ method is due to Carl Friedrich Gauß (1777–1855) and was originally invented by him around 1794. Later on, in 1809, the same author gave it solid statistical interpretation. We will not proceed with statistical properties of χ^2 and LSQ method in general, instead we will rather concentrate on the original idea. And it was like that:

Due to unavoidable experimental uncertainties (then called *errors*), it is unlikely to draw a line representing theoretical outcomes as a function of environment (x) and have the experimental points $m_j(x)$ lying precisely on this curve at the same time. Note that the notion of standard deviation of the measurements was practically unused those days, so the measurements were just numbers. This means there were no denominators in formula (3). Varying the unknown parameters p_j's, we deform the theoretical curve so as to have it running as closely to the experimental points as possible, thus solving the problem.

Remark: Gauß could not speak about the distance (between a theoretical curve and experimental points), at least not in a strict mathematical sense, since this notion was introduced into mathematics many years later, by another German mathematician, Felix Hausdorff (1868–1942), and, independently, in 1905, by Romanian mathematician Dimitrie Pompeiu (1873–1954) [17].

Having known the notion of distance, Gauß would almost certainly use it in his early idea of LSQ method. Indeed, the formula (3) is nothing else as squared Euclidean distance in N-dimensional space, where σ_j serves as a unit length in direction j.

4 Interval version of LSQ method

The problem of global optimization is well studied. Interval researchers have also contributed their share to this field. They have already devised many interval-thinking-inspired, rigorous procedures aiming to solve such tasks. Minimizing χ^2 is obviously one of them. Simple optimization cases, in LSQ sense, were investigated since at least 1990 [18]. Several closely related problems have been successfully solved using rigorous interval algorithms, yet the specific difficulties, recalled in Sec. 3.1, still have not been addressed satisfactorily.

The enormous potential of interval methods for uncertain data processing has been recognized long ago (1993) by Walster and Kreinovich [19]. Kosheleva and Kreinovich (1999) pointed that the cost of interval approach only slightly exceeds the traditional (probabilistic) one [20]. Muñoz and Kearfott (2001) have shown that non-smooth cases should be essentially no more difficult than regular ones [7]. Only one year later (2002), Yang and Kearfott heralded 'new paradigm' and 'new way of thinking about data fitting' [21]. Unfortunately, they did not show how to practically implement their ideas other than for linear equation set. Finally, (2005) Zhilin [13] found the solution for the case when measurements are (multi)linearly related to the environment.

4.1 Correct interval form of χ^2

How to 'translate' the fundamental formula (3) into its interval counterpart? The first idea is to simply replace real-valued measurements m_j by their interval representations $\mathbf{m}_j$, apply the same trick to theoretical predictions $t_j \rightarrow \mathbf{t}_j$, and retain real-valued σ_j's. Consequently, the functional $\boldsymbol{\chi}^2$ becomes interval-valued as well. So far, so good, but wait.

In the light of what was said before, we have to rewrite (3) in terms of distance between experimental and theoretical (predicted, simulated) points. Yet the expression $\mathbf{m}_j - \mathbf{t}_j$, nor even $|\mathbf{m}_j - \mathbf{t}_j|$, is *not* the correct mathematical distance! There are two reasons for that:

- it is not a real number, and
- the implication $(\mathbf{m} = \mathbf{t}) \Rightarrow (\mathbf{m} - \mathbf{t} = 0\,)$ is false.

This naturally raises the question: what *is* the distance between two intervals, here $\mathbf{m}$ and $\mathbf{t}$? Following Hausdorff, we will use his metrics, adapted for $\mathbb{IR}$ space by Moore [22] as:

$$d\,(\mathbf{a},\,\mathbf{b}) = \max\,\left(|\,\underline{a} - \underline{b}\,|\,,\,\left|\,\overline{a} - \overline{b}\,\right|\right). \tag{4}$$

Before continuing, let us only note that another function $d'\,(\mathbf{a},\,\mathbf{b}) = C \cdot\, d\,(\mathbf{a},\,\mathbf{b})$, where $C > 0$ is a fixed number, is a correct distance, too. Additionally, the mathematical distance is not expressed neither in miles nor millimeters, or in any other units – it is simply dimensionless. This is important, since using the strict mathematical distance between intervals we will be able to fit our experimental data obtained from two (or more) completely different experiments, provided the sets of unknown parameters, relevant to each experiment separately, have non-empty intersection. Simply speaking – it is possible to fit several curves simultaneously, in a single run.

