Checking if There Exists a Monotonic Function That Is Consistent with the Measurements: An Efficient Algorithm

Kavitha Tupelly[1], Vladik Kreinovich[1], and Karen Villaverde[2]

[1]Department of Computer Science, University of Texas at El Paso
El Paso, TX 79968, USA, ktupelly@utep.edu, vladik@cs.utep.edu
[2]Department of Computer Science, New Mexico State University
Las Cruces, NM 88003, USA, kvillave@cs.nmsu.edu

Abstract

In many problems in science and engineering ranging from astrophysics to geosciences to financial analysis, we know that a physical quantity y depends on the physical quantity x, i.e., $y = f(x)$ for some function $f(x)$, and we want to check whether this dependence is monotonic. Specifically, finitely many measurements of x_i and $y_i = f(x_i)$ have been made, and we want to check whether the results of these measurements are consistent with the monotonicity of $f(x)$. An efficient parallelizable algorithm is known for solving this problem when the values x_i are known precisely, while the values y_i are known with interval uncertainty. In this paper, we extend this algorithm to a more general (and more realistic) situation when both x_i and y_i are known with interval uncertainty.

1 Formulation of the Problem

Checking monotonicity is important In many problems in science and engineering, we know that a physical quantity y depends on the physical quantity x, i.e., $y = f(x)$ for some function $f(x)$, and we want to check whether this dependence is monotonic.

In *spectral analysis*, chemical species are identified by locating local maxima of the spectra; see, e.g., [21, 22, 23]. Thus, to identify the chemical species, we must identify intervals between local extrema, i.e., intervals of monotonicity.

In *radioastronomy*, sources of celestial radio emission and their subcomponents are identified by locating local maxima of the measured brightness of the radio sky. In other words, we are interested in the local maxima of the *brightness*

distribution, i.e., of the function $y(x)$ that describes how the intensity y of the signal depends on the position x of the point from which we receive this signal. Thus, in radioastronomy, we must also identify the intervals of monotonicity.

Elementary particles are identified by locating local maxima of the experimental curves that describe (crudely speaking) the scattering intensity y as a function of energy x. Thus, in elementary particle physics, finding intervals of monotonicity is also important.

In *1-D landscape analysis*, e.g., different mountain slopes are different monotonicity intervals; see, e.g., [1, 2, 5, 6].

In *financial analysis*, it is also important to find intervals of monotonicity because they correspond to growth or decline periods; see, e.g., [8, 9]

In *clustering*, different 1-D clusters correspond to a multi-modal distribution, so clusters can be naturally described as combinations of monotonicity intervals separating local minima of the probability density function; see, e.g., [16, 19, 20].

Local maxima and minima are also used in the methods that accelerate the convergence of the measurement result to the real value of a physical variable, and thus allow the user to estimate this value without waiting for the oscillations to stop [18]. Thus, to accelerate convergence, we must also be able to efficiently find intervals of monotonicity.

We can never guarantee monotonicity, so we should check whether it is possible that the actual dependence is monotonic. In many real-life situations, we do not know the exact shape of the function $f(x)$. Instead, we know the results of measuring x and y in several situations. Based on these measurement results, we want to check whether it is possible that the dependence $f(x)$ is monotonic.

Of course, we can only make finitely many measurements. Even in the ideal case when all the measurements are absolutely precise, i.e., when for each measurement i ($1 \le i \le n$), we know the exact value x_i of the quantity x and the exact value y_i of the quantity y, all we know about the function $f(x)$ is that $f(x_i) = y_i$ for all i from 1 to n. As a result, we have no information about the values of $f(x)$ for $x \neq x_i$. In principle, it is quite possible that the actual function is not monotonic: if it goes up and then down between the next two values of x_i, we will never notice it based on the n measurement results.

In short, based on the measurement results, we can never *guarantee* that the actual function $f(x)$ is monotonic. What we can do is check whether it is *possible* that the function $f(x)$ is monotonic, i.e., whether there exists a monotonic function $f(x)$ that is consistent with all the measurement results.

2 What Is Known

Before we describe our result, let us describe what is already known about this problem.

Simplest case: exact measurements. Let us first consider the simplest case when the measurement errors are negligible, so we can safely assume that we know all the values $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ exactly. In this case, we want to check whether there exists a monotonic function $f(x)$ for which $f(x_i) = y_i$ for all i from 1 to n.

In this case, if such a monotonic function exists, then, by definition of monotonicity, for all i and j for which $x_i < x_j$, we have $f(x_i) \le f(x_j)$, therefore, $y_i \le y_j$.

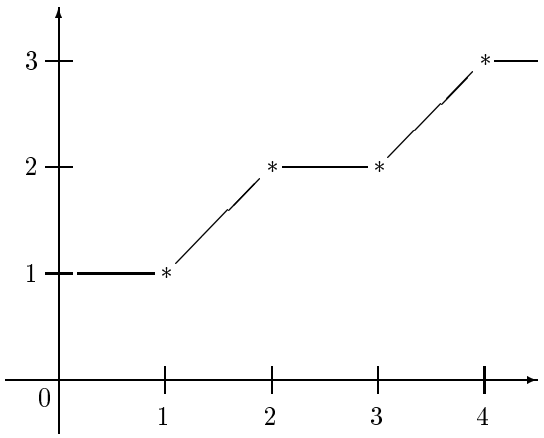
Vice versa, let us assume that for all i and j, $x_i < x_j$ implies $y_i \le y_j$. In this case, we can find a monotonic function $f(x)$ for which $f(x_i) = y_i$. To construct such a function $f(x)$, first, we sort the values x_i into an increasing sequence $x_{(1)} < x_{(2)} < \ldots < x_{(n)}$. Due to our assumed property, the corresponding values of y are also sorted: $y_{(1)} \le y_{(2)} \le \ldots \le y_{(n)}$. We can then design a piecewise linear function that goes through all the points $(x_{(i)}, y_{(i)})$: this function is equal:

- to $y_{(1)}$ for $x \le x_{(1)}$,
- to
$$y_{(i)} + (x - x_{(i)}) \cdot \frac{y_{(i+1)} - y_{(i)}}{x_{(i+1)} - x_{(i)}}$$
for $x_{(i)} \le x \le x_{(i+1)}$, and
- to $y_{(n)}$ for $x \ge x_{(n)}$.

This function is monotonic.

Thus, in this simple case, the existence of a monotonic function that is consistent with all the measurements is equivalent to the following property:

For every i and j, if $x_i < x_j$, then $y_i \le y_j$. (1)



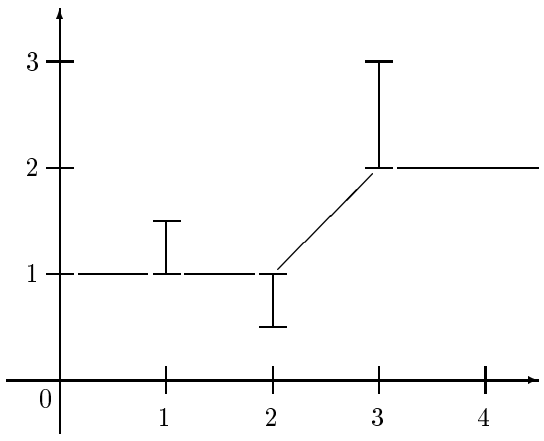
Simplest case: computational complexity. How difficult is it to check this property? If we simply check this property for all $i = 1, \ldots, n$ and all $j = 1, \ldots, n$, then checking this property would require $O(n^2)$ comparisons – i.e., $O(n^2)$ computational steps.

We can check this property faster if, instead, we first sort the values x_i into an increasing sequence $x_{(1)} < x_{(2)} < \ldots < x_{(n)}$ – this requires $O(n \cdot \log(n))$ steps [7] – and then, for every i from 1 to $n-1$, check whether $y_{(i)} \le y_{(i+1)}$ (this requires $O(n)$ steps). Overall, we thus need $O(n \cdot \log(n)) + O(n) = O(n \cdot \log(n))$ steps.

Second case: x is measured exactly, y is known with uncertainty. Let us now consider a more realistic case when we can ignore the measurement error in measuring x but not the measurement error in measuring y. In this case, for each i, the measured value $\widetilde{y}_i$ may be different from the actual value y_i of the measured quantity; in other words, there can be a non-zero measurement error $\Delta y_i \stackrel{\text{def}}{=} \widetilde{y}_i - y_i$.

The manufacturer of a measuring instrument usually provides us with the upper bound Δ_i on the (absolute value of the) measurement error. So, based on the measurement result $\widetilde{y}_i$, the only thing that we can conclude about the actual (unknown) value y_i of the measured quantity is that this value belongs to the interval $[\underline{y}_i, \overline{y}_i]$, where $\underline{y}_i \stackrel{\text{def}}{=} \widetilde{y}_i - \Delta_i$ and $\overline{y}_i \stackrel{\text{def}}{=} \widetilde{y}_i + \Delta_i$.

