Towards the Possibility of Objective Interval Uncertainty

Luc Longpré, Olga Kosheleva, and Vladik Kreinovich

University of Texas at El Paso, El Paso, TX 79968, USA
longpre@utep.edu, olgak@utep.edu, vladik@utep.edu

Abstract. Applications of interval computations usually assume that while we only know an interval containing the actual (unknown) value of a physical quantity, there *is* the exact value of this quantity, and that in principle, we can get more and more accurate estimates of this value. Physicists know, however, that, due to the uncertainty principle, there are limitations on how accurately we can measure the values of physical quantities. One of the important principles of modern physics is *operationalism* – that a physical theory should only use observable properties. This principle is behind most successes of the 20th century physics, starting with relativity theory (vs. un-observable aether) and quantum mechanics. From this viewpoint, it is desirable to avoid using un-measurable exact values and to modify the mathematical formalisms behind physical theories so that they explicitly only take objective uncertainty into account. In this paper, we describe how this can be done for objective interval uncertainty.

Keywords: interval uncertainty, algorithmic randomness, physics

1 Formulation of the Problem

Is interval uncertainty subjective? Applications of interval computations usually assume that while we only know an interval $[\underline{x}, \overline{x}]$ containing the actual (unknown) value of a physical quantity x, there *is* the exact value x of this quantity, and that in principle, we can get more and more accurate estimates of this value.

This assumption is in line with the usual formulations of physical theories – as partial differential equations relating exact values of different physical quantities, fields, etc., at different space-time locations and moments of time; see, e.g., [2]. Physicists know, however, that, due to the uncertainty principle, there are limitations on how accurately we can measure the values of physical quantities [2, 8].

This is not just a theoretical concern: for example, the International Union of Pure and Applied Chemistry (IUPAC) has recently officially recognized that the atomic weight of a chemical element is not an exact number; depending on where the sample came from, the atomic weight may differ within a certain interval; see, e.g., [11].

It is desirable to take objective uncertainty into account. One of the important principles of modern physics is *operationalism* – that a physical theory should only use observable properties. This principle is behind most successes of 20th century physics, starting with relativity theory (vs. un-observable aether) and quantum mechanics. From this viewpoint, it is desirable to avoid using unmeasurable exact values and to modify the mathematical formalisms behind physical theories so that they explicitly only take objective uncertainty into account.

Objective uncertainty is about probabilities. According to quantum physics, we can only predict probabilities of different events. Thus, uncertainty means that instead of exact values of these probabilities, we can only determine intervals; see, e.g., [3, 4].

Let us give a simple example. In the ideal world, atoms of the same element have exactly the same atomic weight, and – if their are radioactive – exactly the same probability p that they will decay within a given moment of time. In such an ideal situation, by taking larger and larger samples and measuring the decay frequency, we can get more and more accurate estimates of the desired decay probability p.

In practice, as we have mentioned, different objects made of the same element have, in general, slightly different average atomic weight – and similarly, they have slightly different average decay probabilities, probabilities that may take different values from the corresponding interval $[\underline{p}, \overline{p}]$. Depending on which objects we take, we may get frequencies close to the lower bound $\underline{p}$ and we may also get frequencies close to the upper bound $\overline{p}$. As we increase the sample size, we will get frequencies oscillating between $\underline{p}$ and $\overline{p}$, without ever converging to a single value.

Formulation of the problem. What is the observational meaning of such interval-valued probabilities?

2 Analysis of the Problem

What is the observational meaning of probability? Probability refers to repeated events: we repeat the same experiment (or perform many similar observations) and record the results as a binary sequence $\omega_1\omega_2\ldots$ For example, when we talk about the probability of a coin falling heads, we mean that we repeatedly flip the coin and record the resulting sequence: for example, we can take $\omega_i = 1$ if the coin falls heads in the i-th experiment, and $\omega_i = 0$ if this coin falls tails.

