

A SYSTEMATIC WAY OF ANALYSING PROOFS IN PROBABILITY THEORY

MORENIKEJI NERI^a, PAULO OLIVA^b, NICHOLAS PISCHKE^c

^a Department of Mathematics, Technische Universität Darmstadt,
Schlossgartenstraße 7, 64289 Darmstadt, Germany,

^b School of Electronic Engineering and Computer Science, Queen Mary University of London,
Mile End Road, London, E1 4NS, United Kingdom,

^c Department of Computer Science, University of Bath,
Claverton Down, Bath, BA2 7AY, United Kingdom,

E-mails: neri@mathematik.tu-darmstadt.de, p.oliva@qmul.ac.uk, nmp39@bath.ac.uk

ABSTRACT. Over extended systems of finite type arithmetic, we utilize a formal representation of the outer measure to define a translation which allows for the systematic formalization of probabilistic statements. As a main result, this translation gives rise to novel *probabilistic logical metatheorems* in the style of proof mining, guaranteeing the extractability of computable bounds from (non-effective) proofs of probabilistic existence statements. We further show how the set-theoretically false principle of uniform boundedness due to Kohlenbach can be used to replicate logically strong continuity properties of probability measures in the context of these bound extraction theorems in a tame way, i.e. without affecting the computational complexity of the resulting bounds in question, all the while guaranteeing the validity of those bounds even over finitely additive probability spaces. This in particular provides a formal perspective on the elimination of the principle of σ -additivity during bound extraction, as previously only observed ad hoc in the practice of proof mining. In that context, we for the first time provide a proof-theoretic treatment of higher-type uniform boundedness principles and related contraction principles via Kohlenbach's monotone variant of Gödel's functional interpretation, which is of independent interest. All together, these new metatheorems provide a systematic proof-theoretic approach towards extracting various types of quantitative information for probabilistic theorems considered in the literature, justifying a range of recent applications to probability theory and stochastic optimization. This paper represents a major logical contribution to a recent advance of bringing the methods of proof mining to bear on probability theory, significantly extending previous work by the first and third author [Forum Math. Sigma, 13, e187 (2025)] in that direction.

Keywords: Proof mining; Metatheorems; Probability theory; Outer measure; Finitely additive measures.

MSC2020 Classification: 03F10, 03F35, 28A12, 60A10

1. INTRODUCTION

1.1. Motivation and background. The nature and description of the computational content of a mathematical theorem is one of the fundamental questions in proof theory, and the development of tools and frameworks to exhibit and categorize this content is a key motivation behind many of the modern developments in the field. In the present paper, we are concerned with the computational content of theorems that are of a probabilistic nature, i.e. which, instead of making absolute statements, make *almost sure* statements.

Indeed, many theorems from probability theory and statistics, including their most striking results, often take the form of almost sure theorems, asserting that a given property holds with probability one. Examples range from convergence results for sampling-based algorithms in

Date: April 9, 2026.

statistics and operations research to zero-one laws, which state that a property can only hold almost always or almost never, with nothing in between. Beyond their qualitative content, such theorems, each in their own right, can clearly be made to offer a related computational challenge. This challenge commonly takes the form of asking whether they can be strengthened quantitatively, showing that an approximate version of the property in question holds up to a chosen degree of probability. This transforms what might seem like a computationally empty statement, that is, the property in question holding with probability one, into a statement with a meaningful computational interpretation, such that the complexity of the related computational information yields a sensible measure of the internal complexity of those statements and, in particular, their proofs.

To illustrate the above with a more precise example, we turn to almost sure convergence results, perhaps the most central and prevalent of such examples due to their ubiquity in statistics and stochastic approximation. Concretely, if $(X_n)_{n \in \mathbb{N}}$ is a stochastic process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, one can show that the probabilistic statement

$$\text{“}(X_n)_{n \in \mathbb{N}} \text{ converges with probability one”}$$

is equivalent to (see e.g. [50]) the property

$$\forall \varepsilon, \lambda > 0 \exists N \in \mathbb{N} (\mathbb{P}(\forall i, j \geq N (|X_i - X_j| \leq \varepsilon)) > 1 - \lambda).$$

A function $\Phi(\varepsilon, \lambda)$ providing a bound (and hence a witness) for the N in the above expression is exactly what is isolated in the probability-theoretic literature as a *rate of almost sure convergence*. From a logical point of view, analogous to rates of convergence for deterministic sequences, such rates are generally only effectively derivable in special circumstances, such as (but not limited to) semi-constructive contexts (see e.g. the discussion in [31]). Correspondingly, there are further equivalent phrasings of almost sure convergence which give rise to quantitatively weaker readings, such as *metastable rates of almost sure convergence* introduced by Avigad, Dean, and Rute [2] (extending the work of Tao [60]), that is, functions $\Phi(\varepsilon, \lambda, g)$ bounding the N in the statement

$$\forall \varepsilon, \lambda > 0 \forall g : \mathbb{N} \rightarrow \mathbb{N} \exists N \in \mathbb{N} (\mathbb{P}(\exists n \leq N \forall i, j \in [n; n + g(n)] (|X_i - X_j| \leq \varepsilon)) > 1 - \lambda).$$

These can be seen as probabilistic generalizations of ordinary rates of metastability in the sense of Tao [61], which over the years have become a central way to assign a computational meaning to deterministic convergence statements in the absence of computable rates of convergence. Indeed, contrary to the above rates of almost sure convergence, these metastable rates (and also more uniform variants as discussed in detail later) are widely available for large classes of classically proven almost sure convergence theorems (as will, in fact, be formally justified for the first time in this paper) and they have been, and continue to be, obtained for major theorems such as the dominated convergence theorem [2], the martingale convergence theorem [50] and the Robbins-Siegmund theorem [51], as well as further applied results from stochastic approximation theory [49], among others.

Such computational considerations are certainly not isolated to convergence theorems, with further examples being discussed later on. The fundamental question of the present paper is then how such probabilistic statements can be formally recognized in the context of logical systems that facilitate the extraction of computational content from proofs, and in particular how the extractability of computational information commonly associated with such statements, as illustrated above, can be guaranteed therein.

1.2. The proof mining program and probability theory. The approach of this paper is based on the formal systems and methods employed in the proof mining program. This research

program originates with the work of Kohlenbach in the 1990s, systematizing and expanding the ideas of Kreisel’s unwinding of proofs [36, 37], and aims at providing the computational content of theorems from mathematics by analysing their (prima facie) non-effective proofs as they are found in the usual literature. While applied in nature, this enterprise is firmly logically supported and substantiated through a range of proof-theoretic results called *logical metatheorems*, mainly based on central proof-theoretic machinery like Gödel’s functional interpretation [15], Kreisel’s modified realizability interpretation [38], as well as Howard’s majorizability [20], and their variants. Concretely, the logical metatheorems used in proof mining are theorems associated with a logical system at hand such that

- (1) the system is suitably designed so that it allows for the formalization of large classes of objects and proofs from the respective area of application,
- (2) the associated logical metatheorem then guarantees that from large classes of proofs carried out in this system, potentially involving a wide range of non-computational principles, one can extract effective and highly uniform computational information for the respective theorem.

Substantiated and supported by these metatheorems, proof mining has now led to hundreds of new applications in core mathematics. We refer to the central monograph [31] for a comprehensive overview of the field, as well as to the surveys [33, 34] for further references and discussions.

This logical machinery was recently extended to probability theory by the first and third author in [46], where corresponding systems and logical metatheorems on bound extraction were presented. While discussed in more detail later on, we already want to highlight here that these systems follow an abstract approach to both the sample and event space, treating them in particular as bounded spaces (we refer to [46] for further discussions on that point, which is of fundamental importance to this enterprise), and that crucially, the main system places strong emphasis on so-called probability contents, that is finitely additive probability measures (see in particular [56]), as the right underlying notion for extending proof mining methods to this area. Concretely, probability contents are functions $\mathbb{P} : S \rightarrow [0, 1]$ on an algebra of sets S over a base set Ω (that is $\emptyset \in S$ and S is closed under complements and finite unions) which satisfy $\mathbb{P}(\Omega) = 1$ and are additive, i.e. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ for disjoint $A, B \in S$. An algebra S becomes a σ -algebra if it is closed under countable unions and over such a σ -algebra, \mathbb{P} becomes a probability measure if it is additionally σ -additive, i.e. $\mathbb{P}(\bigcup_{i \in \mathbb{N}} A_i) = \sum_{i \in \mathbb{N}} \mathbb{P}(A_i)$ for a disjoint collection $(A_i)_{i \in \mathbb{N}} \subseteq S$.