In this case, we know n values $x_1, x_2, \ldots, x_n$ and we know n intervals $[\underline{y}_1, \overline{y}_1], [\underline{y}_2, \overline{y}_2], \ldots, [\underline{y}_n, \overline{y}_n]$. We want to check whether there exists a monotonic function $f(x)$ that is consistent with all these measurement results, i.e., for which $f(x_1) \in [\underline{y}_1, \overline{y}_1], f(x_2) \in [\underline{y}_2, \overline{y}_2], \ldots$



This problem was solved in [24]. Let us describe the algorithm and its motivations. This solution is based on the following result proven in [24]: The

existence of a monotonic function that is consistent with all the measurements is equivalent to the following property:

$$\text{For every } i \text{ and } j, \text{ if } x_i < x_j, \text{ then } \underline{y}_i \le \overline{y}_j. \tag{2}$$

Let us describe the proof of this result. Let us first assume that there exists a monotonic function $f(x)$ for which $f(x_i) \in [\underline{y}_i, \overline{y}_i]$ for every i. Monotonicity means that for every x and x', if $x < x'$, then $f(x) \le f(x')$. In particular, for every i and j, if $x_i < x_j$, then $f(x_i) \le f(x_j)$. Since $f(x_i) \in [\underline{y}_i, \overline{y}_i]$ and $f(x_j) \in [\underline{y}_j, \overline{y}_j]$, we have $\underline{y}_i \le f(x_i) \le f(x_j) \le \overline{y}_j$; therefore, $\underline{y}_i \le \overline{y}_j$. So, in this case, the above property is true.

Let us now assume that the property is true, and let us prove that in this case, there exists a monotonic function $f(x)$ that is consistent with all the measurement results, i.e., for which $f(x_i) \in [\underline{y}_i, \overline{y}_i]$ for all i.

Similarly to the previous case, we can sort the values x_i in increasing order. So, without losing generality, we can assume that the values x_i are already sorted, i.e., that $x_1 \le x_2 \le \ldots \le x_n$. We can now construct the desired monotonic function $f(x)$ as follows.

For every i, we define $f(x_i) \stackrel{\text{def}}{=} \max(\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_i)$, and then – similar to the above case – we use linear extrapolation and interpolation to define the values of $f(x)$ for all other x.

Due to our selection of $f(x_i)$, we have $f(x_1) \le f(x_2) \le \ldots \le f(x_n)$, so the function $f(x)$ is monotonic on the points x_i. Since we used linear extrapolation, the function $f(x)$ is monotonic for all x.

To complete the proof of the above property, it is sufficient to show that thus constructed function is indeed consistent with all the measurement results, i.e., that $\underline{y}_i \le f(x_i) \le \overline{y}_i$. By definition of $f(x_i)$ as a maximum, we have $\underline{y}_i \le f(x_i)$. Let us now prove the second inequality.

Since the values x_i are sorted, we have $x_k < x_i$ for all $k < i$. Due to the property (2), for every $k < i$, we thus have $\underline{y}_k \le \overline{y}_i$. Clearly, this inequality also holds for $k = i$: indeed, $\underline{y}_i \le \overline{y}_i$ is true, since the lower endpoint of an interval cannot exceed its upper endpoint. So, each of the values $\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_i$ does not exceed $\overline{y}_i$. Therefore, the largest of these values – which is $f(x_i)$ – also cannot exceed $\overline{y}_i$. So, we get the desired inequality $f(x_i) \le \overline{y}_i$. The statement is proven.

Case when x is measured exactly and y is known with uncertainty: computational complexity. How difficult is it to check the above property? If we simply check this property for all $i = 1, \ldots, n$ and all $j = 1, \ldots, n$, then checking this property would require $O(n^2)$ comparisons – i.e., $O(n^2)$ computational steps.

We can check this property faster if we use the fact that for sorted x_i ($x_1 < x_2 < \ldots < x_n$), the above property is equivalent to the following auxiliary

property (motivated by the above construction):

$$\text{For every } i, \text{ we have } \max(\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_i) \le \overline{y}_i. \tag{3}$$

Indeed, we have already proven that if the above property (2) is true, then this auxiliary property (3) is true as well. Vice versa, if the property (3) is true, then for every $k < i$, we have $\underline{y}_k \le \max(\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_i) \le \overline{y}_i$ hence $\underline{y}_k \le \overline{y}_i$.

So, in this case, to check the existence of a monotonic function that is consistent with all the measurement results, we can perform the following three stages:

- First, we sort the values x_i into an increasing sequence

  $$x_{(1)} < x_{(2)} < \ldots < x_{(n)}$$

  – this requires $O(n \cdot \log(n))$ steps.

- Then, for every i from 1 to n, we compute the value $f(x_i) = \max(\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_i)$. If we already know $f(x_i)$, then we can compute the next value $f(x_{i+1})$ by using a single operation $f(x_{i+1}) = \max(f(x_i), \underline{y}_{i+1})$. Thus, computing all n values requires n computational steps.

- Finally, for every i from 1 to n, we compare the value $f(x_i)$ computed on Stage 2 with the value $\overline{y}_i$, to check the inequality (3). Each checking requires one comparison, so to check that this inequality holds for all i from 1 to n, we need n comparisons.

Overall, we thus need $O(n \cdot \log(n)) + O(n) + O(n) = O(n \cdot \log(n))$ steps.

3 General Case: Formulation of the Problem

Let us now consider the general case, when neither the measurement error x nor the measurement error in y can be ignored. In this case, both x and y are known with interval uncertainty, and we have the following problem.

Definition 1. *By a* measurement result*, we mean a pair of intervals* $\langle [\underline{x}, \overline{x}], [\underline{y}, \overline{y}] \rangle$.

Definition 2. *By* measurement data*, we mean a finite sequence of measurement results, i.e., a sequence*

$$(\langle [\underline{x}_1, \overline{x}_1], [\underline{y}_1, \overline{y}_1] \rangle, \langle [\underline{x}_2, \overline{x}_2], [\underline{y}_2, \overline{y}_2] \rangle, \ldots, \langle [\underline{x}_n, \overline{x}_n], [\underline{y}_n, \overline{y}_n] \rangle).$$

Definition 3. *We say that a function $f(x)$ is* consistent with the measurement result $\langle [\underline{x}, \overline{x}], [\underline{y}, \overline{y}] \rangle$ *if there exists a value $x \in [\underline{x}, \overline{x}]$ for which $f(x) \in [\underline{y}, \overline{y}]$.*

Definition 4. *We say that a function $f(x)$ is* consistent with the measurement data *if it is consistent with all the measurement results that form this data.*

The main problem is to, given n measurement results, to check whether there exists a monotonic function $f(x)$ that is consistent with all these measurement results. In other words, we want to check whether it is possible to have a monotonic function $f(x)$ and values values $x_i \in [\underline{x}_i, \overline{x}_i]$ for which $y_i \stackrel{\text{def}}{=} f(x_i) \in [\underline{y}_i, \overline{y}_i]$.

Comment. In our formulation, we know that $f(x_i) \in [\underline{y}_i, \overline{y}_i]$ for *some* $x_i \in [\underline{x}_i, \overline{x}_i]$. In some practical situations, we have a different problem: we know that $f(x_i) \in [\underline{y}_i, \overline{y}_i]$ for *all* $x_i \in [\underline{x}_i, \overline{x}_i]$. For this different problem, an efficient algorithm for checking monotonicity was described in [17].

4 General Case: Main Result

Preliminary analysis of the problem: case when a function cannot be monotonic. In order to come up with a good criterion for checking monotonicity, let us describe a case when it is impossible to have a monotonic function consistent with all the measured data. Namely, if there exists i and j for which $\overline{x}_i \le \underline{x}_j$ and $\underline{y}_i > \overline{y}_j$, then, as we will show, such a monotonic function is impossible.

Indeed, according to our definitions, the existence of a monotonic function $f(x)$ consistent with all the measurements would mean that there exists values $x_i \in [\underline{x}_i, \overline{x}_i]$ and $x_j \in [\underline{x}_j, \overline{x}_j]$ for which $f(x_i) \in [\underline{y}_i, \overline{y}_i]$ and $f(x_j) \in [\underline{y}_j, \overline{y}_j]$. Since $x_i \in [\underline{x}_i, \overline{x}_i]$, we have $x_i \le \overline{x}_i$, and similarly, $\underline{x}_j \le x_j$. Therefore, we have $x_i \le \overline{x}_i \le \underline{x}_j \le x_j$, hence $x_i \le x_j$. Due to monotonicity, we must therefore have $f(x_i) \le f(x_j)$. However, from $f(x_i) \ge \underline{y}_i > \overline{y}_j \ge f(x_j)$, we conclude that $f(x_i) > f(x_j)$ – a contradiction. This contradiction shows that in this case, it is impossible to have a consistent monotonic function.