In there terms, the fact that the probability of heads is $1/2$ means that in the limit, when $n \to \infty$, the ratio of 1s in a sequence $\omega_1 \ldots \omega_n$ tends to $1/2$. However, this is only the part of this meaning: for example, for a sequence $0101\ldots$, the ratio tends to $1/2$, but we would not call it a random sequence corresponding to probability $1/2$.

From the practical viewpoint, when we say that a sequence $\omega_1\omega_2\ldots$ is random, we assume that this sequence satisfies *all* the probability laws (such as the

law of large numbers or the Central Limit Theorem); these probability laws are what practitioners use to check whether the sequence is random.

From this viewpoint, if a sequence satisfies all probability laws, then for all practical purposes we can consider it random. Thus, we can formally define a sequence to be random if it satisfies all probability laws. In precise terms, a probability law is a property ℓ which is true with probability 1: $P(\ell) = 1$. So, a sequence is random if it satisfies all the properties which are true with probability 1.

Properties are in 1-1 correspondence with sets – to each property, we can assign the set of all the sequences that satisfy this property and, vice versa, to every set, we can assign a property of belonging to this set. When we talk about probability laws, we mean only properties which can be described by finitely many symbols from a certain formal language; the corresponding sets are known as *definable sets*. Thus, we can say that a sequence is random if it belongs to all definable sets of probability measure 1.

A sequence belongs to a set of measure 1 if and only if it does not belong to its complement $C = -S$ with $P(C) = 0$. So, we can equivalently say that *a sequence is* random *if it does not belong to any definable set of measure 0.* This is, in effect, Kolmogorov-Martin-Löf's (KML) definition of a random sequence; see, e.g., [7].

Each definable set is determined by a finite sequence of symbols. There are no more than countably many finite sequences of symbols, thus, there are countably many definable sets. So, the union of all such sets has measure 0. Therefore, almost all sequences are KML-random.

Probability interval: what is its observational meaning? We have recalled what is an observational meaning of an exact probability p. What is the observational meaning of a probability interval, when instead of a single probability measure we have several possible probability measures?

This is not an easy question: in [1, 6], we have shown that in seemingly reasonable formalizations, every random sequence is actually random relative to one of the possible probability measures. In such a formalization, every random sequence – including the sequence of observations – corresponds to one specific probability measure. In other words, there is a probability, we just do not know it – this is exactly the subjective interval uncertainty that we are trying to avoid.

We consider independent repeated events. Probabilities have direct observational meaning only for repeating events. In mathematical terms, independent repeating events correspond to a *product measure*, when the probability of two events A and B happening in two consequent tests is equal to the product of the corresponding probabilities: $P(A \,\&\, B) = P(A) \cdot P(B)$.

Traditional case. The traditional case is when we know the exact probability p. Then, observable sequences $\omega_1\omega_2\ldots$ are KLM-random relative to a product of p-measures.

It may be that in practice, we do not know the exact value of this probability p, we only know the interval $[\underline{p}, \overline{p}]$ containing this probability. In this case,

we have an interval uncertainty, but this interval uncertainty is *subjective* in the following sense: there is the actual exact value of the probability, the value which can be determined, e.g., by taking larger and larger samples; then the corresponding frequencies will be closer and close to the actual probability.

What we are trying to describe. What we are trying to describe is the case when there is no such objective probability: e.g., the case when the corresponding frequencies do not have an exact limit: limit frequencies oscillate between $\underline{p}$ and $\overline{p}$.

3 Objective Interval Uncertainty: Definitions and the First Result

Definition 1. We say that a sequence is $[\underline{p}, \overline{p}]$-*random* if it is random for some product measure with $p_i \in [\underline{p}, \overline{p}]$.