In parallel to and since the work [46], various case studies of proof mining in probability theory have been carried out, which include, beyond those related to probabilistic convergence theorems mentioned before, also work on strong laws of large numbers [42, 43, 44] and ergodic theory [45], on stochastic optimization [48, 52, 53, 54] and on zero-one laws [55].

1.3. The contributions of this paper. The main system introduced in [46] “only” axiomatises probability contents. It is already argued therein that this restriction does not appear to be a real limitation on the applicability of the system, as it seemed sufficient for essentially all applications of that approach developed at the time of its release. These were however quite limited, and further developments (in particular the case studies mentioned above) quickly showed that the applicability of this system is a subtle issue, not least through the system’s restricted vocabulary, which required substantial additional work tailored ad hoc to each such case study. The fundamental question thereby arose of how theorems and their proofs, coming from ordinary probability theory, can be systematically recognized in the language of that system to render these metatheorems applicable and to hence guarantee the extractability of

related quantitative information. Or, in other words: How can the property (1) of logical metatheorems highlighted above be justified for the systems presented in [46]?

In this paper, we provide an answer to this question by presenting a systematic approach to the general task of assigning a computational meaning, both classically and constructively, to probabilistic statements which covers the previously discussed examples, as well as many more, in a uniform manner, and moreover facilitates a similarly systematic approach for the extraction of that content. In particular, the results of this paper substantiate the previously mentioned case studies released since as applications of this general logical approach initiated in [46]. Hence, while first proposed [46], it is only through the present work that we can now really recognize the systematic applicability and suitability of this approach via the tame theory of probability contents.

More concretely, the present paper proceeds as follows: In the language of those systems from [46], we utilize a formal representation of the outer measure to define a translation which allows for the systematic formalization of probabilistic statements. In its most special case, this translation associates to any formula φ another formula $\mathbb{P}[\varphi] = 1$ which expresses that φ holds almost surely. This translation can then be combined with the various proof-theoretic tools employed previously to establish the logical metatheorems for the systems for contents in [46] to associate with that derived statement $\mathbb{P}[\varphi] = 1$ a precise computational meaning.

In particular, this gives rise to dedicated new *probabilistic* logical metatheorems for the extraction of computational information from proofs of probabilistic theorems, extending the *basic* logical metatheorems previously derived in [46]. In similarity to the bound extraction theorems from [46], also these new metatheorems presented here guarantee a high degree of uniformity of the extractable bounds, in the sense that they will be independent of all parameters relating to the underlying probability space, and in particular of the measure. Establishing some of these new metatheorems is highly non-trivial, and makes crucial use of essentially all known techniques and devices from the proof mining program, in particular of extensions of a set-theoretically false principle of uniform boundedness introduced by Kohlenbach [30] as further discussed below. Throughout, we also simultaneously focus both on metatheorems tailored to classical as well as constructive reasoning, the latter of which are considered in the context of probability theory here for the first time and are derived by following the work by Gerhardy and Kohlenbach [13]. We will then in particular discuss how these probabilistic metatheorems give rise to various well-known quantitative notions from probability theory, covering the examples and case studies mentioned before, and thereby illustrating that the new approach of this paper is in particular practically natural and useful.

All of these considerations are, in particular, made over the systems for probability contents which guarantees that the resulting extracted bounds will be valid in that context. This has particular consequences in the light of our second set of contributions, where we provide an algebra for expressions of the form $\mathbb{P}[\varphi] = 1$ based on the logical structure, and in particular the quantifier structure, of φ . In that context, under the use of a principle of uniform boundedness due to Kohlenbach [28, 30], we show that the quantifiers internal to φ can be, in some cases, “pulled out” of the expression $\mathbb{P}[\varphi] = 1$. These prenexations of probabilistic statements are intimately related to continuity principles of the set-function, and in general only valid in the context of probability measures. However, as the aforementioned uniform boundedness principle is admissible in the context of the usual logical metatheorems employed in proof mining (such as those developed in this paper), even though it is generally set-theoretically false, these prenexation results may be used in proofs, while still guaranteeing that the resulting bounds are true for general probability contents. In that way, we illustrate how uniform boundedness

principles can be employed to use facts about probability measures in proofs for probability contents, highlighting great potential importance of such principles for the future of proof mining and probability. For that reason, all the logical metatheorems discussed in the present paper accommodate such uniform boundedness principles explicitly. In that context, we in particular for the first time provide a proof-theoretic treatment of these uniform boundedness principles in higher types, which allow for such prenexation principles even for uncountable quantifiers, which in that way are false even for probability measures. To accommodate such results also in a semi-constructive context, we not only rely on the principle of uniform boundedness, but also on the dual so-called “contra-collection” principle, which prominently features in the context of the bounded functional interpretation of the second author and Ferreira [12], and is treated here for the first time in the context of the monotone functional interpretation. These results, extending the reach of the monotone functional interpretation, are in particular of independent interest even outside applications to probability theory. Moreover, as hinted on before, we want to emphasize that these principles feature crucially in our approach to some of the new probabilistic metatheorems. Lastly, we provide a type of conservativity result for the systems considered here regarding the principle of σ -additivity. Concretely, we show that in a classical context, σ -additivity can actually be derived from the uniform boundedness principles, which provides a formal perspective on the elimination of σ -additivity during bound extraction in proof mining, as was previously only observed ad hoc in practice. All these results in particular further affirm the suitability of probability contents, not measures, as the fundamental measure-theoretic notion in proof mining and probability theory, as mentioned before.

1.4. Related and future work. Beyond previously mentioned contributions, we of course want to highlight that present work will allow for the development of many further applications of proof mining to probability theory. Indeed, some of the previously mentioned case studies (see in particular [44, 45, 48, 49, 52, 53]) were developed in tandem with the present work, where the logical perspectives developed here were crucial in guiding the formulation of quantitative variants of probabilistic statements and the extraction of the respective computational content. In the context of [49] in particular, the present logical approach was crucial for developing the associated notion of a finitary martingale, as well as the associated quantitative theory. Further applications stemming from the present work, in particular to stochastic optimization, are already underway. Aside from those concrete applications, we want to mention two avenues for future work stemming from this paper.

At first, we want to highlight that the boundedness of the sample and event space as well as their associated uniform boundedness and contra-collection principles play a crucial role in this work. While not discussed here in any great detail, since we generally follow the formal setup of [46], this highlights the very interesting question for future work of whether the present topic can benefit from the use of the *bounded functional interpretation* due to Ferreira and the second author [12] (see also [10, 11]), a proof interpretation inspired by the paradigm of the monotone functional interpretation of Kohlenbach [25] (and conceptually already going back to [24]) but where such aspects are dealt with more fundamentally, in particular the contra-collection principle.

Similarly, while also not discussed here in any further detail, we want to mention that while the present results are formulated as tailored to the proof-theoretic framework from [46], various ideas of the present paper could perhaps also be used in the context of model-theoretic approaches to uniform bounds, in particular as, e.g., articulated in the context of positive bounded logic [5, 8] (see also [46] for a more detailed discussion on the differences and similarities of the present proof-theoretic methodology to such results from model theory). In particular, the

general translation induced by the outer measure, mapping formulas to probabilistic formulas, could be used in such a context to broaden the applicability of the related model-theoretic results on the existence of uniform bounds, as it can be phrased in the language of positive bounded logic. In any case however, we want to highlight that these model-theoretic results can only guarantee existence of such uniform bounds and, in difference to the proof-theoretic approach, do not offer considerations on complexity or an approach to extracting them from proofs.

Further, it would be very interesting to compare the present use of the outer measure to similar such uses in model theory (see e.g. [22]), and especially to probability quantifiers (see in particular [21], among many others) or probabilistic logic (see in particular [58]). However, we want to emphasize that the goal of the present paper is quite distinct from these and related works, and is particularly rooted in the perspective of (applied) proof theory, where such a notion is, to the best of our knowledge, employed for the first time in the present paper.