We will now show that this is the only case when it is impossible to have a monotonic function that is consistent with the measurement results. In other words, we will show, that if for every i and j for which $\overline{x}_i \le \underline{x}_j$, we have $\underline{y}_i \le \overline{y}_j$ (so that $\underline{y}_i > \overline{y}_j$ is impossible), then there exists a monotonic function that is consistent with all the measurement results.

Theorem. *For every measurement data*

$$(\langle [\underline{x}_1, \overline{x}_1], [\underline{y}_1, \overline{y}_1] \rangle, \langle [\underline{x}_2, \overline{x}_2], [\underline{y}_2, \overline{y}_2] \rangle, \ldots, \langle [\underline{x}_n, \overline{x}_n], [\underline{y}_n, \overline{y}_n] \rangle),$$

the existence of the monotonic function that is consistent with the measurement data is equivalent to the following condition:

$$\textit{For every } i \textit{ and } j, \textit{ if } \overline{x}_i \le \underline{x}_j, \textit{ then } \underline{y}_i \le \overline{y}_j. \tag{4}$$

Discussion. The inequality $\overline{x}_i \le \underline{x}_j$ means that no matter what the values x_i and x_j are within the corresponding intervals $[\underline{x}_i, \overline{x}_i]$ and $[\underline{x}_j, \overline{x}_j]$, we always have $x_i \le x_j$. In other words, based on the measurement data, we can guarantee that necessarily $x_i \le x_j$.

There is an area of logic called *modal logic* that deals with statements of the type "necessarily A" or "possibly A". In modal logic, "necessarily A" is usually denoted as $\Box A$. So, the condition "necessarily $x_i \le x_j$" (that is equivalent to $\overline{x}_i \le \underline{x}_j$) can be reformulated as $\Box(x_i \le x_j)$; see, e.g., [4, 10, 11, 14].

Similarly, the condition $\underline{y}_i \le \overline{y}_j$ means that it is *possible* to have $y_i \le y_j$, i.e., that there exist values $y_i \in [\underline{y}_i, \overline{y}_i]$ and $y_j \in [\underline{y}_j, \overline{y}_j]$ for which $y_i \le y_j$. In modal logic notations, "possible" is denoted by $\Diamond$, so "possible $y_i \le y_j$" can be denoted as $\Diamond(y_i \le y_j)$.

Indeed, all we know about the (unknown) actual value y_i of the quantity y measured in i-th measurement is that $y_i \in [\underline{y}_i, \overline{y}_i]$, i.e., $\underline{y}_i \le y_i \le \overline{y}_i$. Similarly, all we know about the (unknown) actual value y_j of the quantity y measured in j-th measurement is that $y_j \in [\underline{y}_j, \overline{y}_j]$, i.e., $\underline{y}_j \le y_j \le \overline{y}_j$.

- If it is possible that $y_i \le y_j$, i.e., if $y_i \le y_j$ for some $y_i \in [\underline{y}_i, \overline{y}_i]$ and $y_j \in [\underline{y}_j, \overline{y}_j]$ for which $y_i \le y_j$, then, due to $\underline{y}_i \le y_i$ and $y_j \le \overline{y}_j$, we have $\underline{y}_i \le y_i \le y_j \le \overline{y}_j$, hence $\underline{y}_i \le \overline{y}_j$.
- Vice versa, if $\underline{y}_i \le \overline{y}_j$, then the inequality $y_i \le y_j$ holds for $y_i = \underline{y}_i \in [\underline{y}_i, \overline{y}_j]$ and $y_j = \overline{y}_j \in [\underline{y}_j, \overline{y}_j]$. Therefore, in this case, it is indeed possible to have $y_i \le y_j$.

So, the inequality $\underline{y}_i \le \overline{y}_j$ is equivalent to $\Diamond(y_i \le y_j)$.

We have shown that the inequality $\overline{x}_i \le \underline{x}_j$ can be rewritten as $\Box(x_i \le x_j)$, and that the inequality $\underline{y}_i \le \overline{y}_j$ can be rewritten as $\Diamond(y_i \le y_j)$. Thus, the condition (4) can be reformulated as follows:

$$\text{For every } i \text{ and } j, \text{ if } \Box(x_i \le x_j), \text{ then } \Diamond(y_i \le y_j). \tag{5}$$

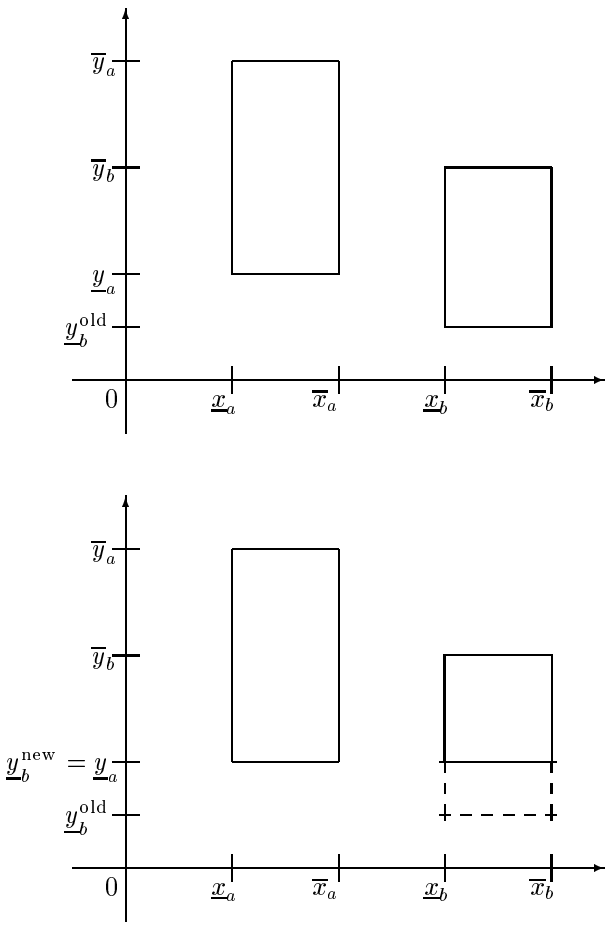
Proof. 1°. We have already shown that if the condition (4) is not satisfied, then we cannot have a monotonic function that is consistent with all the measurement results. Thus, to complete the proof, we must show that if the condition (4) is satisfied, then there exists a monotonic function $f(x)$ that is consistent with all the measurement results.

In this proof, we will therefore assume that the condition (4) is satisfied, and we will construct the desired monotonic function $f(x)$.

2°. Let us show that we can simplify (specifically, narrow down) some of the x- and y-intervals without affecting the existence of a monotonic function that is consistent with all the measurement results, and without changing the condition (4).

The corresponding transformations will enable us to reduce the general case of the theorem to the case when the corresponding x- and y-intervals have already been simplified.

2.1°. Let us first show that for each a and b, if $\overline{x}_a \le \underline{x}_b$, then we can replace the value $\underline{y}_b$ with the new value $\underline{y}_b^{\text{new}} = \max(\underline{y}_a, \underline{y}_b)$ without affecting the existence of a monotonic function that is consistent with all the measurement results and without changing the condition (4).



2.1.1°. Let us first show that the new lower endpoint $\underline{y}_b^{\rm new}$ of b-th y-interval is consistent with the unchanged upper endpoint of this interval, i.e., that $\underline{y}_b^{\rm new} \le \overline{y}_b$.

Indeed, since we assumed that $\overline{x}_a \le \underline{x}_b$, then, due to the condition (4), we will have $\underline{y}_a \le \overline{y}_b$. Clearly, we also have $\underline{y}_b \le \overline{y}_b$. From these two inequalities, we can conclude that $\max(\underline{y}_a, \underline{y}_b) \le \overline{y}_b$, i.e., that $\underline{y}_b^{\rm new} \le \overline{y}_b$.

2.1.2°. Let us now show that the replacement of the old interval $[\underline{y}_b, \overline{y}_b]$ with the new interval $[\underline{y}_b^{\rm new}, \overline{y}_b]$ does not affect the existence of a monotonic function that is consistent with all the measurement results.

By definition, $\underline{y}_b^{\rm new} = \max(\underline{y}_a, \underline{y}_b) \ge \underline{y}_b$, so we replace the lower endpoint $\underline{y}_b$ with a value that is either the same or larger. As a result, the new interval for y_b is a subset of the old one: $\mathbf{y}_b^{\rm new} \subseteq \mathbf{y}_b$. All other x- and y-intervals remain unchanged.

So, if there exists a monotonic function $f(x)$ that is consistent with the new measurement results, i.e., for which, for some $x_i \in \mathbf{x}_i$, we have $f(x_i) \in \mathbf{y}_i$ for $i \ne b$ and $f(x_b) \in \mathbf{y}_b^{\rm new}$, then the value $f(x_b)$ belongs to a larger interval $\mathbf{y}_b$ as well, so the same monotonic function is consistent with the old measurement results.