Definition 2. *We say that a sequence $\omega_1\omega_2\ldots$ is objectively $[\underline{p}, \overline{p}]$-random if this sequence is $[\underline{p}, \overline{p}]$-random, and it is not $[\underline{q}, \overline{q}]$-random for any proper subinterval $[\underline{q}, \overline{q}] \subset [\underline{p}, \overline{p}]$.*

Proposition 1. *For every interval $[\underline{p}, \overline{p}]$, there exist objectively $[\underline{p}, \overline{p}]$-random sequences.*

Proof. We will show that any sequence $\omega_1\omega_2\ldots$ corresponding to p_i for which $\liminf p_i = \underline{p}$ and $\limsup p_i = \overline{p}$ is objectively $[\underline{p}, \overline{p}]$-random.

Since $p_i \in [\underline{p}, \overline{p}]$, this sequence is $[\underline{p}, \overline{p}]$-random. Let us prove that this sequence $\omega_1\omega_2\ldots$ is not $[\underline{q}, \overline{q}]$-random for any proper subinterval $[\underline{q}, \overline{q}] \subset [\underline{p}, \overline{p}]$, i.e., that it is not random w.r.t. any sequence $q_i \in [\underline{q}, \overline{q}]$.

It is known that if two measures are mutually singular, then no sequence is random w.r.t. both measures. For product measures, singularity is equivalent to the following equality (see, e.g., [7, 8]):

$$\sum_{i=1}^{\infty}\left[(\sqrt{p_i}-\sqrt{q_i})^2+\left(\sqrt{1-p_i}-\sqrt{1-q_i}\right)^2\right]=+\infty.$$

For a proper subinterval, either $\underline{p} < \underline{q}$ or $\overline{q} < \overline{p}$. Without loss of generality, let us consider the case when $\underline{p} < \underline{q}$.

When $\liminf p_i = \underline{p}$ then, for every $\varepsilon > 0$, there are infinitely many i for which $\sqrt{p_i} \le \sqrt{\underline{p}} + \varepsilon$. For these i, we have $q_i \ge \underline{q}$, so $\sqrt{q_i} \ge \sqrt{\underline{q}}$. Thus, $\sqrt{q_i} - \sqrt{p_i} \ge \sqrt{\underline{q}} - \left(\sqrt{\underline{p}} + \varepsilon\right) = \left(\sqrt{\underline{q}} - \sqrt{\underline{p}}\right) - \varepsilon$. For $\varepsilon = (\sqrt{\underline{q}} - \sqrt{\underline{p}})/2$, we have $\sqrt{q_i} - \sqrt{p_i} > \varepsilon > 0$ and therefore, the above sum is infinite. So, a $\{p_i\}$-random sequence $\omega_1\omega_2\ldots$ cannot be $\{q_i\}$-random. The proposition is proven.

4 Objective Interval Uncertainty: A Stronger Definition and the Second Result

Discussion. We want to describe the idea that all we know is an interval $[\underline{p}, \overline{p}]$.

The above definition means, in effect, that all the values p_i from the sequence p_i are in between $\underline{p}$ and $\overline{p}$ and that, even if we dismiss finitely many probabilities, no narrower interval contains all the remaining values of p_i.

In general, however, such sequences may satisfy additional laws, in addition to $p_i \in [\underline{p}, \overline{p}]$. For example, if we have $p_{2i} = \underline{p}$ and $p_{2i+1} = \overline{p}$, then we satisfy the above condition – but we also satisfy the additional condition, that all even-placed probabilities are equal to $\underline{p}$ and all odd-placed probabilities are equal to $\overline{p}$.

Is it possible to have a sequence of probabilities p_i whose *only* meaningful property is that all these values are from the interval $[\underline{p}, \overline{p}]$? In other words, is it possible to find a sequence p_i which does not satisfy any other meaningful property?

What is a "meaningful property". In order to answer the above question, we need to formalize what is meant by a meaningful property.

In foundations of mathematics, the main object is a set. Properties naturally correspond to sets: namely, to each property, we can put into correspondence the set of all the sequences that satisfy this property. In these terms, describing what we mean by a property is equivalent to describing the corresponding sets.

First, a meaningful property must be described by a finite sequence of symbols in an appropriate mathematical language. Corresponding sets are known as *definable* sets. It is important to realize that while every example of a set that we can give is definable, not all sets are definable: for example, there are more than countably many subsets of the set of all natural numbers, but since there are only countably many finite sequences, there are only countably many definable sets.