Lastly, it would be interesting to consider the connection between uniform boundedness principles and continuity theorems for probability measures, as developed in this paper, also from a model-theoretic perspective, in particular in light of the observation that uniform boundedness represents a sort-of proof-theoretic version of saturation for ultraproducts in model theory, as e.g. centrally discussed in [17] where many such applications of saturation are shown to be captured from a proof-theoretic perspective via uniform boundedness. In particular, it would be interesting to investigate the relationship between the present uses of uniform boundedness in probability theory and the construction of the Loeb measure in nonstandard analysis (see the seminal work [41] as well as [8]), where saturation is similarly employed to extend finitely additive measures to σ -additive ones, and hence used to establish continuity properties. Investigations on this last aspect are already underway.

1.5. The outline of the paper. The paper is now organised as follows: Section 2 provides the necessary details on the systems from [46] as well as the central proof-theoretic devices used in this paper and the associated *basic* logical metatheorems, in particular their extensions to incorporate higher-type uniform boundedness and contra-collection. Section 3 motivates and introduces our general translation based on a formal representation of the outer measure in the language of the systems from [46]. Section 4 provides the various results on the algebra of our translation, including the prenexation results based on the principles of uniform boundedness and contra-collection. Moreover, in that section we highlight in particular how the uniform boundedness principles provide a formal perspective on the elimination of σ -additivity from proofs in the course of a proof-theoretic analysis. Section 5 then provides the general discussion of the computational content of probabilistic statements as suggested by our translation, resulting in metatheorems specially tailored to these probabilistic circumstances. The final Section 6 discusses how these respective probabilistic logical metatheorems proved earlier give rise to various natural quantitative notions from the literature, and beyond, relating them in particular to the previously mentioned case studies.

2. THE FORMAL FRAMEWORK

The present section serves effectively as preliminaries to the present paper, laying out the essential details of the formal approach towards proof mining and probability theory developed in [46], but also develops crucial new material like further variations arising by considering semi-constructive versions and extensions via uniform boundedness and contra-collection principles, as required here. While we will only rely on rather little probability theory, which we will

introduce as needed, we refer to [23] for a general reference on probability theory and to [56] for the central reference for the theory of contents.

2.1. Systems for probability contents. We begin with the essential details of the formal approach towards a system geared for bound extractions from proofs in the theory of probability contents, as introduced in [46]. The main system defined for that purpose in [46] is the system $\mathcal{F}^\omega[\mathbb{P}]$ which, as common for logical systems employed in the proof mining program, arises as an extension of systems of finite type arithmetic.

Concretely, as in [31] and [62], let WE-PA^ω be a weakly extensional variant of Peano arithmetic over all *finite types* T defined by

$$0 \in T \quad \text{and} \quad \text{if } \rho, \tau \in T \text{ then } \rho(\tau) \in T.$$

We denote the *pure types* using natural numbers, by recursively setting $n + 1 := 0(n)$. The only primitive relation symbol of WE-PA^ω is $=_0$, representing equality at type 0. In particular, equality at higher types is defined via the abbreviation

$$x^{\tau(\xi)} =_{\tau(\xi)} y^{\tau(\xi)} \quad := \quad \forall z^\xi (xz =_\tau yz)$$

and for higher type inequalities, we similarly set $x^{\tau(\xi)} \geq_{\tau(\xi)} y^{\tau(\xi)} := \forall z^\xi (xz \geq_\tau yz)$.¹ Besides equality at type 0, the system contains constants for successor and zero, constants for the combinators of Schönfinkel [57] allowing for lambda abstraction, and lastly, constants allowing for simultaneous primitive recursion in the sense of Gödel [15] and Hilbert [19], all with their corresponding natural defining axioms. We do not spell out these constants or axioms as they do not feature explicitly in this work. Instead, we refer the reader to [31].

Furthermore, as opposed to the full axiom of extensionality, the system crucially only contains the following *quantifier-free extensionality rule*

$$\text{(QF-ER)} \quad \frac{\varphi_0 \rightarrow s =_\rho t}{\varphi_0 \rightarrow r[s/x^\rho] =_\tau r[t/x^\rho]}$$

where φ_0 is a quantifier-free formula, s and t are terms of type ρ and r is a term of type τ . We refer the reader to [31] for further discussions on the (necessary) restrictions on extensionally in systems amenable to program extraction.

Over WE-PA^ω and its extensions, we rely on a representation of real numbers as fast converging Cauchy sequences of rational numbers with a fixed Cauchy modulus 2^{-n} (see Chapter 4 in [31] for details), that is as objects of type $1 = 0(0)$. We will not need any precise details of this representation and just recall that in that context, the usual arithmetical operations like $+\mathbb{R}$, $\cdot\mathbb{R}$, $|\cdot|_{\mathbb{R}}$, etc., are primitive recursively definable through closed terms and the relations $=_{\mathbb{R}}$ and $<_{\mathbb{R}}$, etc., are representable via formulas of the underlying language. Furthermore, these relations are not decidable but are given by Π_1^0 - and Σ_1^0 -formulas, respectively. We generally drop subscripts from arithmetical operations and often also reduce typing in the notation throughout for readability.

Instead of relying on representations of probability content spaces in finite type arithmetic, which would come with all sorts of additional questions and complications, the central approach used in [46] is to represent the sample and event space using new, additional abstract types, following the central modern paradigm of proof mining developed by Kohlenbach, starting with the seminal work [29]. Concretely, consider now the extended set of finite types $T^{\Omega, S}$ by

$$0, \Omega, S \in T^{\Omega, S} \quad \text{and} \quad \text{if } \rho, \tau \in T^{\Omega, S} \text{ then } \rho(\tau) \in T^{\Omega, S},$$

¹Note that \geq_0 is definable from the only primitive relation $=_0$ using cut-off subtraction.

with the types Ω and S serving as abstract representations of the sample and event space, respectively. The language of the system $\mathcal{F}^\omega[\mathbb{P}]$ for a probability content over an algebra of events now arises from WE-PA $^\omega$ (lifted to this extended set of types) by adding the constants eq with type $0(\Omega)(\Omega)$ and \in with type $0(S)(\Omega)$ for the sample space, \cup with type $S(S)(S)$, $(\cdot)^c$ with type $S(S)$ and \emptyset with type S for the event space, and lastly \mathbb{P} of type $1(S)$ for the probability content.² The axioms of $\mathcal{F}^\omega[\mathbb{P}]$ then specify that

$$x^\Omega =_\Omega y^\Omega := \text{eq}xy =_0 0$$

is an equivalence relation representing equality on Ω , that the abstract element relation \in behaves as expected relative to the set-theoretic operations \cup and $(\cdot)^c$ for union and complement as well as the constant \emptyset for the empty set and lastly that \mathbb{P} is probability content on the algebra represented by the type S .³ As in [46], we introduce appropriate abbreviations to make formal expressions less rigid: We typically write A^c instead of $(A)^c$, $x \in A$ instead of $\in xA =_0 0$ and $x \notin A$ instead of $\in xA \neq_0 0$. Also, we define $\Omega := \emptyset^c$ and $A \cap B := (A^c \cup B^c)^c$. Arbitrary finite unions and intersections are then definable from these operators via recursion, and we refer to [46] for further discussions. Equality on S as well as inclusion are also derived notions which we obtain by setting

$$A =_S B := \forall x^\Omega (x \in A \leftrightarrow x \in B) \quad \text{and} \quad A \subseteq_S B := \forall x^\Omega (x \in A \rightarrow x \in B).$$

Then, $=_S$ is provably an equivalence relation and \subseteq_S is provably a partial order w.r.t. that equality. For the sole purpose of extending the inequality relation to all types, we define $x^\Omega \geq_\Omega y^\Omega := \mathbb{P}(\Omega) \geq_{\mathbb{R}} \mathbb{P}(\Omega)$ and $A^S \geq_S B^S := \mathbb{P}(A) \geq_{\mathbb{R}} \mathbb{P}(B)$, following [46].

The last addition to $\mathcal{F}^\omega[\mathbb{P}]$ are two choice principles, the principle of quantifier-free choice QF-AC as well as the principle of dependent choice DC (see e.g. Chapter 11 in [31]), the latter of those giving access to full classical analysis.

Remark 2.1. In similarity to [46], we remark that while the main system $\mathcal{F}^\omega[\mathbb{P}]$, through DC, contains a rather large amount of comprehension, this is just to highlight the potential strength of systems which are amenable to the present methods. All the while, as with [46], no logical considerations made in this paper in any way depend on the presence of that comprehension, nor do they entail principles of such strength, and hence can be just as well developed over a range of weak fragments, such as those introduced in [26] based on the Grzegorzczuk hierarchy [16].