Remark: The Euclidean distance between theoretical and experimental points is not the only thinkable distance. We will discuss other N-dimensional metrics in the other article. We will stick, however, to the Moore-Hausdorff distance as one-dimensional metrics in $\mathbb{IR}$, no matter that there are other choices, see for example [23] and [24]. Note also that Moore-Hausdorff distance specialized to the real line simplifies to the familiar form: $d(a, b) = |\,a - b\,|$, as one might expect.

4.2 Analyzing χ^2 term by term

Consider now again a single term in (3). What is the distance between the predicted value $\mathbf{t}$ and the true value $m^\star$ ($m^\star$ is a real number, not interval)? All we know about $m^\star$ is that it satisfies the double inequality: $\underline{m} \leqslant m^\star \leqslant \overline{m}$, with exact value of $m^\star$ remaining unknown. According to the definition (4) we have ($m^\star = [m^\star, m^\star] = \mathbf{m}^\star$):

$$d\,(\mathbf{t},\,m^\star) = \max\left(\left|\underline{t} - m^\star\right|,\,\left|\overline{t} - m^\star\right|\right) \,\in\, \mathbb{R} \tag{5}$$

We don't know which internal point of $\mathbf{m}$ is equal to $m^\star$. Suppose for a moment that m' coincides[2] with one of the endpoints of $\mathbf{m}$, say $m' = \underline{m}$. Call the current distance $d(\mathbf{t}, m' = \underline{m}) = \xi$. What happens when m' gradually moves to $\overline{m}$ – the other endpoint of $\mathbf{m}$?

[2]For a moment we will be dealing with some m', not even necessarily contained in $[\underline{m}, \overline{m}]$, rather than with $m^\star$. This is just for purity: $m^\star$ is a fixed number and we want m' to be variable.

If $\mathbf{t}$ and $\mathbf{m}$ are disjoint then ξ will linearly increase or decrease, depending on the relative position of $\mathbf{t}$ and $\mathbf{m}$ on a real axis. We will finally get either $d\,(\mathbf{t}, \overline{m}) = \xi + \mathrm{w}(\mathbf{m})$ or $d\,(\mathbf{t}, \overline{m}) = \xi - \mathrm{w}(\mathbf{m})$ (this must be a positive number as $\mathbf{t} \cap \mathbf{m} = \varnothing$). Consequently we have bounded the true distance $d\,(\mathbf{t}, m^\star)$ with uncertainty $|d\,(\mathbf{t}, \overline{m}) - d\,(\mathbf{t}, \underline{m})| = \mathrm{w}(\mathbf{m})$. This is indeed very remarkable result, especially when compared with the so called natural interval extension of a basic building block of χ^2, namely the expression '$\mathbf{t} - \mathbf{m}$'. There we always have $\mathrm{w}(\mathbf{t} - \mathbf{m}) = \mathrm{w}(\mathbf{t}) + \mathrm{w}(\mathbf{m}) > \mathrm{w}(\mathbf{m})$. In our approach the uncertainty of a distance in question never exceeds the uncertainty of an individual measurement and – surprisingly – it does *not* depend on current accuracy of unknown parameters.

It remains to show how this result changes when $\mathbf{t}$ and $\mathbf{m}$ overlap. Previously, thanks to the disjointness of $\mathbf{t}$ and $\mathbf{m}$, only one argument of (5) was 'active' at all times, meaning that its value defined the distance. Now, at some point m' (not necessarily $m^\star$), the two arguments can exchange their roles and the other one may become 'active'. If this happens (it may not) then the direction of change of ξ will change too, thus 'un-doing' the already acquired change. Therefore the final change of ξ cannot even reach $\mathrm{w}(\mathbf{m})$.