Second, it is reasonable to only consider *observable* properties, i.e., properties whose validity can be determined based on observations. Let us show that, because of this requirement, it is reasonable to require that for each observable property, the corresponding set of sequences is *closed* in the sense of component-wise convergence: if for all k, the sequences $p^{(k)} = \{p_i^{(k)}\}$ belong to this set, then their limit $p = \{p_i\}$, with $p_i \stackrel{\text{def}}{=} \lim\limits_k p_i^{(k)}$, should also belong to the corresponding set.

Indeed, in practice, we do not observe probabilities, we only observe *frequencies* which are close to probabilities. If the actual probabilities are the limit values $(p_1, p_2, \ldots)$, this means that for every $\varepsilon > 0$ and for a sufficiently large sample, we will observe frequencies f_i which are ε-close to these limit value p_i. Since p is the limit of $p^{(k)}$, for sufficiently large k, the values p_i – and thus, the frequencies f_i – are close to the probabilities $p_i^{(k)}$, and therefore, consistent with the assumption that the actual probabilities are $p_i^{(k)}$ (and thus, with the assumption that the actual probabilities satisfy the given property).

So, if the actual probabilities are equal to the limit, then no matter how large a sample we take, the resulting observations will always be consistent with the given property. Thus, it is reasonable to add the limit p to the set of all probability sequences that satisfy the given property. Because of this argument,

in the following text, we will assume that for each observable property, the corresponding set of sequences is closed.

The third property of the corresponding sets comes from the need to distinguish trivial unavoidable "properties" like $p_1 = \underline{p}$ – properties that do not really restrict any values beyond a few first ones – from non-trivial properties that we are trying to avoid. In other words, we need to formulate the idea that if we only know approximate values of the first n probabilities, then we cannot guarantee that the corresponding property will be satisfied.

This requirement can be described in precise terms, if on the set of all the sequences p we introduce a topology in which the basis is formed by "boxes" $(\underline{p}_1, \overline{p}_1) \times \ldots \times (\underline{p}_n, \overline{p}_n)$ corresponding to different n and different bounds $\underline{p}_i$ and $\overline{p}_i$. (Convergence in this topology corresponds to the above point-wise convergence.) In terms of this topology, the above requirement means that within every element from the basis – and thus, within every open set – there should be a sequence that does not belong to the corresponding set S. For closed sets, this requirement means that the set S is *nowhere dense*.

Summarizing, we can formalize our requirements by saying that by a meaningful property, we mean a closed nowhere dense definable set, and that the actual sequence of probabilities p should not belong to any of such sets.

Definition 3. *Let $\mathcal{S}$ be a set of all sequences $p_i \in [\underline{p}, \overline{p}]$, with topology whose basis is formed by the boxes $(\underline{p}_1, \overline{p}_1) \times \ldots \times (\underline{p}_n, \overline{p}_n)$.*

- *By a* meaningful property*, we mean a definable closed nowhere dense set $S \subseteq \mathcal{S}$.*
- *We say that a sequence p satisfies a meaningful property S if $p \in S$.*
- *We say that a sequence $p \in \mathcal{S}$* has no other properties *if p does not satisfy any meaningful property S.*

Proposition 2. *For every two definable values $\underline{p} < \overline{p}$, there exists a sequence p_i for which $p_i \in [\underline{p}, \overline{p}]$ for all i and which has no other properties.*

Mathematical comment. Not only there *exist* such sequences, but there are *many* such sequences: as we can see from the proof, "almost all" sequences $p \in \mathcal{S}$ (almost all in some reasonable sense) have no other properties.

Practical comment. The above description is, at this stage, very theoretical. We do not have a full understanding of how to check whether an experimentally observed sequence has no other properties. One consequence that we *can* check is that the limit frequencies should fill the whole interval $[\underline{p}, \overline{p}]$ and not be bound by any narrower subinterval. In other words, as we increase the sample size, the frequencies should always fill this whole interval.