Remark 2.2. The main system $\mathcal{F}^\omega[\mathbb{P}]$ allows for various mathematically natural and proof-theoretically tame extensions to deal with various objects and notions from mathematical practice, e.g. random variables and their Lebesgue integrals as discussed in [46]. These extensions then, in particular, potentially populate the term structure of the type S with more non-trivial ground terms. While we do not discuss these extensions here, and just phrase everything for

²The system $\mathcal{F}^\omega[\mathbb{P}]$ as defined in [46] also contains a constant c_Ω of type Ω , witnessing the non-emptiness of Ω , and which in [46] serves a rather technical purpose in the proof of the associated logical metatheorems that plays no further role in this paper. We hence omit any further discussion thereof.

³While most axioms of probability contents are easily defined via computationally empty universal axioms, recognizing that the law of disjoint unions can be specified in a way that is admissible in the context of bound extraction theorems is not immediately straightforward and requires considerations on the uniform boundedness of the new abstract types. Further, the system $\mathcal{F}^\omega[\mathbb{P}]$ as defined in [46] actually replaces the law of disjoint unions with the inclusion exclusion principle together with the monotonicity of the content for reasons of practicality in concrete formalizations. In fact, we want to highlight a necessary correction to [46] which misses the crucial (universal, and hence unproblematic) axiom that $\mathbb{P}(\Omega) =_{\mathbb{R}} 1$. While the precise axiomatization of \mathbb{P} therefore carries various subtleties with it, these do not feature in the present paper any further and so we simply refer the interested reader to the discussion in [46].

the simplest system $\mathcal{F}^\omega[\mathbb{P}]$, it should be noted that essentially all considerations of this paper extend mutatis mutandis to these broader settings.

As we also consider semi-constructive variants of our results throughout, we will also rely on an intuitionistic variant of $\mathcal{F}^\omega[\mathbb{P}]$. This variant, denoted by $\mathcal{F}_i^\omega[\mathbb{P}]$, will be based on WE-HA^ω , that is WE-PA^ω with the law of excluded middle removed, and additionally includes the principles the full axiom of choice AC , Markov's principle M^ω and the independence of premise principle for universal formulas IP_\forall^ω (see e.g. Chapters 5 and 8 in [31]) instead of QF-AC and DC . Aside from these, $\mathcal{F}_i^\omega[\mathbb{P}]$ arises from WE-HA^ω in exactly the same way as $\mathcal{F}^\omega[\mathbb{P}]$ arises from WE-PA^ω , that is with the same additional types, constants and axioms.

2.2. Uniform boundedness, contra-collection and metatheorems on program extraction. We now discuss the main results for the systems introduced previously, that is the logical metatheorems on bound extraction. While the basic versions of the classical variants of these results were established in [46], the semi-constructive variants are for the first time stated in this paper. These, however, arise through a rather immediate combination of the constructions from [46] with the considerations on semi-constructive systems given by Gerhardy and Kohlenbach in [13], so that we are content with proof sketches for the resulting metatheorems tailored to probability theory.

Slightly more extensive modifications are required to incorporate Kohlenbach's uniform boundedness principle, as considered in [30] (see also already the earlier works [25, 27, 28] for variants of this principle not phrased in the context of abstract types). While our arguments here follow the previous works [17, 27, 30], we actually consider this principle here for the first time in types higher than the base type 0 in the context of the monotone functional interpretation. This requires some care in adapting the previous arguments, and we therefore provide additional details.

The most extensive modifications arise through our treatment of the so-called *principle of contra-collection*, an abstract principle related to weak König's lemma (arising, classically, as the contraposition of the associated uniform boundedness principle, itself an abstract extension of the FAN principle from intuitionism), isolated in this extended form for the first time in the context of the bounded functional interpretation [12]. Also this principle is treated here for the first time in the context of the monotone functional interpretation (outside of classical contexts), and although we only consider its variant for the base type 0, together with some extensions to type 1, this principle requires various non-trivial arguments which we provide in full detail.

The first key logical tool used in the context of the logical metatheorems, and in the rest of the paper for that matter, is Gödel's functional (Dialectica) interpretation.

Definition 2.3 ([15]). The Dialectica interpretation $\varphi^D = \exists \underline{x} \forall \underline{y} \varphi_D(\underline{x}, \underline{y})$ of a formula φ in the language of $\mathcal{F}^\omega[\mathbb{P}]$ (or its variants) is defined via the following recursion on the structure of the formula:

- (1) $\varphi^D := \varphi_D := \varphi$ for φ being a prime formula.

If $\varphi^D = \exists \underline{x} \forall \underline{y} \varphi_D(\underline{x}, \underline{y})$ and $\psi^D = \exists \underline{u} \forall \underline{v} \psi_D(\underline{u}, \underline{v})$, we set

- (2) $(\varphi \wedge \psi)^D := \exists \underline{x}, \underline{u} \forall \underline{y}, \underline{v} (\varphi \wedge \psi)_D$

where $(\varphi \wedge \psi)_D(\underline{x}, \underline{u}, \underline{y}, \underline{v}) := \varphi_D(\underline{x}, \underline{y}) \wedge \psi_D(\underline{u}, \underline{v})$,

- (3) $(\varphi \vee \psi)^D := \exists z^0, \underline{x}, \underline{u} \forall \underline{y}, \underline{v} (\varphi \vee \psi)_D$

where $(\varphi \vee \psi)_D(z^0, \underline{x}, \underline{u}, \underline{y}, \underline{v}) := (z = 0 \rightarrow \varphi_D(\underline{x}, \underline{y})) \wedge (z \neq 0 \rightarrow \psi_D(\underline{u}, \underline{v}))$,

- (4) $(\varphi \rightarrow \psi)^D := \exists \underline{U}, \underline{Y} \forall \underline{x}, \underline{v} (\varphi \rightarrow \psi)_D$

where $(\varphi \rightarrow \psi)_D(\underline{U}, \underline{Y}, \underline{x}, \underline{v}) := \varphi_D(\underline{x}, \underline{Y} \underline{x} \underline{v}) \rightarrow \psi_D(\underline{U} \underline{x}, \underline{v})$,

- (5) $(\exists z^\tau \varphi(z))^D := \exists z, \underline{x} \forall \underline{y} (\exists z^\tau \varphi(z))_D$
 where $(\exists z^\tau \varphi(z))_D(z, \underline{x}, \underline{y}) := \varphi_D(\underline{x}, \underline{y}, z)$,
- (6) $(\forall z^\tau \varphi(z))^D := \exists \underline{X} \forall z, \underline{y} (\forall z^\tau \varphi(z))_D$
 where $(\forall z^\tau \varphi(z))_D(\underline{X}, z, \underline{y}) := \varphi_D(\underline{X}z, \underline{y}, z)$.

Beyond this, in the classical context we will also rely on a negative translation to transfer classical to semi-constructive proofs. As in previous works, we consider the following rather minimal variant due to Kuroda.

Definition 2.4 ([40]). The (Kuroda) negative translation φ' of φ is defined as $\varphi' := \neg\neg\varphi^*$, with φ^* defined recursively via

- (1) $\varphi^* := \varphi$, if φ is atomic,
- (2) $(\varphi \circ \psi)^* := \varphi^* \circ \psi^*$, for $\circ \in \{\wedge, \vee, \rightarrow\}$,
- (3) $(\exists x^\tau \varphi)^* := \exists x^\tau \varphi^*$,
- (4) $(\forall x^\tau \varphi)^* := \forall x^\tau \neg\neg\varphi^*$.

The last tool that will feature in the present paper is Kreisel's modified realizability interpretation, which can be employed in semi-constructive contexts *without* Markov's principle.

Definition 2.5 ([38, 39]). For any formula φ in the language of $\mathcal{F}^\omega[\mathbb{P}]$ (or its variants), we define its modified realizability interpretation $\underline{x} \text{ mr } \varphi$ by recursion on the structure of φ :

- (1) $\langle \rangle \text{ mr } \varphi := \varphi$ for a prime formula φ where $\langle \rangle$ is the empty tuple.