Vice versa, let us assume that there exists a monotonic function $f(x)$ that is consistent with the old measurement results, i.e., for which, for some $x_i \in \mathbf{x}_i$, we have $f(x_i) \in \mathbf{y}_i$ for $i \ne b$ and $f(x_b) \in \mathbf{y}_b$. In particular, for $x_a \in \mathbf{x}_a$ and $x_b \in \mathbf{x}_b$, we have $x_a \le \overline{x}_a$ and $\underline{x}_b \le x_b$. Since we assumed that $\overline{x}_a \le \underline{x}_b$, we can therefore conclude that $x_a \le x_b$. Since the function $f(x)$ is monotonic, we can therefore conclude that $f(x_a) \le f(x_b)$. Since $f(x_a) \in \mathbf{y}_a$, we conclude that $\underline{y}_a \le f(x_a)$, therefore, from $f(x_a) \le f(x_b)$, we can deduce that $\underline{y}_a \le f(x_b)$.

Similarly, from $f(x_b) \in \mathbf{y}_b$, we can conclude that $\underline{y}_b \le f(x_b) \le \overline{y}_b$. Since $f(x_b)$ is not smaller than both $\underline{y}_a$ and $\underline{y}_b$, it is therefore not smaller than the largest of them, i.e., $\max(\underline{y}_a, \underline{y}_b) \le f(x_b) \le \overline{y}_b$. In other words, we conclude that $f(x_b) \in \mathbf{y}_b^{\rm new}$. So, if the monotonic function $f(x)$ is consistent with the old measurement results, then the same monotonic function $f(x)$ is consistent with the new measurement results as well.

These two conclusions show that the replacement of the old interval $[\underline{y}_b, \overline{y}_b]$ with the new interval $[\underline{y}_b^{\rm new}, \overline{y}_b]$ does not indeed affect the existence of a monotonic function that is consistent with all the measurement results.

2.1.3°. The objective of Part 2.1° of this proof is to show that for each a and b, if $\overline{x}_a \le \underline{x}_b$, then we can replace the value $\underline{y}_b$ with the new value $\underline{y}_b^{\rm new} = \max(\underline{y}_a, \underline{y}_b)$ without affecting the existence of a monotonic function that is consistent with all the measurement results and without changing the condition (4).

We have already proven that this replacement does not affect the existence of a monotonic function that is consistent with all the measurements.

To complete the proof of Part 2.1°, let us now show that the replacement of the old interval $[\underline{y}_b, \overline{y}_b]$ with the new interval $[\underline{y}_b^{\rm new}, \overline{y}_b]$ does not change the

condition (4).

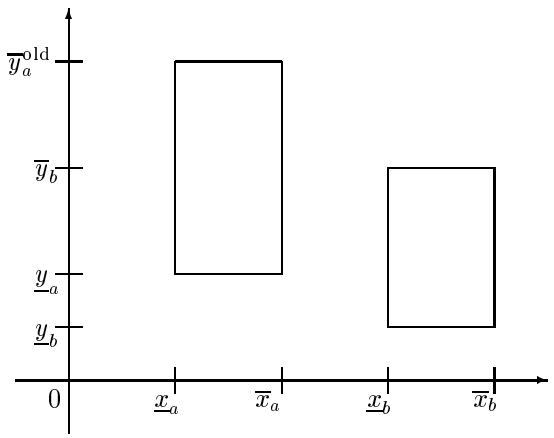
Indeed, we do not change the values $\underline{x}_i$, $\overline{x}_j$, and $\overline{y}_j$, the only change we made is the change in $\underline{y}_b$. Therefore, the only particular case of condition (4) that we need to check is the case when $i = b$, i.e., when $\underline{y}_i$ is the value $\underline{y}_b$ that we change. In other words, assuming that the condition (4) was originally true, we must prove that this condition (4) still holds for the new value $\underline{y}_b$, i.e., that if $\overline{x}_b \le \underline{x}_j$, then $\underline{y}_b^{\text{new}} \le \overline{y}_j$.

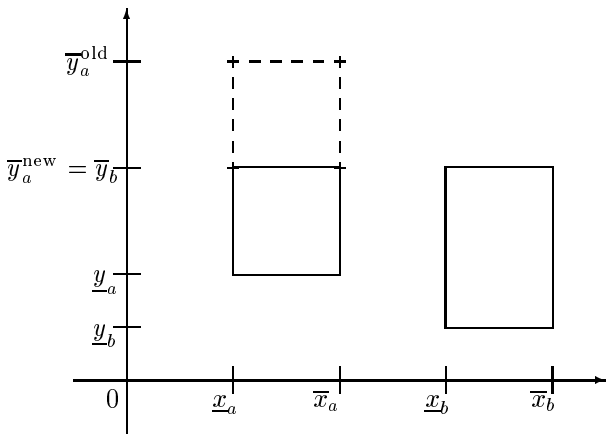
By definition, $\underline{y}_b^{\text{new}} = \max(\underline{y}_a, \underline{y}_b)$. Since the condition (4) is satisfied for the original value $\underline{y}_b$, from $\overline{x}_b \le \underline{x}_j$, we conclude that $\underline{y}_b \le \overline{y}_j$.

Now, the indices a and b are selected in such a way that $\overline{x}_a \le \underline{x}_b$. Since, for every interval, the lower endpoint cannot exceed the upper endpoint, we have $\underline{x}_b \le \overline{x}_b$, and we thus conclude that $\overline{x}_a \le \overline{x}_b$. Since $\overline{x}_b \le \underline{x}_j$, we therefore conclude that $\overline{x}_a \le \underline{x}_j$. Since the condition (4) holds for the original bounds $\underline{y}_i$, we get $\underline{y}_a \le \overline{y}_j$.

We have proven that $\overline{y}_j$ is not smaller than both $\underline{y}_a$ and $\underline{y}_b$. Thus, $\overline{y}_j$ is not smaller than the largest of these two values, i.e., $\max(\underline{y}_a, \underline{y}_b) \le \overline{y}_j$. This is exactly the desired inequality. The statement is proven.

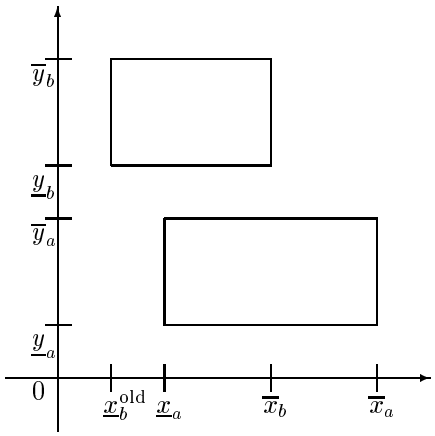
2.2°. Similarly, we can show that for each a and b, if $\overline{x}_a \le \underline{x}_b$, then we can replace the value $\overline{y}_a$ with the new value $\overline{y}_a^{\text{new}} = \min(\overline{y}_a, \overline{y}_b)$ without affecting the existence of a monotonic function that is consistent with all the measurement results and without changing the condition (4).

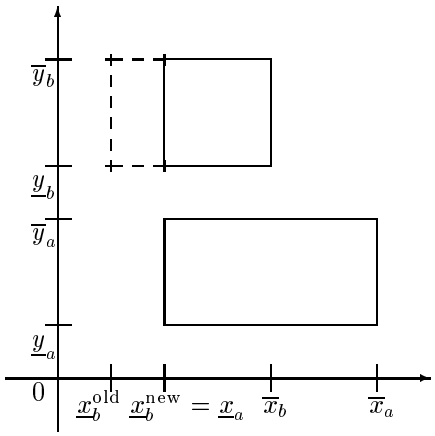




Comment. In 2.1° and 2.2°, we described transformations that simplify the input when $\overline{x}_a \leq \underline{x}_b$ for some a and b. Let us now show that similar transformations can be performed when, instead of considering pairs (a, b) for which $\overline{x}_a \leq \underline{x}_b$, we consider pairs for which $\overline{y}_a < \underline{y}_b$.

2.3°. Let us now show that for each a and b, if $\overline{y}_a < \underline{y}_b$, then we can replace the value $\underline{x}_b$ with the new value $\underline{x}_b^{\text{new}} = \max(\underline{x}_a, \underline{x}_b)$ without affecting the existence of a monotonic function that is consistent with all the measurement results and without changing the condition (4).





2.3.1°. Let us first show that the condition (4) implies the following condition:

$$\text{For every } i \text{ and } j, \text{ if } \overline{y}_i < \underline{y}_j, \text{ then } \underline{x}_i < \overline{x}_j. \tag{6}$$

Indeed, suppose that $\underline{y}_i < \overline{y}_j$ is true. If $\underline{x}_i < \overline{x}_j$ is false, i.e., if $\overline{x}_j \le \underline{x}_i$, then, due to (4), we would have $\underline{y}_j \le \overline{y}_i$ – which contradicts to our assumption that $\underline{y}_i < \overline{y}_j$. This contradiction shows that the inequality $\underline{x}_i < \overline{x}_j$ cannot be false, i.e., it must be true.