However, we may recall that in statistics, in addition to observing frequencies, there are other criteria that describe p-random sequences – e.g., according to the Central Limit Theorem, deviations of frequencies from the probability must be, asymptotically, normally distributed. It is therefore desirable to come up with similar criteria for the case of sequences which are $[\underline{p}, \overline{p}]$-random for some non-degenerate interval $\underline{p} < \overline{p}$.

Proof. By definition, a sequence p has no other properties if its does belong to any property-related set S (in the sense of Definition 3), i.e., equivalently, if it does not belong to the union U of all such sets S.

Each property-related set S is, by definition, a definable closed nowhere dense set. As we have mentioned, there are no more than countably many definable objects, so U is a union of countably many closed nowhere dense sets. Such unions are known as *meager* sets, or sets of *first Baire category*. It is known that the set of all sequences is *not* meager; this is the main gist of the corresponding *Baire's theorem*; see, e.g., [9]. Thus, there are sequences p which do not belong to U, i.e., which have no other properties. Moreover, "almost all" sequences p – in the sense of all sequences except for a meager set – do not belong to U, i.e., have no other properties. The proposition is proven.

5 Why This Is Interesting: Objective Interval Uncertainty Can Potentially Help in Solving NP-Hard Problems Faster

What we do in this section. Objective interval uncertainty means that the corresponding series of repeated experiments, the sequence of observations $\omega_1\omega_2\ldots$ is random with respect to some sequence of probabilities p_i for which $p_i \in [\underline{p}, \overline{p}]$ and which has no other property.

In this section, we prove that by using such sequences ω, it is, in principle, possible to drastically speed up the solution of NP-complete problems.

Practical comment. It should be emphasized that our result only says that it is *theoretically* possible to speed up the solution of NP-hard problem. At this point, we do not know how to actually achieve such a speed-up – but we hope that the proof of theoretically possibility of the speed-up will eventually lead to practical algorithms.

Mathematical comment. From the mathematical viewpoint, the result from this section is a modification of a similar result from [5].

What is an NP problem? Brief reminder. In practice, we often need to find a solution that satisfies a given set of constraints or at least check that such a solution is possible. Once we have a candidate for the solution, we can feasibly check whether this candidate indeed satisfies all the constraints. In theoretical computer science, "feasibly" is usually interpreted as computable in polynomial time, i.e., in time bounded by a polynomial of the length of the input.

A problem of checking whether a given set of constraints has solution is called a *problem of the class NP* if we can check, in polynomial time, whether a given candidate is a solution; see, e.g., [10]. Examples of such problem includes checking whether a given graph can be colored in 3 colors, checking whether a given propositional formula – i.e., formula of the type

$$(v_1 \vee \neg v_2 \vee v_3)\,\&\,(v_4 \vee \neg v_2 \vee \neg v_5)\,\&\,\ldots$$

is satisfiable, i.e., whether this formula is true by some combination of the propositional variables v_i.

Each problem from the class NP can be algorithmically solved by trying all possible candidates. For example, we can check whether a graph can be colored by trying all possible assignments of colors to different vertices of a graph, and we can check whether a given propositional formula is satisfiable by trying all 2^n possible combinations of true-or-false values $v_1, \ldots, v_n$. Such exhaustive search algorithms require computation time like 2^n, time that grows exponentially with n. For medium-size inputs, e.g., for $n \approx 300$, the resulting time is larger than the lifetime of the Universe. So, these exhaustive search algorithms are not practically feasible.

It is not known whether problems from the class NP can be solved feasibly (i.e., in polynomial time): this is a famous open problem $\mathrm{P} \stackrel{?}{=} \mathrm{NP}$. It is known, however, there are problems in the class NP which are *NP-complete* in the sense that every problem from the class NP can be reduced to this problem. Reduction means, in particular, that if we can find a way to efficiently solve one NP-complete problem, then, by reducing other problems from the class NP to this problem, we can thus efficiently solve all the problems from the class NP.