Further, if $\underline{x} \text{ mr } \varphi$ and $\underline{y} \text{ mr } \psi$ are the modified realizability interpretations of φ and ψ , respectively, then:

- (2) $\underline{x}, \underline{y} \text{ mr } (\varphi \wedge \psi) := \underline{x} \text{ mr } \varphi \wedge \underline{y} \text{ mr } \psi$,
- (3) $z^0, \underline{x}, \underline{y} \text{ mr } (\varphi \vee \psi) := (z =_0 0 \rightarrow \underline{x} \text{ mr } \varphi) \wedge (z \neq_0 0 \rightarrow \underline{y} \text{ mr } \psi)$,
- (4) $\underline{Y} \text{ mr } (\varphi \rightarrow \psi) := \forall \underline{x} (\underline{x} \text{ mr } \varphi \rightarrow \underline{Y} \text{ mr } \psi)$,
- (5) $\underline{X} \text{ mr } \forall w^\rho \varphi(w) := \forall w^\rho (\underline{X}w \text{ mr } \varphi(w))$,
- (6) $z^\rho, \underline{x} \text{ mr } \exists w^\rho \varphi(w) := \underline{x} \text{ mr } \varphi(z)$.

Next to the above proof interpretations, we will also crucially rely on the notion of *majorizability*, originally introduced by Howard [20] and later extended by Bezem [7] to the notion of *strong majorizability*. While majorizability features crucially throughout, it is in particular also the latter notion which allows our classical system $\mathcal{F}^\omega[\mathbb{P}]$ to include the strong dependent choice principle DC, which is interpreted using bar-recursion for which the structure of all strongly majorizable functionals provides a model. Moreover, it generally allows us to include a very strong uniform boundedness principle later (see Definitions 2.8 and 2.10). We here consider an extension of the respective notion of majorizability to the abstract types Ω and S for the use in the theory of probability contents in [46]. In that context, as in the case of the first logical metatheorems with abstract types considered in [14, 29], majorants of functionals with types $\tau \in T^{\Omega, S}$ will have type $\widehat{\tau} \in T$, defined recursively via $\widehat{0} := 0$, $\widehat{\Omega} := 0$, $\widehat{S} := 0$ and $\widehat{\tau(\xi)} := \widehat{\tau}(\widehat{\xi})$. We then introduce our extension of Bezem's notion of strong majorizability semantically as follows, tied to the structure of all strongly majorizable functionals:

Definition 2.6 ([46], following [14, 29] and [7]). Let Ω be a non-empty set, $S \subseteq 2^\Omega$ be an algebra and \mathbb{P} be a probability content on S . The structure $\mathcal{M}^{\omega, \Omega, S}$ and the majorizability

relation \succsim_ρ are defined by

$$\left\{ \begin{array}{l} \mathcal{M}_0 := \mathbb{N}, n \succsim_0 m := n, m \in \mathbb{N} \wedge n \geq m, \\ \mathcal{M}_\Omega := \Omega, n \succsim_\Omega x := n \in \mathbb{N}, x \in \Omega \wedge n \geq \mathbb{P}(\Omega), \\ \mathcal{M}_S := S, n \succsim_S A := n \in \mathbb{N}, A \in S \wedge n \geq \mathbb{P}(A), \\ x^* \succsim_{\tau(\xi)} x := x^* \in \mathcal{M}_{\hat{\tau}}^{\mathcal{M}_{\hat{\xi}}} \wedge x \in \mathcal{M}_{\tau}^{\mathcal{M}_{\xi}} \\ \quad \wedge \forall y^* \in \mathcal{M}_{\hat{\xi}}, y \in \mathcal{M}_{\xi} (y^* \succsim_{\xi} y \rightarrow x^* y^* \succsim_{\tau} xy) \\ \quad \wedge \forall y^*, y \in \mathcal{M}_{\hat{\xi}} (y^* \succsim_{\hat{\xi}} y \rightarrow x^* y^* \succsim_{\hat{\tau}} x^* y), \\ \mathcal{M}_{\tau(\xi)} := \left\{ x \in \mathcal{M}_{\tau}^{\mathcal{M}_{\xi}} \mid \exists x^* \in \mathcal{M}_{\hat{\tau}}^{\mathcal{M}_{\hat{\xi}}} : x^* \succsim_{\tau(\xi)} x \right\}. \end{array} \right.$$

The other main structure featuring in the metatheorems is the structure of all set-theoretic functionals:

Definition 2.7. Let Ω be a non-empty set, $S \subseteq 2^\Omega$ be an algebra and \mathbb{P} be a probability content on S . The structure $\mathcal{S}^{\omega, \Omega, S}$ is defined via $\mathcal{S}_0 := \mathbb{N}$, $\mathcal{S}_\Omega := \Omega$, $\mathcal{S}_S := S$ and $\mathcal{S}_{\tau(\xi)} := \mathcal{S}_{\tau}^{\mathcal{S}_{\xi}}$.

The logical metatheorems, classical and constructive, then arise via a combination of a soundness result for any of the previous proof interpretations, by which one extracts realizing terms with types from $T^{\Omega, S}$ for a given $\forall\exists$ -theorem, together with a subsequent majorization of these realizers, resulting in bounding terms with types from T which are validated in a model based on $\mathcal{M}^{\omega, \Omega, S}$. From this, one then can transfer the results back to $\mathcal{S}^{\omega, \Omega, S}$ if the types occurring in the theorem are “low enough”, formally encapsulated by the notion of small and admissible types as defined in [17] (derived from the same notion introduced in [14, 29] under a different name). This combination of a proof interpretation and majorization effectively amounts to an application of the monotone variants of the interpretations in question, that is of the monotone functional interpretation [25] (recall also the previous references) or the monotone modified realizability interpretation [27]. However, following e.g. [29], we here prefer the presentation where these two processes are firmly disentangled.

Now, as mentioned previously, the crucial “novelty” of the bound extraction theorem for the system $\mathcal{F}^\omega[\mathbb{P}]$ as stated herein, compared to the one presented in [46], is the inclusion of a uniform boundedness principle for the types Ω and S . This principle, pioneered in [25, 26] by Kohlenbach, and then developed further substantially in [17, 27, 28, 30, 31], has rather strong and even set-theoretically false conclusions in general. Nevertheless, it can be used admissibly in the context of bound extraction theorems to derive set-theoretically correct and uniform bounds. In the present paper, as already discussed in the introduction, we will in particular show how this principle allows our systems to formally use contents and associated outer contents, to a rather large degree, as if they were probability measures and associated outer measures, all while staying admissible in the context of bound extraction theorems, so that rather strikingly the resulting theorems will nevertheless still be true for contents. Even further, we in fact allow for uniform boundedness *in higher types* in this paper, which has particularly interesting implications in the context of probability theory.

In any case, the uniform boundedness principles considered in this paper take the following form:⁴

⁴As apparent, while extended to higher types, the uniform boundedness principles considered in this paper are only defined for pure types, as this allows for a quite smooth development of the associated theory while being sufficient for the mathematical considerations of this paper. It is however quite possible that the present considerations also extend to arbitrary higher types.

Definition 2.8 (extending [30]). Let ρ be a pure type. The principle $\exists\text{-UB}^{\Omega,\rho}$ is defined as

$$\begin{cases} \forall y^{\alpha(0)} (\forall k^0, x^\alpha, \underline{z}^\beta \exists w^\rho \varphi_\exists(y, k, \min_\alpha(x, yk), \underline{z}, w)) \\ \rightarrow \exists \chi^{\rho(0)} \forall k^0, x^\alpha, \underline{z}^\beta \exists w \leq_\rho \chi k \varphi_\exists(y, k, \min_\alpha(x, yk), \underline{z}, w) \end{cases}$$

where α is a type of the form $0(\sigma_k) \dots (\sigma_1)$, each type in β is of the form $\Omega(\tau_m) \dots (\tau_1)$ and φ_\exists is an existential formula where the types of all quantifiers are admissible (see [17] and also [14, 29]). The principle $\exists\text{-UB}^{S,\rho}$ is defined analogously, using S instead of Ω in the conditions on the types in the tuple β .