5 Main result

The analysis just presented is intuitive and easy to follow, but still crude. It says nothing about the values of bounds, being limited only to their separation. This is not enough to be useful in practice. In particular, it is easy to show that m' at which the two arguments of (5) are equal to each other is $m' = \left(\underline{t} + \overline{t}\right)/2$ – the center of interval $\mathbf{t}$ (but only when m' also belongs to $\mathbf{m}$, otherwise there is no such point and switching of roles does not occur). For such m', the distance $d\,(\mathbf{t}, m') = (\overline{t} - \underline{t})/2$. With this fact in mind, we are able to derive the exact bounds for the distance ρ, between a predicted value $\mathbf{t}$ of an experimental outcome, under known environment $\mathbf{x}$, and the true value $m^\star \in \mathbf{m}$, in the same circumstances $\mathbf{x}$. They are following:

- when $\mathrm{c}(\mathbf{t}) \in \mathbf{m}$:

$$\begin{aligned} \text{lower bound:} \quad \underline{\rho} &= \tfrac{1}{2}\mathrm{w}(\mathbf{t}) \\ \text{upper bound:} \quad \overline{\rho} &= \max\left[\, d\,(\mathbf{t}, \underline{m}),\, d\,(\mathbf{t}, \overline{m})\,\right] \end{aligned} \tag{6}$$

- when $\mathrm{c}(\mathbf{t}) \notin \mathbf{m}$:

$$\begin{aligned} \text{lower bound:} \quad \underline{\rho} &= \min\left[\, d\,(\mathbf{t}, \underline{m}),\, d\,(\mathbf{t}, \overline{m})\,\right] \\ \text{upper bound:} \quad \overline{\rho} &= \max\left[\, d\,(\mathbf{t}, \underline{m}),\, d\,(\mathbf{t}, \overline{m})\,\right], \end{aligned} \tag{7}$$

where $d\,(\cdot, \cdot)$ is a Moore-Hausdorff distance between intervals. Please note that generally

$$d\,(\mathbf{t}, \mathbf{m}) \neq \max\left[\, d\,(\mathbf{t}, \underline{m}),\, d\,(\mathbf{t}, \overline{m})\,\right], \tag{8}$$

so the above bounds cannot be written in a more compact form.

In addition, we have introduced a new symbol, ρ, for the mathematically correct distance between the true measured quantity and its interval theoretical estimate. $\boldsymbol{\rho}$ stands for its interval enclosure (note the bold font). The intuitively shown relation $\mathrm{w}(\boldsymbol{\rho}) \leqslant \mathrm{w}(\mathbf{m})$ remains true.

6 Discussion

In conclusion, the recommended form of interval version of χ^2 functional is:

$$\boldsymbol{\chi}^2 = \sum_{j=1}^{N} \left[\frac{\boldsymbol{\rho}\left(\mathbf{t}_j\left(\mathbf{p}_1, \ldots, \mathbf{p}_k\right),\ \mathbf{m}_j\right)}{\mathrm{w}(\mathbf{m}_j)} \right]^2, \tag{9}$$

where $\boldsymbol{\rho} = \left[\underline{\rho}, \overline{\rho}\right]$ is given by (6) and (7), respectively. To recover the unknown parameters $\mathbf{p}_1, \ldots, \mathbf{p}_k$, one has to find a global minimum of (9) with respect to those parameters. This task may be accomplished using procedures developed mostly by Luc Jaulin and Éric Walter (set-inversion methods) as well as ideas first put forward by Shary [25]. Especially the Jaulin's and Walter's algorithm SIVIA looks very promissing: it spares a great deal of computing time by bisecting only some boxes, leaving intact those already classified as *good*. No matter which approach will be adopted, it is clear that no crisp values of unknown parameters can be ever obtained, just because all $\mathbf{m}$'s are intervals. On the other hand, this very feature is highly desirable, since the uncertainties of searched parameters are evaluated very credibly and as nearly a side effect, with no extra effort. This observation sheds new light not only on the problem of experimental data fitting in general, but also substantially changes our perspective on reliable estimates of uncertainties in indirect measurements.