So, it is very important to be able to efficiently solve even one NP-complete problem. (By the way, both above example of NP problems – checking whether a graph can be colored in 3 colors and whether coloring a propositional formula is satisfiable – are NP-complete.)

How to represent instances of an NP-complete problem. For each NP-complete problem $\mathcal{P}$, its instances are sequences of symbols. In the computer, each such sequence is represented as a sequence of 0s and 1s. Thus, we can append 1 in front of this sequence and interpret the resulting sequence as a binary code of a natural number i (we need to add 1 in front, so that different sequences transform into different numbers, otherwise 0 and 00 will lead to the same number).

In principle, not all natural numbers i correspond to instances of a problem $\mathcal{P}$; we will denote the set of all natural numbers which correspond to such instances by $S_{\mathcal{P}}$. For each $i \in S_{\mathcal{P}}$, the correct answer (true or false) to the i-th instance of the problem $\mathcal{P}$ will be denoted by $s_{\mathcal{P},i}$.

Easier-to-solve and harder-to-solve NP-complete problems. For some easier-to-solve problems, there are feasible algorithms which solve "almost all" instances, in the sense that for each n, the proportion of instance $i \leq n$ for which the problem is solved by this algorithm tends to 1. In this case, while the worst-case complexity is still exponential, in practice, almost all problems can be feasibly solved.

A more challenging case are harder-to-solve NP-complete problems, for which no feasible algorithm is known that would solve almost all instances. In this section, we show that our method works on all NP-complete problems, both easier-to-solve and harder-to-solve ones.

What we mean by using physical observations in computations. We assume that the sequence ω_i comes from observations. In addition to performing

computations, our computational device can, given a natural number i, use the result ω_i of the corresponding i-th observation in its computations. In other words, given an integer i, we can produce ω_i.

In precise theory-of-computation terms, this means computations that use the sequence ω as an oracle; see, e.g., [10].

Comment. Since we are interested in feasible (= polynomial time) computations, the code should be set up in such a way that the overall time of an experiment does not exceed a polynomial of the length of the number i. This can be done, e.g., if we explicitly add maximum waiting time into the description of the experiment, by adding as many 0s as the time that we plan to wait.

Definition 4. *By a $[\underline{p},\overline{p}]$-algorithm $\mathcal{A}$, we mean an algorithm which uses, as an oracle, a sequence ω_i which is random with respect to a probability measure determined by a sequence p_i for which $p_i \in [\underline{p},\overline{p}]$ for all i and which has no other properties.*

Notation. The result of applying an algorithm $\mathcal{A}$ using ω_i to an input i will be denoted by $\mathcal{A}(\omega, i)$.

Definition 5. *Let $\mathcal{P}$ be an NP-complete problem. We say that a feasible $[\underline{p},\overline{p}]$-algorithm $\mathcal{A}$* solves almost all instances of $\mathcal{P}$ *if for every $\varepsilon > 0$ and $\delta > 0$ and for every integer n, there exists an integer $N \geq n$ for which, with probability $\geq 1-\delta$, the proportion of the instances $i \leq N$ of the problem $\mathcal{P}$ which are correctly solved by $\mathcal{A}$ is greater than $1-\varepsilon$:*

$$\mathrm{Prob}\left(\frac{\#i\{i \leq N : i \in S_{\mathcal{P}} \,\&\, \mathcal{A}(\omega, i) = s_{\mathcal{P},i}\}}{\#i\{i \leq N : i \in S_{\mathcal{P}}\}} > 1-\varepsilon\right) \geq 1-\delta.$$

Comment. The restriction to sufficiently long inputs $N \geq n$ makes perfect sense: for short inputs, NP-completeness is not an issue: we can perform exhaustive search of all possible bit sequences of length 10, 20, and even 30. The challenge starts when the length of the input is high.