This formulation is indeed very general and, following [17], represents a carefully defined intensional version of the ‘‘usual’’ uniform boundedness principle

$$\begin{cases} \forall y^{\alpha(0)} (\forall k^0, x \leq_\alpha yk, \underline{z}^\beta \exists w^\rho \varphi_\exists(y, k, x, \underline{z}, w)) \\ \rightarrow \exists \chi^{\rho(0)} \forall k^0, x \leq_\alpha yk, \underline{z}^\beta \exists w \leq_\rho \chi k \varphi_\exists(y, k, x, \underline{z}, w). \end{cases}$$

While a necessary restriction in order to stay admissible in the context of bound extraction theorems, these two formulations coincide for sentences that are extensional (see [17], to which we refer for further discussions in that vein). In any way, previous considerations of uniform boundedness in the context of the monotone functional interpretation were confined to principles with $\rho = 0$ (in the above notation). It should however be stressed that this was not really out of a technical necessity, but rather since all mathematically meaningful applications of uniform boundedness were covered by uniform boundedness in that lowest type. In that way, a key contribution of the present paper is even the consideration of higher type uniform boundedness in the context of the monotone functional interpretation per se, but in particular that it uncovers probability theory as a mathematical area where these higher principles have mathematically useful consequences.

The following is now the classical metatheorem for probability contents, as established in [46], extended to include the above uniform boundedness principles. For that purpose, we denote the union of all instances of $\exists\text{-UB}^{\Omega,\rho}$ and $\exists\text{-UB}^{S,\rho}$, for pure types ρ , by $\exists\text{-UB}^{\Omega,\omega}$ and $\exists\text{-UB}^{S,\omega}$, respectively. These principles can now be treated by a rather simple extension of the usual approach to the uniform boundedness principles in the context of the monotone functional interpretation as given in [30] (see also [17]). In particular, this extension (for the case $\rho = 1$, with the higher cases following a similar pattern) is due to U. Kohlenbach, who kindly allowed us to include it here. In every other regard, the proof is a straightforward amalgamation of the proof given in [46] together with the discussions in [17, 30], so that we confine ourselves to the treatment of $\exists\text{-UB}^{\Omega,\omega}$ and $\exists\text{-UB}^{S,\omega}$.

Theorem 2.9 (extending [30] and [46]). *Let τ be admissible (see [17] and also [14, 29]) and let $\varphi_\exists(x, n)$ be an existential formula of $\mathcal{F}^\omega[\mathbb{P}]$ (with only x, n free) such that all internal quantifiers have admissible types. If*

$$\mathcal{F}^\omega[\mathbb{P}] + \exists\text{-UB}^{\Omega,\omega} + \exists\text{-UB}^{S,\omega} \vdash \forall x^\tau \exists n^0 \varphi_\exists(x, n),$$

then one can extract a partial functional $\Phi : \mathcal{S}_\tau \rightarrow \mathbb{N}$ which is total and bar-recursively computable on \mathcal{M}_τ (in the sense of Spector [59]) and such that

$$\mathcal{S}^{\omega,\Omega,S} \models \forall x^{*\hat{\tau}} \forall x \lesssim_\tau x^* \exists n \leq_0 \Phi(x^*) \varphi_\exists(x, n)$$

holds for $\mathcal{S}^{\omega,\Omega,S}$ defined via any non-empty set Ω , algebra $S \subseteq 2^\Omega$, and probability content \mathbb{P} on S (with suitable interpretations of the additional constants).⁵ Further:

⁵For example, the constant \mathbb{P} is interpreted in the models $\mathcal{M}^{\omega,\Omega,S}$ or $\mathcal{S}^{\omega,\Omega,S}$ via $[\mathbb{P}]_{\mathcal{M}^{\omega,\Omega,S}} = [\mathbb{P}]_{\mathcal{S}^{\omega,\Omega,S}} = \lambda A^S.(\mathbb{P}(A))_\circ$ for the given probability content \mathbb{P} , where $(\cdot)_\circ$ selects a canonical type 1 representation of a given

- (1) If $\widehat{\tau}$ is of degree 1, then Φ is a total computable functional.
- (2) We may have tuples instead of single variables x, n .
- (3) The system used in the premise may be extended with any further set of axioms \beth of type Δ (see e.g. [46] for a definition in the context of $\mathcal{F}^\omega[\mathbb{P}]$), in which case the conclusion is true whenever $\mathcal{S}^{\omega, \Omega, S} \models \beth$.
- (4) If the claim is proved without DC, then τ may further be arbitrary and Φ will be a total functional on $\mathcal{S}_{\widehat{\tau}}$ that is primitive recursive in the sense of Gödel [15] and Hilbert [19].
- (5) If the claim can be proved without using either DC or uniform boundedness, then plain majorization (see e.g. Chapter 6 in [31]) can be used instead of strong majorization.

Proof. We only discuss the treatment of the higher uniform boundedness principle $\exists\text{-UB}^{\Omega, \rho}$ for a pure type $\rho = 0(\xi)$, the principle $\exists\text{-UB}^{S, \rho}$ is treated analogously. As with $\exists\text{-UB}^{\Omega, 0}$ and any other uniform boundedness principle for that matter (see e.g. [17, 30]), we actually treat an associated stronger (type Δ) axiom $F^{\Omega, \rho}$ defined as

$$\forall \Phi^{\rho \underline{\beta}^t(\alpha)(0)}, y^{\alpha(0)} \exists X^{\alpha(0)(\xi)} \leq_{\alpha(0)(0)} \lambda i^0. y \exists \underline{Z}^{\beta(0)(\xi)} \forall u^\xi, k^0, x^\alpha, \underline{z}^\beta \left(\Phi(k, X(u, k), \underline{Z}(u, k), u) \geq_0 \Phi(k, \min_\alpha(x, yk), \underline{z}, u) \right).$$

where, given $\beta = (\beta_1, \dots, \beta_m)$, we write $\underline{\beta}^t := (\beta_m) \dots (\beta_1)$ and $\underline{\beta}(0) := (\beta_1(0), \dots, \beta_m(0))$. To see that treating $F^{\Omega, \rho}$ suffices, we have the following claim:

Claim: $\mathcal{F}^\omega[\mathbb{P}] + F^{\Omega, \rho} \vdash \exists\text{-UB}^{\Omega, \rho}$.

Proof: For that, let y be given and assume that $\forall k^0, x^\alpha, \underline{z}^\beta \exists w^\rho \varphi_\exists(y, k, \min_\alpha(x, yk), \underline{z}, w)$. By QF-AC, we get a functional Φ such that

$$\forall k^0, x^\alpha, \underline{z}^\beta \varphi_\exists(y, k, \min_\alpha(x, yk), \underline{z}, \Phi(k, x, z)).$$

Provably, without any assumptions, one has $\min_\alpha(\min_\alpha(x, yk), yk) =_\alpha \min_\alpha(x, yk)$. Thus, by the quantifier-free extensionality rule QF-ER, we have

$$\forall k^0, x^\alpha, \underline{z}^\beta \varphi_\exists(y, k, \min_\alpha(x, yk), \underline{z}, \Phi(k, \min_\alpha(x, yk), z)).$$

Using $F^{\Omega, \rho}$, we get an X and \underline{Z} such that

$$\forall u^\xi, k^0, x^\alpha, \underline{z}^\beta \left(\Phi(k, X(u, k), \underline{Z}(u, k), u) \geq_0 \Phi(k, \min_\alpha(x, yk), \underline{z}, u) \right).$$

In particular, for any k, x and \underline{z} , we have

$$\lambda u^\xi. \Phi(k, X(u, k), \underline{Z}(u, k), u) \geq_\rho \Phi(k, \min_\alpha(x, yk), \underline{z}),$$

so that $\chi := \lambda k^0, u^\xi. \Phi(k, X(u, k), \underline{Z}(u, k), u)$ satisfies the conclusion of $\exists\text{-UB}^{\Omega, \rho}$. ■

To now formally deal with $F^{\Omega, \rho}$ in the context of a metatheorem, we proceed like with any other type Δ axiom (see e.g. [17, 46]) and consider the witnessed Skolemization $\widetilde{F}^{\Omega, \rho}$ defined as

$$\forall \Phi^{\rho \underline{\beta}^t(\alpha)(0)}, y^{\alpha(0)}, u^\xi, k^0, x^\alpha, \underline{z}^\beta \left(\mathcal{X}(\Phi, y, u) \leq_{\alpha(0)} y \wedge \Phi(k, \mathcal{X}(\Phi, y, u, k), \underline{\mathcal{Z}}(\Phi, y, u, k), u) \geq_0 \Phi(k, \min_\alpha(x, yk), \underline{z}, u) \right).$$

real as defined in [29]. While this operation is naturally non-effective, it is well-behaved w.r.t. majorization which serves all intents and purposes. In this paper, we do not rely on these aspects any further and generally refer to [46] for further discussions on this in relation to probability theory, as well as the interpretations of the other constants.