Once the measurements are completed, the values of all $\mathbf{m}$'s, as well as their widths, are fixed. This makes possible to use widths of measurements as the natural unit lengths in every direction of N possible. That is why $\mathbf{m}$'s are present in denominators of all components of the sum (9). When we speak of distances this is the only choice of appropriating individual weights to all measurements. There is no space left for arbitrariness, as it sometimes happens in other versions of the so called *weighted LSQ regression*.

As the measurements are fixed during calculations, the interval enclosure $\boldsymbol{\rho}$ of the distances between the true unknown values $m_j^\star$ and their theoretical predictions $\mathbf{t}_j$, is perfect and cannot be further improved. Moreover, those widths are uniquely determined by the accuracy of the real measurements, not guessed or necessarily subjectively evaluated by a human expert. The optimal widths are very fortunate, since the lower is the width of the interval extension of a function being minimized, the more precisely the global minimum may be located. At least at this respect our procedure definitely beats the natural approach.

Important last minute note: The interval enclosure $\boldsymbol{\rho}$, decribed here, lacks an important property: it is *not* inclusion monotonic. The lack of inclusion monotonicity is rather rare and went undetected during first numerical experiments. Take for example $\mathbf{m} = [10, 35]$, $\mathbf{t} = [0, 60]$, $\mathbf{t}' = [0, 30]$, and $\mathbf{t}'' = [30, 60]$. Obviously $\mathbf{t}' \subset \mathbf{t}$, $\mathbf{t}'' \subset \mathbf{t}$, $\mathbf{t}' \cup \mathbf{t}'' = \mathbf{t}$ and $\boldsymbol{\rho}(\mathbf{t}, \mathbf{m}) = [30, 50]$. On the other hand $\boldsymbol{\rho}(\mathbf{t}', \mathbf{m}) = [15, 35] \not\subset [30, 50]$ and, similarly, $\boldsymbol{\rho}(\mathbf{t}'', \mathbf{m}) = [25, 50] \not\subset [30, 50]$. This renders $\boldsymbol{\rho}$, the interval enclosure of the Hausdorff distance between intervals, completely unsuitable for our purpose. Paradoxically – it is too accurate!

In the second part of this work a connection between interval hulls of solutions and statistical description of their uncertainties will be demonstrated and discussed.

Acknowledgments

This work was done as a part of author's statutory activities at the Institute of Physics, Polish Academy of Sciences.

Bibliography

[1] The website `http://cs.utep.edu/interval-comp/main.html` is an excellent and up-to-date entry point to the wonderful world of interval computations.

[2] Luc Jaulin and Eric Walter, *Guaranteed Nonlinear Parameter Estimation via Interval Computations*, Conf. on Numerical Analysis with Automatic Result Verification, Lafayette, Feb. 25th–March 3rd, 1993, pp. 61–75

[3] Luc Jaulin and Éric Walter, *Guaranteed nonlinear parameter estimation from bounded-error data via interval analysis*, Math. and Comput. in Simulation **35**, 123–137, 1993

[4] Luc Jaulin and Éric Walter, *Set Inversion via Interval Analysis for Nonlinear Bounded-error Estimation*, Automatica **2**9, 1053, 1993

[5] L. Jaulin and É. Walter, *Guaranteed Parameter Bounding for Nonlinear Models with Uncertain Experimental Factors*, Automatica **3**5, 849–856, 1999

[6] Luc Jaulin, *Interval constraint propagation with application to bounded-error estimation*, Automatica **3**6, 1547, 2000

[7] Humberto Muñoz and R.B. Kearfott, *Interval Robustness in Nonsmooth Nonlinear Parameter Estimation*, unpublished preprint, `http://interval.louisiana.edu/preprints/2001_robustness.pdf`

[8] L. Jaulin and É. Walter, *Nonlinear Bounded-Error Parameter Estimation Using Interval Computation*, in: *Granular computing: an emerging paradigm*, Physica-Verlag GmbH Heidelberg, pp. 58–71, 2001