Proposition 3. *For every NP-complete problem $\mathcal{P}$, there exists a feasible $[\underline{p},\overline{p}]$-algorithm $\mathcal{A}$ that solves almost all instances of $\mathcal{P}$.*

Comment. In other words, we show that if there is objective interval uncertainty, then, *theoretically*, the use of the corresponding physical observations makes all NP-complete problems easier-to-solve (in the above-described sense).

Of course, as we have mentioned earlier, this does not mean that we already have an efficient algorithm for solving NP-complete problems – but this theoretical possibility is encouraging, and we hope that it will eventually lead to efficient algorithms.

Proof. We know that for every i, the probability p_i that $\omega_i = 1$ is in between $\underline{p}$ and $\overline{p}$. Thus, for every two numbers $N \gg N'$, the proportion of values ω_i $(i = N, N+1, \ldots, N'-1)$ which are equal to 1, should be either within the interval $[\underline{p},\overline{p}]$ or at least close to this interval. Let us use this property to design the desired algorithm $\mathcal{A}$.

A value p from the interval $[\underline{p},\overline{p}]$ is:

– closer to $\overline{p}$ if it is larger than the midpoint $\widetilde{p} \stackrel{\text{def}}{=} \dfrac{\underline{p}+\overline{p}}{2}$ and
– closer to $\underline{p}$ if p is smaller than the midpoint.

The midpoint itself is equidistant from both endpoints $\underline{p}$ and $\overline{p}$.

Let us therefore select an increasing sequence $N_1 < N_2 < \ldots$, and take:

- $\mathcal{A}(\omega, i) = 1$ if the proportion of values $\omega_i = 1$ between N_i and N_{i+1} is greater than or equal to the midpoint $\widetilde{p}$, and
- $\mathcal{A}(\omega, i) = 0$ if this proportion is smaller than $\widetilde{p}$.

Let us prove that, for an appropriate sequence N_i, this algorithm indeed solves almost all instances of the given problem $\mathcal{P}$.

The proposition states that for very $\varepsilon > 0$, $\delta > 0$, and n, there exists an integer $N \geq n$ for which the above inequality holds. To prove the existence of such an N, let us consider the set T of all sequences p for which, for all $N \geq n$, this inequality *does not* hold. We will show that this set T is definable, closed, and nowhere dense. By definition of a sequence that has no other properties (Definition 3), this would imply that the actual sequence p *does not* belong to this set T – and thus, there exists the desired value N, which is exactly what the proposition claims.

Definability is easy: we just had defined this set. Closeness is also rather easy to prove; it can be proven similarly to a similar closeness proof in [5].

The non-trivial part is nowhere density. To prove that the set T is nowhere dense, it is sufficient, for each finite starting sequence $p_1, \ldots, p_n$, to produce an infinite extension p for which the desired integer $N \geq n$ exists (and which, thus, does not belong to the set T).

We will take a sequence p_i all whose elements are either equal to the lower endpoint $\underline{p}$ or to the upper endpoint $\overline{p}$. Specifically, for all the values between N_i and N_{i+1}, we will take:

- $p_i = \overline{p}$ if $s_{\mathcal{P},i} = 1$,
- $p_i = \underline{p}$ if $s_{\mathcal{P},i} = 0$, and
- any of these two values if $i \notin S_{\mathcal{P}}$.

Let us show that for an appropriate choice of the sequence N_i, with probability $\geq 1 - \delta$, for all the values i from n to $N = \dfrac{n}{\varepsilon}$, we will have $\mathcal{A}(\omega, i) = s_{\mathcal{P},i}$. This will imply that the proportion of such i is indeed greater than $1 - \varepsilon$ with probability $\geq 1 - \delta$.