2.3.2°. Let us now show that the new lower endpoint $\underline{x}_b^{\text{new}}$ of b-th x-interval is consistent with the unchanged upper endpoint of this interval, i.e., that $\underline{x}_b^{\text{new}} \le \overline{x}_b$.

Indeed, since we assumed that $\overline{y}_a < \underline{y}_b$, then, due to the condition (6), we will have $\underline{x}_a < \overline{x}_b$. Clearly, we also have $\underline{x}_b \le \overline{x}_b$. From these two inequalities, we can conclude that $\max(\underline{x}_a, \underline{x}_b) \le \overline{x}_b$, i.e., that $\underline{x}_b^{\text{new}} \le \overline{x}_b$.

2.3.3°. Let us now show that the replacement of the old interval $[\underline{x}_b, \overline{x}_b]$ with the new interval $[\underline{x}_b^{\text{new}}, \overline{x}_b]$ does not affect the existence of a monotonic function that is consistent with all the measurement results.

By definition, $\underline{x}_b^{\text{new}} = \max(\underline{x}_a, \underline{x}_b) \ge \underline{x}_b$, so we replace the lower endpoint $\underline{x}_b$ with a value that is either the same or larger. As a result, the new interval for x_b is a subset of the old one: $\mathbf{x}_b^{\text{new}} \subseteq \mathbf{x}_b$. All other x- and y-intervals remain unchanged.

So, if there exists a monotonic function $f(x)$ that is consistent with the new measurement results, i.e., for which, for some $x_b \in \mathbf{x}_b^{\text{new}}$ and for some $x_i \in \mathbf{x}_i$ for $i \ne b$, we have $f(x_i) \in \mathbf{y}_i$. If $x_b \in \mathbf{x}_b^{\text{new}}$, then the value x_b belongs to a larger interval $\mathbf{x}_b$ as well, so the same monotonic function is consistent with the old measurement results.

Vice versa, let us assume that there exists a monotonic function $f(x)$ that is consistent with the old measurement results, i.e., for which, for some $x_b \in \mathbf{x}_b^{\rm new}$ and for some $x_i \in \mathbf{x}_i$ for $i \neq b$, we have $f(x_i) \in \mathbf{y}_i$. In particular, for $x_a \in \mathbf{x}_a$ and $x_b \in \mathbf{x}_b$, we have $f(x_a) \leq \overline{y}_a$ and $\underline{y}_b \leq f(x_b)$. Since we assumed that $\overline{y}_a < \underline{y}_b$, we can therefore conclude that $f(x_a) \leq f(x_b)$. Since the function $f(x)$ is monotonic, we can therefore conclude that $x_a < x_b$ (because $x_a \geq x_b$ would then imply $f(x_a) \geq f(x_b)$). Since $x_a \in \mathbf{x}_a$, we conclude that $\underline{x}_a \leq x_a$, therefore, from $x_a < x_b$, we can deduce that $\underline{x}_a \leq x_b$.

Similarly, from $x_b \in \mathbf{x}_b$, we can conclude that $\underline{x}_b \leq x_b \leq \overline{x}_b$. Since x_b is not smaller than both $\underline{x}_a$ and $\underline{x}_b$, it is therefore not smaller than the largest of them, i.e., $\max(\underline{x}_a, \underline{x}_b) \leq x_b \leq \overline{x}_b$. In other words, we conclude that $x_b \in \mathbf{x}_b^{\rm new}$. So, if the monotonic function $f(x)$ is consistent with the old measurement results, then the same monotonic function $f(x)$ is consistent with the new measurement results as well.

These two conclusions show that the replacement of the old interval $[\underline{x}_b, \overline{x}_b]$ with the new interval $[\underline{x}_b^{\rm new}, \overline{x}_b]$ does not indeed affect the existence of a monotonic function that is consistent with all the measurement results.

2.3.4°. Finally, let us show that the replacement of the old interval $[\underline{x}_b, \overline{x}_b]$ with the new interval $[\underline{x}_b^{\rm new}, \overline{x}_b]$ does not change the condition (4).

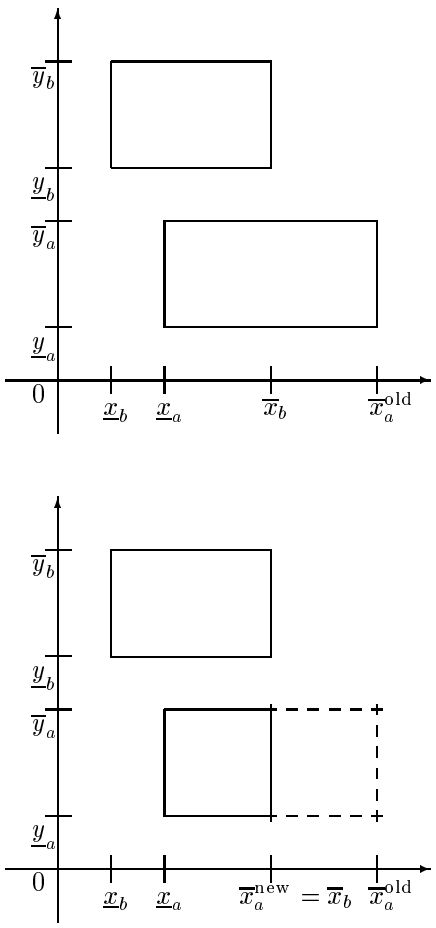
Indeed, we do not change the values $\underline{y}_i$, $\overline{y}_j$, and $\overline{x}_j$, the only change we made is the change in $\underline{x}_b$. Therefore, the only particular case of condition (4) that we need to check is the case when $i = b$, i.e., when $\underline{x}_i$ is the value $\underline{x}_b$ that we change. In other words, assuming that the condition (4) was originally true, we must prove that this condition (4) still holds for the new value $\underline{x}_b$, i.e., that if $\overline{x}_i \leq \underline{x}_b^{\rm new}$, then $\underline{y}_i \leq \overline{x}_b$.

By definition, $\underline{x}_b^{\rm new} = \max(\underline{x}_a, \underline{x}_b)$. If the new value is equal to the old value, then there is no change and, since the property (4) was true, it will be true again.

The only case when we need to prove the property (4) is when the value of $\underline{x}_b$ actually changes, i.e., when $\underline{x}_b^{\rm new} = \underline{x}_a$. In this case, $\overline{x}_i \leq \underline{x}_b^{\rm new}$ means $\overline{x}_i \leq \underline{x}_a$. Since the property (4) holds for the original intervals, we can conclude that $\underline{y}_i \leq \overline{y}_a$.

The transformation is applied if $\overline{y}_a < \underline{y}_b$. From $\underline{y}_i \leq \overline{y}_a$ and $\overline{y}_a < \underline{y}_b$, we conclude that $\underline{y}_i < \underline{y}_b$. Since $\underline{y}_b \leq \overline{y}_b$, we conclude that $\underline{y}_i \leq \overline{y}_b$. This is exactly the desired inequality. The statement is proven.

2.4°. Similarly, we can prove that for each a and b, if $\overline{y}_a < \underline{y}_b$, then we can replace the value $\overline{x}_a$ with the new value $\overline{x}_a^{\rm new} = \min(\overline{x}_a, \overline{x}_b)$ without affecting the existence of a monotonic function that is consistent with all the measurement results and without changing the condition (4).



2.5°. In Parts 2.1–2.4 of this proof, we showed that in some situations, we can narrow down some of the x- and y-intervals without affecting the existence of a monotonic function that is consistent with all the measurement results, and without changing the condition (4).

In some cases, the transformation is applicable, but it does not change the actual $x-$ and y-intervals. For example, in Part 2.1, we showed that for each a and b, if $\overline{x}_a \le \underline{x}_b$, then we can replace the value $\underline{y}_b$ with the new value $\underline{y}_b^{\text{new}} = \max(\underline{y}_a, \underline{y}_b)$. In this case, if $\overline{x}_a \le \underline{x}_b$ and $\underline{y}_a \le \underline{y}_b$, then $\underline{y}_b^{\text{new}} = \max(\underline{y}_a, \underline{y}_b) = \underline{y}_b$, so the intervals do not change.

Let us say that the transformation is *applicable* if we can apply it and decrease the width of at least one interval.

If one of these transformations is applicable, we apply it. If after that, some other transformation is applicable, we apply this other transformation, etc. Let us show that this process will stop after finitely many transformation steps.