[9] Marek W. Gutowski, *Prosta dostatecznie gruba*, Postępy Fizyki, **53**(4), 181—192, 2002 (in Polish) (*Fat enough straight line*, Advances in Physics, bimonthly of Polish Physical Society)

[10] Voschinin Alexander, Tyurin Alexander, *Interval identification of time series parameters using readings with bounded errors*, 5th International Scientific-Technical Conf. ProcessControl, Pordubice 2002, paper R-210, 7 pages

[11] M.H. van Emden, *Using the duality principle to improve lower bounds for the global minimum in nonconvex optimization*, Second COCOS workshop on intervals and optimization, 2003 (published version: *Using Propagation for Solving Complex Arithmetic Constraints*, `http://arxiv.org/abs/cs/0309018`)

[12] I. Braems, N. Ramdani, A. Boudenne, L. Jaulin, L. Ibos, É. Walter, and Y. Candau, *New set-membership techniques for parameter estimation in presence of model uncertainty*, Proc. of the 5th Int. Conf. on Inverse Problems in Engineering: Theory and Practice, Cambridge, UK, 11–15 July 2005

[13] Sergei I. Zhilin, *On Fitting Empirical Data under Interval Error*, Reliable Computing **11**, 433–442, 2005

[14] Maarten van Emden, *Constraint-Driven Global Optimization*, 13th International Symposium on Scientific Computing Computer Arithmetic and Verified Numerical Computations SCAN'2008, El Paso, Texas, September 29 – October 3, pp. 144–145, 2008

[15] L. Jaulin, J-L. Godet, É. Walter, A. Elliasmine, and Y. Le Duff, *Light scattering data analysis via set inversion*, J. Phys. A: Math. Gen. **30**, 7733–7738, 1993

[16] Maëlenn Aufray, Adrien Brochier, and Wulff Possart, *Set Inversion via Interval Analysis applied to dielectric spectroscopy*, the talk given at SWIM 2008, June 19–20th, Montpelier, France

[17] T. Barsan, D. Tiba, *One hundred years since the introduction of the set distance by Dimitrie Pompeiu*, in: *System modeling and optimization*, Springer, New York, pp. 35–39, 2006

[18] Svetoslav M. Markov, *Least-square approximations under interval input data*, Contributions to Computer Arithmetic and Self-Validating Numerical Methods, C. Ulrich (editor), J.C. Balzer AG, Scientific Publishing Co. (C) IMACS 1990, pp. 133–147

[19] G. WIlliam Walster and Vladik Kreinovich, *For Uknown-but-Bounded Errors, Interval Estimates are Often Better Than Averaging*, ACM SIGNUM Newsletter **31**, 6–19, 1996, `http://www.cs.utep.edu/vladik/1993/tr93-31b.ps.gz`

[20] Olga Kosheleva and Vladik Kreinovich, *Error Estimation for Indirect Measurements: Interval Computation Problem Is (Slightly) Harder Than a Similar Probabilistic Computational Problem*, Reliable Computing **5**, 81–95, 1999

[21] Jie Yang and R. Baker Kearfott *Interval Linear and Nonlinear Regression – New Paradigms, Implementations, and Experiments or New Ways of Thinking About Data Fitting*, talk given at the Seventh SIAM Conference on Optimization, May 20-22, 2002, Toronto, Canada, `http://interval.louisiana.edu/preprints/2002_SIAM_minisymposium.ps`

[22] R.E. Moore *Interval Analysis* Prentice Hall, Englewood Cliffs, NJ, 1966

[23] Francesco Palumbo and Antonio Irpino, *Multidimensional Interval-Data: Metrics and Factorial Analysis*, Applied Stochastic Models and Data Analysis, Brest, France May 17–20, 2005, pp. 689–698, `http://webhouse.unimc.it/economia/repo/39/689.pdf`

[24] Antonio Irpino and Rosanna Verde, *Dynamic clustering of interval data using a Wasserstein-based distance*, Pattern Recognition Letters **29**, 1648–1658, 2008

[25] Sergei P. Shary, *A Surprising Approach in Interval Global Optimization*, Reliable Computing **7**, 497–505, 2001