For each i, we consider the arithmetic average of $k_i \stackrel{\text{def}}{=} (N_{i+1} - 1) - N_i$ independent 0-1 random values each of which is equal to 1 with some probability p (namely, either with probability $\underline{p}$ or with probability $\overline{p}$). It is known that this arithmetic average is, in the limit $k_i \to \infty$, normally distributed – this fact is a particular case of the Central Limit Theorem. The mean value of this average is equal to the corresponding probability p, and the standard deviation decreases, with k_i, as $\dfrac{1}{\sqrt{k_i}}$. Let us use these facts to estimate the probability that with $\underline{p} = 0$ we will have $\mathcal{A}(\omega, i) = 1$ or vice versa. In other words, we are interested

in the probability that the average differs from its expected values by at least the half-width $w \stackrel{\text{def}}{=} \dfrac{\overline{p} - \underline{p}}{2}$. For a normal distribution with mean μ and standard deviation σ, asymptotically, this probability is proportional to $\exp\left(-\dfrac{w^2}{2\sigma^2}\right)$, i.e., to $\exp(-\text{const} \cdot k_i)$. If we select k_i in such a way that $\exp(-\text{const} \cdot k_i) \le \dfrac{1}{i^2}$, i.e., $k_i = \text{const} \cdot \ln(i)$, then the probability that this happens for one of the values i cannot exceed the sum of the probabilities corresponding to different i, and is, thus, smaller than the sum $\sum\limits_{i=n}^{\infty} \dfrac{1}{i^2}$. Thus, the sum tends to 0 and is, therefore, smaller than δ for all sufficiently large n.

So, we get the desired property if we find N_i for which $k_i \approx N_{i+1} - N_i \sim \text{const} \cdot \ln(i)$. This approximate equality is true if we take $N_i = i \cdot \ln(i)$.

For this choice of N_i, computing $\mathcal{A}(\omega, i)$ requires $N_{i+1} - N_i \sim \ln(i)$ calls to the oracle – a number which is a linear function of the bit length of an integer i. Thus, this algorithm is indeed feasible. The proposition is proven.

Acknowledgments

This work was supported in part by the National Science Foundation grants HRD-0734825, HRD-1242122, and DUE-0926721. The authors are thankful to all the participants of the 16th International Symposium on Scientific Computing, Computer Arithmetic, and Validated Numerics SCAN'2014 (Würzburg, German, September 21–26, 2014) for valuable discussions, and to the anonymous referees for valuable suggestions.

References

1. Cheu, D., Longpré, L.: Towards the possibility of objective interval uncertainty in physics, Reliable Computing 15(1), 43–49 (2011)
2. Feynman, R., Leighton, R., Sands, M.: The Feynman Lectures on Physics, Addison Wesley, Boston, Massachusetts (2005)
3. Gorban, I.I.: Theory of Hyper-Random Phenomena, Ukrainian National Academy of Sciences Publ., Kyiv (2007) in Russian
4. Gorban, I.I.: Hyper-random phenomena: definition and description, Information Theories and Applications 15(3), 203–211 (2008)
5. Kosheleva, O., Zakharevich, M., Kreinovich, V.: If many physicists are right and no physical theory is perfect, then by using physical observations, we can feasibly solve almost all instances of each NP-complete problem, Mathematical Structures and Modeling 31, 4–17 (2014)
6. Kreinovich, V., Longpré, L.: Pure quantum states are fundamental, mixtures (composite states) are mathematical constructions: an argument using algorithmic information theory, International Journal on Theoretical Physics 36(1), 167–176 (1997)
7. Li, M., Vitányi, P.: An Introduction to Kolmogorov Complexity and Its Applications, Springer, Berlin-Heidelberg, (2008)

8. Longpré, L., Kreinovich, V.: When are two wave functions distinguishable: a new answer to Pauli's question, with potential application to quantum cosmology, International Journal of Theoretical Physics 47(3), 814–831 (2008)
9. Oxtoby, J.C.: Measure and Category: A Survey of the Analogies between Topological and Measure Spaces, Springer Verlag, New York, Heidelberg, Berlin (1980)
10. Papadimitriou, C.: Computational Complexity, Addison Welsey, Reading, Massachusetts (1994)
11. Wester, M.E., et al.: Atomic weights of the elements 2011 (IUPAC Technical Report), Pure and Applied Chemistry 85(5), 1047–1078 (2013)