where \mathcal{X} and $\underline{\mathcal{Z}}$ are new constants of type $\alpha(0)(\xi)(\alpha(0))(\rho\beta^t(\alpha)(0))$ and $\underline{\beta}(0)(\xi)(\alpha(0))(\rho\beta^t(\alpha)(0))$, respectively. Note that \mathcal{X} and $\underline{\mathcal{Z}}$ solve the functional interpretation of $F^{\Omega,\rho}$ and, as $(F^{\Omega,\rho})'$ is intuitionistically implied by $F^{\Omega,\rho}$, they can also be used to solve that of $(F^{\Omega,\rho})'$. It remains to give an interpretation of these new constants in the model based on $\mathcal{M}^{\omega,\Omega,S}$ such that $\tilde{F}^{\Omega,\rho}$ is satisfied and that they are majorized by closed terms from \mathcal{A}^ω . For that, let Φ, y from $\mathcal{M}^{\omega,\Omega,S}$ be given with majorants Φ^*, y^* . Then, for any x, \underline{z} and k in $\mathcal{M}^{\omega,\Omega,S}$, $\min_\alpha(x, yk)$ is majorized by y^*k and \underline{z} by the tuple of constant-1 functionals \underline{z}^* , as all elements of \underline{z} have value type Ω . Hence, we have

$$\forall x \in \mathcal{M}_\alpha, \underline{z} \in \mathcal{M}_{\underline{\beta}} \left(\Phi^*(k, y^*k, \underline{z}^*) \succeq_\rho \Phi(k, \min_\alpha(x, yk), \underline{z}) \right).$$

For any $u \in \mathcal{M}_\xi$, we can define

$$\text{Max}_{\Phi, y, k}(u) := \max\{\Phi(k, \min_\alpha(x, yk), \underline{z}, u) \mid x \in \mathcal{M}_\alpha \text{ and } \underline{z} \in \mathcal{M}_{\underline{\beta}}\}$$

which is well-defined as $\Phi^*(k, y^*k, \underline{z}^*, u^*) \geq_0 \Phi(k, \min_\alpha(x, yk), \underline{z}, u)$ for any such x and z , where $u^* \succeq_\xi u$ (which exists as $u \in \mathcal{M}_\xi$). In particular, as the maximum is attained, we have

$$\forall \Phi, y, u, k \in \mathcal{M}^{\omega,\Omega,S} \exists x, \underline{z} \in \mathcal{M}^{\omega,\Omega,S} (x \leq_\alpha yk \wedge \Phi(k, x, \underline{z}, u) =_0 \text{Max}_{\Phi, y, k}(u)).$$

Using the axiom of choice, we get functionals \mathfrak{X} and \mathfrak{Z} such that $\mathfrak{X}(\Phi, y, u) \leq y$ and

$$\Phi(k, \mathfrak{X}(\Phi, y, u, k), \mathfrak{Z}(\Phi, y, u, k), u) =_0 \text{Max}_{\Phi, y, k}(u) \geq \Phi(k, \min_\alpha(x, yk), \underline{z}, u)$$

for any $\Phi, y, k, x, \underline{z}$ and any i . We set $[\mathcal{X}]_{\mathcal{M}^{\omega,\Omega,S}} := \mathfrak{X}$ and $[\mathcal{Z}]_{\mathcal{M}^{\omega,\Omega,S}} := \mathfrak{Z}$, which are well-defined as the constant-1 functionals majorize \mathfrak{Z} and since $\lambda\Phi^*, y^*, u.y^*$ majorizes \mathfrak{X} . It is clear that these functionals satisfy $\tilde{F}^{\Omega,\rho}$. \square

Note in particular that the extractable bounds, as guaranteed by the above metatheorem, subscribe to a high degree of uniformity in the sense that they will be independent of all parameters relating to the underlying probability content space (see also again the discussion in [46]).

We now present the semi-constructive variants of the above metatheorem. In fact, we will discuss two variants of a semi-constructive metatheorem here, one based on the Dialectica interpretation and the other based on modified realizability. While we focus on the Dialectica interpretation throughout, we still highlight the metatheorem based on the modified realizability interpretation, as if a proof does not rely on those aspects of the theory underpinned by Markov's principle, then it provides a just as viable tool, with the added bonus that one can allow an even larger class of non-computational principles in proofs without voiding the bound extraction theorems (see in particular also the discussions in [31]).

Crucially, also here we can allow a uniform boundedness principle which will similarly include types higher than 0. In contrast to the classical case, the main benefit of the semi-constructive case even allows for this principle to be unrestricted regarding the formula (see e.g. the discussion in [27]). In that way, the principles considered in the present paper take the following form:

Definition 2.10 (extending [25] and [30]). Let ρ be a pure type. The principle $\text{UB}^{\Omega,\rho}$ is defined as

$$\left\{ \begin{array}{l} \forall y^{\alpha(0)} (\forall k^0, x^\alpha, \underline{z}^\beta \exists w^\rho \varphi(y, k, \min_\alpha(x, yk), \underline{z}, w) \\ \rightarrow \exists \chi^{\rho(0)} \forall k^0, x^\alpha, \underline{z}^\beta \exists w \leq_\rho \chi k \varphi(y, k, \min_\alpha(x, yk), \underline{z}, w) \end{array} \right.$$

where α is a type of the form $0(\sigma_k) \dots (\sigma_1)$, each type in $\underline{\beta}$ is of the form $\Omega(\tau_m) \dots (\tau_1)$ and φ is any formula. The principle $\text{UB}^{S,\rho}$ is defined analogously, using S instead of Ω in the conditions on the types in the tuple $\underline{\beta}$.

As commented on above, in the presence of extensionality for φ , the above principle in fact implies the stronger uniform boundedness principle

$$\left\{ \begin{array}{l} \forall y^{\alpha(0)} (\forall k^0, x \leq_{\alpha} yk, \underline{z}^{\beta} \exists w^{\rho} \varphi(y, k, x, \underline{z}, w) \\ \rightarrow \exists \chi^{\rho(0)} \forall k^0, x \leq_{\alpha} yk, \underline{z}^{\beta} \exists w \leq_{\rho} \chi k \varphi(y, k, x, \underline{z}, w)). \end{array} \right.$$

We now state the resulting metatheorem as derived using the monotone Dialectica interpretation. While this metatheorem is new to the literature, we also here do not expand on the proof as it arises via an easy combination of the considerations from [46] with [13] together with those of [25] for the uniform boundedness principle, adapted to the abstract types in a similar way as in [30, 17] and to higher types as in the proof of the previous Theorem 2.9. Similar to before, we write $\text{UB}^{\Omega,\omega}$ for the union of the principles $\text{UB}^{\Omega,\rho}$, as well as $\text{UB}^{S,\omega}$ for the union of the principles $\text{UB}^{S,\rho}$, both for pure types ρ .

Theorem 2.11. *Let τ be admissible and let $\varphi(x, n)$ be any formula of $\mathcal{F}_i^{\omega}[\mathbb{P}]$ (with only x, n free) such that all positively occurring universal quantifiers and all negatively occurring existential quantifiers have small types and all other types are admissible. If*

$$\mathcal{F}_i^{\omega}[\mathbb{P}] + \text{UB}^{\Omega,\omega} + \text{UB}^{S,\omega} \vdash \forall x^{\tau} \exists n^0 \varphi(x, n),$$

then one can extract a functional $\Phi : \mathcal{S}_{\tau} \rightarrow \mathbb{N}$ which is primitive recursive (in the sense of Gödel [15] and Hilbert [19]) and such that

$$\mathcal{S}^{\omega,\Omega,S} \models \forall x^{*\hat{\tau}} \forall x \lesssim_{\tau} x^* \exists n \leq_0 \Phi(x^*) \varphi(x, n)$$

holds for $\mathcal{S}^{\omega,\Omega,S}$ defined via any non-empty set Ω , algebra $S \subseteq 2^{\Omega}$, and probability content \mathbb{P} on S (and with suitable interpretations of the additional constants). Further:

- (1) As before, we may have tuples instead of single variables x, n and the system used in the premise may be extended with any further set of axioms \beth of type Δ (see e.g. [46] for a definition in the context of $\mathcal{F}^{\omega}[\mathbb{P}]$), in which case the conclusion is true whenever $\mathcal{S}^{\omega,\Omega,S} \models \beth$.
- (2) If the proof does not use Markov's principle M^{ω} , then (using the monotone modified realizability interpretation) the conclusion remains valid even in the context of the independence of premise principle IP_{ef}^{ω} for existential-free formulas (see e.g. Chapter 5 in [31]) and any further set of axioms \beth of type Θ (see e.g. Chapter 7 in [31]), in which case the conclusion is true whenever $\mathcal{S}^{\omega,\Omega,S} \models \beth$.
- (3) If the claim is proved without uniform boundedness, then plain majorization (see e.g. Chapter 6 in [31]) can be used instead of strong majorization.