Indeed, each transformation changes one of the endpoints of one of the intervals. For example, the only transformation that changes the lower endpoint $\underline{y}_b$ of a y-interval $[\underline{y}_b, \overline{y}_b]$ is the transformation described in Part 2.1: if this transformation is applicable, it replaces the original value $\underline{y}_b$ with the new value $\underline{y}_b^{\text{new}} = \underline{y}_a$, i.e., with the lower endpoints of one of the y-intervals.

In the beginning, before we start applying all these transformations, we have a set $\underline{Y} \stackrel{\text{def}}{=} \{\underline{y}_1, \ldots, \underline{y}_n\}$ of at most n different values $\underline{y}_1, \ldots, \underline{y}_n$ of lower endpoints of y-intervals. After each transformation of type 2.1, one of these lower endpoints changes its value, but since it changes its value to one of the existing lower endpoint values, the new value stays within the set $\underline{Y}$. Therefore, for every a, after each transformation, the new value $\underline{y}_a$ belongs to the original finite set $\underline{Y}$.

After each transformation, each y-interval either stays the same or is replaced by its proper subinterval. Thus, for every a, after each transformation, the lower endpoint $\underline{y}_a$ either remains the same or is replaced by a larger value. Since we stay within the finite set $\underline{Y}$ of no more than n elements, we can only have finitely many such increases (no more than n); so, for every a, no more than n transformations of type 2.1 can change (increase) the lower endpoint $\underline{y}_a$.

Each transformation of type 2.1, when it is applicable, changes (increases) one of the values $\underline{y}_a$. There are n such values, and each value can be changed no more than n times. So, overall, there can be no more than $n \cdot n = n^2$ transformations of type 2.1.

Similarly, there can be only finitely many transformations of the other three types, so overall, we can have only finitely many transformations.

2.6°. We have just shown that after we apply the transformations 2.1–2.4, eventually, we will have a situation in which no more transformations will be applicable. The fact that transformations of type 2.1 are not applicable means the following:

$$\text{For every } a \text{ and } b, \text{ if } \overline{x}_a \le \underline{x}_b, \text{ then } \underline{y}_a \le \underline{y}_b. \tag{7}$$

Similarly, the fact that transformations of types 2.2–2.4 are not applicable mean that the following three conditions hold:

$$\text{For every } a \text{ and } b, \text{ if } \overline{x}_a \le \underline{x}_b, \text{ then } \overline{y}_a \le \overline{y}_b. \tag{8}$$

$$\text{For every } a \text{ and } b, \text{ if } \overline{y}_a < \underline{y}_b, \text{ then } \underline{x}_a \le \underline{x}_b. \tag{9}$$

$$\text{For every } a \text{ and } b, \text{ if } \overline{y}_a < \underline{y}_b, \text{ then } \overline{x}_a \le \overline{x}_b. \tag{10}$$

Therefore, to prove the theorem, it is sufficient to consider the case when the given intervals satisfy the conditions (7)–(10).

3°. To complete the proof of the theorem, we must therefore show that if the given intervals $[\underline{x}_i, \overline{x}_i]$ and $[\underline{y}_i, \overline{y}_i]$ satisfy the conditions (4) and (7)–(10),

then there exists a monotonic function that is consistent with all the given measurement results $\langle [\underline{x}_i, \overline{x}_i], [\underline{y}_i, \overline{y}_i] \rangle$.

3.1°. If two measurement results are the same, i.e., if $[\underline{x}_i, \overline{x}_i] = [\underline{x}_j, \overline{x}_j]$ and $[\underline{y}_i, \overline{y}_i] = [\underline{y}_j, \overline{y}_j]$, then consistency with i-th measurement result means the same as consistency with j-th measurement result. Thus, in this situation, it is sufficient to consider only one of these two identical measurement results.

In other words, without losing generality, we can assume that all measurement results are different, i.e., either their x-intervals differ, or their y-interval differ.

3.2°. Let us sort the measurement results in lexicographic order $\prec$, so that $i \prec k$ if and only if:

- either $\underline{x}_i < \underline{x}_k$,
- or $\underline{x}_i = \underline{x}_k$ and $\underline{y}_i < \underline{y}_k$,
- or $\underline{x}_i = \underline{x}_k$, $\underline{y}_i = \underline{y}_k$, and $\overline{y}_i < \overline{y}_k$, or
- or $\underline{x}_i = \underline{x}_k$, $\underline{y}_i = \underline{y}_k$, $\overline{y}_i = \overline{y}_k$, and $\overline{x}_i < \overline{x}_k$.

Without losing generality, we can assume that the measurement results are already sorted in this order.

3.3°. In order to describe f, we must first define some auxiliary quantities.

On the x-axis, we have lower and upper endpoints of x-intervals. Let ε denote the smallest of the non-zero distances between these endpoints.

For each i, let $v(i)$ denote the number of values $k < i$ for which $\underline{x}_k = \underline{x}_i$.

3.4°. Let us now construct the desired function f. We will construct it by applying linear interpolation to the values (x_i, y_i), where

$$y_i \stackrel{\text{def}}{=} \max(\underline{y}_1, \ldots, \underline{y}_i), \tag{11}$$

and x_i is defined as follows:

$$x_i \stackrel{\text{def}}{=} \underline{x}_i + v(i) \cdot \frac{\varepsilon}{2n}. \tag{12}$$

To complete the proof, we need to prove two things:

- that the resulting function is monotonic, and
- that the resulting function is consistent with all the measurement results $\langle [\underline{x}_i, \overline{x}_i], [\underline{y}_i, \overline{y}_i] \rangle$.

3.5°. Let us first prove that the function defined by linear extrapolation of the values (x_i, y_i) is indeed monotonic. For that, it is sufficient to prove that the corresponding values are monotonic, i.e., that $a < b$ implies $x_a < x_b$ and $y_a \le y_b$.

By definition of y_i, it is easy to see that when $a < b$, we have $y_a \le y_b$ – because then y_b is the maximum of the same values $\underline{y}_1, \ldots, \underline{y}_a$ as y_a and of some additional values $\underline{y}_{a+1}, \ldots, \underline{y}_b$ as well.

It is therefore sufficient to prove that if $a < b$, then $x_a \le x_b$. Indeed, by definition of the lexicographic order, $a < b$ means that we either have $\underline{x}_a < \underline{x}_b$ or $\underline{x}_a = \underline{x}_b$. Let us consider these cases one by one.

3.5.1°. Let us first consider the case when $\underline{x}_a < \underline{x}_b$.

By definition of x_b, we have $x_b \ge \underline{x}_b$.

On the other hand, since ε is the smallest non-zero distance between different x-endpoints, we conclude that ε cannot exceed the distance between $\underline{x}_a$ and $\underline{x}_b$, i.e., $\varepsilon \le \underline{x}_b - \underline{x}_a$. Therefore, $\underline{x}_a + \varepsilon \le \underline{x}_b$.

By definition of $v(a)$, we have $v(a) \le n$, hence

$$x_a = \underline{x}_a + v(a) \cdot \frac{\varepsilon}{2n} \le \underline{x}_a + n \cdot \frac{\varepsilon}{2n} = \underline{x}_a + \frac{\varepsilon}{2} < \underline{x}_a + \varepsilon,$$

so $x_a < \underline{x}_a + \varepsilon \le \underline{x}_b \le x_b$ and $x_a < x_b$.

3.5.2°. Let us now consider the case when $\underline{x}_a = \underline{x}_b$.

In this case, by definition of $v(i)$, we have $v(b) = v(a) + 1$, hence $x_a < x_b$.

3.6°. Let us now prove that the resulting function $f(x)$ is consistent with all the measurement results.

By definition of consistency, this means that for every i, we must find a value $x_i \in [\underline{x}_i, \overline{x}_i]$ for which $f(x_i) \in [\underline{y}_i, \overline{y}_i]$. We will show that these inclusions hold for our selected x_i, when $f(x_i) = y_i$.

3.6.1°. Let us first prove that for all i, we have $x_i \in [\underline{x}_i, \overline{x}_i]$, i.e., $\underline{x}_i \le x_i \le \overline{x}_i$.

By definition of x_i, we always have $\underline{x}_i \le x_i$, so it is sufficient to prove that $x_i \le \overline{x}_i$.

If the i-th x-interval is non-degenerate, i.e., if $\underline{x}_i < \overline{x}_i$, then, since ε is the smallest non-zero distance between different x-endpoints, we conclude that ε cannot exceed the distance between $\underline{x}_i$ and $\overline{x}_i$, i.e., $\varepsilon \le \overline{x}_i - \underline{x}_i$. Therefore, $\underline{x}_i + \varepsilon \le \overline{x}_i$.

On the other hand, by definition of $v(i)$, we have $v(i) \le n$, hence

$$x_i = \underline{x}_i + v(i) \cdot \frac{\varepsilon}{2n} \le \underline{x}_i + n \cdot \frac{\varepsilon}{2n} = \underline{x}_i + \frac{\varepsilon}{2} < \underline{x}_i + \varepsilon,$$

so $x_i < \underline{x}_i + \varepsilon \le \overline{x}_i$.

Let us now prove that if the i-th x-interval is degenerate, i.e., if $\overline{x}_i = \underline{x}_i$, then $v(i) = 0$, hence $x_i = \underline{x}_i \in [\underline{x}_i, \overline{x}_i]$. To prove that $v(i) = 0$, we must show that in this case, we cannot have $\underline{x}_{i-1} = \underline{x}_i$. Let us prove this impossibility by reduction to a contradiction.

Indeed, let us assume that $\underline{x}_{i-1} = \underline{x}_i$. Since the $(i-1)$-th measurement result precedes the i-th result in the lexicographic order, and $\underline{x}_{i-1} = \underline{x}_i$, we have one of the following three possibilities:

1. either $\underline{y}_{i-1} < \underline{y}_i$,
2. or $\underline{y}_{i-1} = \underline{y}_i$ and $\overline{y}_{i-1} < \overline{y}_i$,
3. or $\underline{y}_{i-1} = \underline{y}_i$, $\overline{y}_{i-1} = \overline{y}_i$, and $\overline{x}_{i-1} < \overline{x}_i$.

Since the i-th interval is degenerate and $\underline{x}_{i-1} = \underline{x}_i$, we have $\overline{x}_i = \underline{x}_i = \underline{x}_{i-1}$ hence $\overline{x}_i \le \underline{x}_{i-1}$.

1. Due to the property (7), we conclude that $\underline{y}_i \le \underline{y}_{i-1}$, so we cannot have the first case.
2. Due to the property (8), we conclude that $\overline{y}_i \le \overline{y}_{i-1}$, so we cannot have the second case.
3. Since $\overline{x}_i \le \underline{x}_{i-1}$, we cannot have the third case either.

The contradiction shows that for the degenerate i-th x-interval, we cannot have $\underline{x}_{i-1} = \underline{x}_i$, so for such intervals, $v(i) = 0$, and $x_i \in [\underline{x}_i, \overline{x}_i]$.

The statement is proven.

3.6.2°. Let us now prove that for all i, we have $y_i \in [\underline{y}_i, \overline{y}_i]$, i.e., $\underline{y}_i \le y_i \le \overline{y}_i$.

By definition of y_i as the largest of i values $\underline{y}_1, \ldots, \underline{y}_i$, we have $\underline{y}_i \le y_i$, so it is sufficient to prove that $y_i \le \overline{y}_i$.

By definition of y_i as the largest of the values $\underline{y}_k$ $(k \le i)$, it is therefore sufficient to prove that all these i values do not exceed $\overline{y}_i$, i.e., that $k \le i$ implies that $\underline{y}_k \le \overline{y}_i$.

For $k = i$, we clearly have $\underline{y}_i \le \overline{y}_i$, so it is sufficient to prove this inequality for $k < i$. We will prove this inequality by reduction to a contradiction.

Indeed, let us assume that for some $k < i$, we have $\underline{y}_k > \overline{y}_i$. Since k precedes i in lexicographic order, we have one of the following two possibilities:

1. either $\underline{x}_k < \underline{x}_i$,
2. or $\underline{x}_k = \underline{x}_i$ and $\underline{y}_k \le \underline{y}_i$.

Let us show that none of these cases is possible:

1. Since we have the property (9), from $\overline{y}_i < \underline{y}_k$, we conclude that $\underline{x}_i \le \underline{x}_k$, so the first case is impossible.

2. We cannot have the second case because then, $\underline{y}_k \le \underline{y}_i \le \overline{y}_i$ hence $\underline{y}_k \le \overline{y}_i$, and we assumed that $\underline{y}_k > \overline{y}_i$.

The contradiction shows that $\underline{y}_k > \overline{y}_i$ is impossible, hence $\underline{y}_k \le \overline{y}_i$ and $y_i \in [\underline{y}_i, \overline{y}_i]$.

The theorem is proven.

5 General Case: Algorithm

Possible straightforward algorithm. According to our Theorem, in order to check whether there exists a monotonic function that is consistent with the measurement data, it is sufficient to check the condition (4).

How difficult is it to check this condition? If we simply check this property for all $i = 1, \ldots, n$ and all $j = 1, \ldots, n$, then checking this condition would require $O(n^2)$ comparisons – i.e., $O(n^2)$ computational steps.

Faster algorithm: idea. We can check this condition faster if we use the fact that the condition (4) is equivalent to the following auxiliary property:

$$\text{For every } i, \text{ we have } \underline{y}_i \le \min_{j:\underline{x}_j \ge \overline{x}_i} \overline{y}_j. \tag{13}$$

Faster algorithm: description. To check the existence of a monotonic function that is consistent with all the measurement results, we can perform the following four-stage algorithm:

- First, we sort the values $\underline{x}_i$ into an increasing sequence – this requires $O(n \cdot \log(n))$ steps. We correspondingly re-order the values $\overline{x}_i$, $\underline{y}_i$, and $\overline{y}_i$. After this stage, we can assume that the values $\underline{x}_i$ are sorted:

$$\underline{x}_1 \le \underline{x}_2 \le \cdots \le \underline{x}_n.$$

- Then, for every i from 1 to n, we compute the value $M_i \stackrel{\text{def}}{=} \min(\overline{y}_n, \overline{y}_{n-1}, \ldots, \overline{y}_i)$. Here, $M_n = \overline{y}_n$. If we already know M_i, then we can compute the previous value M_{i-1} by using a single operation $M_{i-1} = \min(M_i, \overline{y}_{i-1})$. Thus, computing all n values requires n computational steps.

- For each i from 1 to n, we can now use binary search (see, e.g., [7]) to find the integer $m(i)$ for which $\underline{x}_{m(i)-1} < \overline{x}_i \le \underline{x}_{m(i)}$ (if such a value exists). Each binary search requires $\log(n)$ computational steps; thus, n such searches require $O(n \cdot \log(n))$ steps.

- Finally, for every i from 1 to n for which $m(i)$ exists, we check whether $\underline{y}_i \le M_{m(i)}$:

- if this inequality holds for all such i, then the measurement data is consistent with monotonicity;
- otherwise, the function $f(x)$ cannot be monotonic.

Each checking requires one comparison, so to check that this inequality holds for all i from 1 to n, we need n comparisons.

Overall, we thus need $O(n \cdot \log(n)) + O(n \cdot \log(n)) + O(n) + O(n) = O(n \cdot \log(n))$ steps.

Faster algorithm: numerical example. In general, our algorithm can be applied to an arbitrary measurement data

$$(\langle [\underline{x}_1, \overline{x}_1], [\underline{y}_1, \overline{y}_1] \rangle, \langle [\underline{x}_2, \overline{x}_2], [\underline{y}_2, \overline{y}_2] \rangle, \ldots, \langle [\underline{x}_n, \overline{x}_n], [\underline{y}_n, \overline{y}_n] \rangle).$$

Let us illustrate the above algorithm on the example of $n = 3$ measurement results:

$$(\langle [1, 5], [2, 5] \rangle, \langle [4, 6], [4, 7] \rangle, \langle [3, 4], [3, 8] \rangle).$$

- On the first stage of the above algorithm, we soft the values $\underline{x}_1 = 1$, $\underline{x}_2 = 4$, and $\underline{x}_3 = 3$ into an increasing sequence $1 \le 3 \le 4$. After we correspondingly rearrange the measurement results, we get the following measurement data:

$$(\langle [1, 5], [2, 5] \rangle, \langle [3, 4], [3, 8] \rangle, \langle [4, 6], [4, 7] \rangle).$$

- On Stage 2, we compute the values M_n:
  - $M_3 = \overline{y}_3 = 7$,
  - $M_2 = \min(M_3, \overline{y}_2) = \min(7, 8) = 7$, and
  - $M_1 = \min(M_2, \overline{y}_1) = \min(7, 5) = 5$.
- On Stage 3, we find the values $m(i)$ for $i = 1, \ldots, n$.
  - For $i = 1$, the value $\overline{x}_1 = 5$ is larger than the largest upper x-endpoint $\underline{x}_3 = 4$, so there is no value $m(1)$.
  - For $i = 2$, we have $\underline{x}_2 = 3 < \overline{x}_2 \le \underline{x}_3 = 4$, so $m(2) = 3$.
  - For $i = 3$, the value $\overline{x}_3 = 6$ is larger than the largest upper x-endpoint $\underline{x}_3 = 4$, so there is no value $m(3)$.
- Finally, on Stage 4, we check whether $\underline{y}_2 \le M_{m(2)} = M_3$. For our data, this inequality turns into $3 < 7$ and is, thus, satisfied. We therefore conclude that there exists a monotonic function that is consistent with the measurement results.

Possibility of parallelization. For large n, we may want to further speed up computations if we have several processors working in parallel. For the case when the values x_i are known precisely, parallelization is described in [25]. Let us show how this parallelization can be extended to the general case when both x_i and y_i are known with interval uncertainty.

In the general case, all fours stages of the above algorithm can be parallelized by known techniques. In particular, Stage 1 is a particular case of a general *prefix-sum* problem, in which we must compute the values a_n, $a_n * a_{n-1}$, $a_n * a_{n-1} * a_{n-2}$, ..., for some associative operation $*$ (in our case, $* = \min$).

If we have a potentially unlimited number of processors, then we can do the following (see, e.g., [13], for the information on how to parallelize the corresponding stages):

- on Stage 1, we can sort the values $\underline{x}_i$ in time $O(\log(n))$;
- on Stage 2, we can compute the values M_i (i.e., solve the prefix-sum problem) in time $O(\log(n))$;
- on Stage 3, we can use n processors, each of which compute the corresponding value $m(i)$ in time $O(\log(n))$;
- finally, on Stage 4, we can use n processors, each of which checks the corresponding inequality in time $O(1)$.

As a result, we can check monotonicity in time

$$O(\log(n)) + O(\log(n)) + O(\log(n)) + O(1) = O(\log(n)).$$

If we have $p < n$ processors, then we can:

- on Stage 1, sort n values in time $O((n \cdot \log(n))/p + \log(n))$ [13];
- on Stage 2, compute the values M_i in time $O(n/p + \log(p))$ [3];
- on Stage 3, we subdivide n indices i between p processors, so each processor computes $m(i)$ for n/p indices i; computing each index requires $\log(n)$ time, so the overall time is $(n/p) \cdot \log(n) = O((n \cdot \log(n))/p)$;
- finally, on Stage 4, each of p processors checks the desired inequality for its n/p indices; this requires time $O(n/p)$.

Overall, we thus need time

$$O\left(\frac{n \cdot \log(n)}{p} + \log(n)\right) + O\left(\frac{n}{p} + \log(p)\right) + O\left(\frac{n \cdot \log(n)}{p}\right) + O\left(\frac{n}{p}\right) =$$

$$O\left(\frac{n \cdot \log(n)}{p} + \log(n) + \log(p)\right).$$

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209, by the Future Aerospace Science and Technology Program (FAST) Center for Structural Integrity of Aerospace Systems, effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant F49620-00-1-0365, by NSF grants EAR-0112968, EAR-0225670, and EIA-0321328, and the Army Research Laboratories grant DATM-05-02-C-0046.

References

[1] M. S. Aguiar, A. C. R. Costa, and G. P. Dimuro, "ICTM: an interval tessellation-based model for reliable topographic segmentation", *Numerical Algorithms* (to appear).

[2] M. S. Aguiar, G. P. Dimuro, A. C. R. Costa, R. K. S. Silva, and V. Kreinovich, "HPC-ICTM: the interval categorizer tesselation-based model for high performance computing", *Proceedings of the Workshop on State-of-the-Art in Scientific Computing PARA'04*, Lyngby, Denmark, June 20–23, 2004.

[3] G. E. Blelloch, "Prefix sums and their applications", In: J. H. Reif, *Synthesis of Parallel Algorithms*, Morgan Kaufmann, San Mateo, California, 1993, pp. 35–60.

[4] B. Bouchon-Meunier and V. Kreinovich, "From Interval Computations to Modal Mathematics: Applications and Computational Complexity", *ACM SIGSAM Bulletin*, 1998, Vol. 32, No. 2, pp. 7–11.

[5] D. D. Coblentz, V. Kreinovich, B. S. Penn, and S. A. Starks, "Towards Reliable Sub-Division of Geological Areas: Interval Approach" *Proceedings of the 19th International Conference of the North American Fuzzy Information Society NAFIPS'2000*, Atlanta, Georgia, July 13–15, 2000, pp. 368–372.

[6] D. D. Coblentz, V. Kreinovich, B. S. Penn, and S. A. Starks, "Towards Reliable Sub-Division of Geological Areas: Interval Approach", In: L. Reznik and V. Kreinovich (eds.), *Soft Computing in Measurements and Information Acquisition*, Springer-Verlag, Berlin-Heidelberg, 2003, pp. 223–233.

[7] Th. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, MIT Press, Cambridge, MA, 2001.

[8] G. Deboeck, "Fuzzy logic, neural networks, and genetic algorithms for financial trading systems", *NAFIPS/IFIS/NASA'94, Proceedings of the*

First International Joint Conference of The North American Fuzzy Information Processing Society Biannual Conference, The Industrial Fuzzy Control and Intelligent Systems Conference, and The NASA Joint Technology Workshop on Neural Networks and Fuzzy Logic, San Antonio, December 18–21, 1994 (plenary talk).

[9] G. J. Deboeck, K. Villaverde, and V. Kreinovich, "Interval Methods for Presenting Performance of Financial Trading Systems", *Reliable Computing*, 1995, Supplement (Extended Abstracts of APIC'95: International Workshop on Applications of Interval Computations, El Paso, TX, Febr. 23–25, 1995), pp. 67–70.

[10] E. Gardeñes, H. Mielgo, and A. Trepat, "Modal intervals: reason and ground semantics", In: K. Nickel (ed.), *Interval Mathematics 1985*, Lecture Notes in Computer Science, Vol. 212, Springer-Verlag, Berlin, Heidelberg, 1986, pp. 27–35.

[11] E. Gardeñes, M. A. Sainz, L. Jorba, R. Calm, R. Estela, H. Mielgo, and A. Trepat, "Modal Intervals", *Reliable Computing*, 2001, Vol. 7, No. 2, pp. 77–111.

[12] I. J. Good and R. A. Gaskins, "Density Estimation and Bump-Hunting by the Penalized Likelihood Method Exemplified by Scattering and Meteorite Data", *J. Amer. Stat. Soc.*, 1980, Vol. 75, pp. 42–56.

[13] J. Jájá, *An Introduction to Parallel Algorithms*, Addison-Wesley, Reading, MA, 1992.

[14] V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, *Computational Complexity and Feasibility of Data Processing and Interval Computations*, Kluwer, Dordrecht, 1997.

[15] V. Kreinovich, H. T. Nguyen, G. P. Dimuro, A. C. R. Costa, and B. R. C. Bedregal, "A New Differential Formalism for Interval-Valued Functions and Its Potential Use in Detecting 1-D Landscape Features", *Proceedings of the International Conference on Information Technology InTech'03*, Chiang Mai, Thailand, December 17–19, 2003, pp. 491–498.

[16] V. Kreinovich, E. J. Pauwels, S. Ferson, and L. Ginzburg, "A Feasible Algorithm for Locating Concave and Convex Zones of Interval Data and Its Use in Statistics-Based Clustering", *Numerical Algorithms* (to appear).

[17] J. Lorkowski and V. Kreinovich, "If we measure a number, we get an interval. What if we measure a function or an operator?", *Reliable Computing*, 1996, Vol. 2, No. 3, pp. 287–298.

[18] K. Nickel, "Interval Acceleration of Convergence", in: R. E. Moore (editor), *Reliability in Computing*, Academic Press, N.Y., 1988, pp. 151–169.

[19] E. J. Pauwels and G. Frederix, "Image Segmentation by Nonparametric Clustering Based on the Kolmogorov-Smirnov Distance", In: D. Vernon (ed.), *Proceedings of the 6th European Conference on Computer Vision ECCV'2000, Dublin, Ireland, June 26–July 1, 2000, Part II*, Springer Lecture Notes in Computer Science, Vol. 1843, 2000, pp. 85–99.

[20] E. J. Pauwels, G. Frederix, and G. Caenaen, "Image segmentation based on statistically principled clustering", *Proc. of 2001 International IEEE Conference on Image processing*, 2001, Vol. 3, pp. 66–69.

[21] B. W. Silverman, *On a Test for Multimodality Based on Kernel Density Estimates*, Technical Summary Report No. 21–81, U. S. Army Math Research Center, Madison, WI, February 1981.

[22] D. J. Thompson, "Spectrum Estimation Techniques for Characterization and Development of WT4 Waveguide", *Bell Syst. Technical Journal*, Part 1, November 1977, Vol. 56, p. 1769; Part 2, December 1977, Vol. 57, p. 1983.

[23] D. J. Thompson, "Spectrum Estimation and Harmonic Analysis", *Proceedings IEEE*, 1982, Vol. 70, pp. 1055–1096.

[24] K. Villaverde and V. Kreinovich, "A linear-time algorithm that locates local extrema of a function of one variable from interval measurement results," *Interval Computations*, 1993, No. 4, pp. 176-194.

[25] K. Villaverde and V. Kreinovich, "Parallel algorithm that locates local extrema of a function of one variable from interval measurement results", *Reliable Computing*, 1995, Supplement (Extended Abstracts of APIC'95: International Workshop on Applications of Interval Computations, El Paso, TX, Febr. 23–25, 1995), pp. 212–219.