As before, also the bounds guaranteed by this semi-constructive metatheorem are highly uniform, being again independent of all parameters relating to the underlying probability content space.

We now provide further extensions in this semi-constructive context to also accommodate the so-called principle of *contra-collection*, which has similar mathematically useful consequences than uniform boundedness later on. This principle is motivated from its use in the context of the bounded functional interpretation [12], but introduced here in a style that is suitably adapted to the monotone functional interpretation for the first time.

Definition 2.12. The principle $\text{QF-CC}^{\Omega,0}$ is defined as

$$\forall k^0 \exists \underline{z}^\beta \forall n \leq_0 k \varphi_0(z, n) \rightarrow \exists \underline{z}^\beta \forall n^0 \varphi_0(z, n),$$

where each type in $\underline{\beta}$ is of the form $\Omega(\tau_m) \dots (\tau_1)$ and φ_0 is a quantifier-free formula. The principle $\text{QF-CC}^{S,0}$ is defined analogously, using S instead of Ω in the conditions on the types in the tuple $\underline{\beta}$.

We shall show that under suitable restrictions on the formula, a version of Theorem 2.11 holds in the presence of $\text{QF-CC}^{S,0}$ and $\text{QF-CC}^{\Omega,0}$. However, to provide a treatment, we require a certain “limited” use of these principles, which nevertheless suffices for all applications.

Concretely, in what follows, we write $\oplus \text{QF-CC}^{\Omega,0}$ (viz. $\oplus \text{QF-CC}^{S,0}$) in the specification of a system to signify that, in the context of a derivation, the principle $\text{QF-CC}^{\Omega,0}$ (viz. $\text{QF-CC}^{S,0}$) must not be used in the proof of the premise of an application of the quantifier-free rule of extensionality QF-ER . This in particular has the consequence that while systems with the quantifier-free extensionality rule, which do not satisfy the deduction theorem as a whole, nevertheless then satisfy the deduction theorem relative to the premise $\text{QF-CC}^{\Omega,0}$ (viz. $\text{QF-CC}^{S,0}$), as can be immediately verified by an induction on the length of the proof.

This availability of a fragment of the deduction theorem will be crucial for us in the following result to discharge these premises in the context of a bound extraction theorem. Indeed, we will be able to deal with this additional premise of contra-collection as follows: Using the deduction theorem, we move each such application to the premise of an implication and strengthen this principle to a formula \sqsupset of type Δ , motivated along the treatment of weak König’s lemma given by Avigad and Feferman [3] (which is slightly different compared to the first treatment of weak König’s lemma in the context of the (monotone) functional interpretation given by Kohlenbach [24]). To this compound statement, we then apply the previous semi-constructive metatheorem, which extracts quantitative information on the conclusion that is verified in the context of a so-called ε -weakening \sqsupset_ε (see e.g. Chapter 10 in [31]), which can then be discharged.

Theorem 2.13. *Let τ be admissible and let $\varphi_\exists(x^\tau, n^0)$ be an existential formula of $\mathcal{F}_i^\omega[\mathbb{P}]$ (with only x, n free) such that all internal quantifiers have admissible types. If*

$$\mathcal{F}_i^\omega[\mathbb{P}] + \text{UB}^{\Omega,\omega} + \text{UB}^{S,\omega} \oplus \text{QF-CC}^{\Omega,0} \oplus \text{QF-CC}^{S,0} \vdash \forall x^\tau \exists n^0 \varphi_\exists(x, n),$$

then one can extract a functional $\Phi : \mathcal{S}_\tau \rightarrow \mathbb{N}$ which is primitive recursive (in the sense of Gödel [15] and Hilbert [19]) and such that

$$\mathcal{S}^{\omega,\Omega,S} \models \forall x^{*\hat{\tau}} \forall x \lesssim_\tau x^* \exists n \leq_0 \Phi(x^*) \varphi_\exists(x, n)$$

holds for $\mathcal{S}^{\omega,\Omega,S}$ defined via any non-empty set Ω , algebra $S \subseteq 2^\Omega$, and probability content \mathbb{P} on S (and with suitable interpretations of the additional constants).

We have the same remark (1) as in Theorem 2.11.

Proof. We only discuss the treatment of $\text{QF-CC}^{\Omega,0}$, the principle $\text{QF-CC}^{S,0}$ can be treated analogously and simultaneously. So suppose we have

$$\mathcal{F}_i^\omega[\mathbb{P}] + \text{UB}^{\Omega,\omega} + \text{UB}^{S,\omega} \oplus \text{QF-CC}^{\Omega,0} \vdash \forall x^\tau \exists n^0 \varphi_\exists(x, n).$$

As a proof is finite, only finitely many instances of the principle $\text{QF-CC}^{\Omega,0}$ are used. For simplicity, we assume it is only one such principle, say for the formula ψ_0 , to which we still refer by $\text{QF-CC}^{\Omega,0}$. As \oplus allows for the deduction theorem, we have

$$\mathcal{F}_i^\omega[\mathbb{P}] + \text{UB}^{\Omega,\omega} + \text{UB}^{S,\omega} \vdash \text{QF-CC}^{\Omega,0} \rightarrow \forall x^\tau \exists n^0 \varphi_\exists(x, n).$$

Consider the formula

$$\beth := \exists \underline{z}^\beta \forall k^0 (\exists \underline{z}^* \forall m \leq_0 k \psi_0(\underline{z}^*, m) \rightarrow \psi_0(\underline{z}, k)).$$

It is clear that $\beth \rightarrow \text{QF-CC}^{\Omega,0}$ is provable in $\mathcal{F}_i^\omega[\mathbb{P}]$, and so we have,

$$\mathcal{F}_i^\omega[\mathbb{P}] + \text{UB}^{\Omega,\omega} + \text{UB}^{S,\omega} \vdash \beth \rightarrow \forall x^\tau \exists n^0 \varphi_\exists(x, n).$$

Thus, we have

$$\mathcal{F}_i^\omega[\mathbb{P}] + \text{UB}^{\Omega,\omega} + \text{UB}^{S,\omega} \vdash \exists \underline{z}^\beta \forall k^0 \psi(\underline{z}, k) \rightarrow \forall x^\tau \exists n^0 \varphi_0(x, n).$$

where

$$\psi(\underline{z}, k) := \exists \underline{z}^* \forall n \leq_0 k \varphi_0(\underline{z}^*, n) \rightarrow \varphi_0(\underline{z}, k).$$

Using additional epsilon-terms (see e.g. [17]), which we suppress, we can regard ψ as quantifier-free. By \mathcal{M}^ω and IP_\forall^ω , we have that $\mathcal{F}_i^\omega[\mathbb{P}] + \text{UB}^{\Omega,\omega} + \text{UB}^{S,\omega}$ proves

$$\mathcal{F}_i^\omega[\mathbb{P}] + \text{UB}^{\Omega,\omega} + \text{UB}^{S,\omega} \vdash \forall x^\tau \forall \underline{z}^\beta \exists n^0 \exists k^0 (\psi(\underline{z}, k) \rightarrow \varphi_\exists(x, n)).$$

Using Theorem 2.11, we get that there exists a functional Φ such that for all $x \in \mathcal{M}_\tau$, $x^* \in \mathcal{M}_{\hat{\tau}}$, if $x^* \gtrsim x$, then

$$\mathcal{M}^{\omega,\Omega,S} \models \forall \underline{z}^\beta \exists n \leq_0 \Phi(x^*) \exists k \leq_0 \Phi(x^*) (\psi(\underline{z}, k) \rightarrow \varphi_\exists(x, n)).$$

This implies

$$\mathcal{M}^{\omega,\Omega,S} \models \beth_\epsilon \rightarrow \exists n \leq_0 \Phi(x^*) \varphi_\exists(x, n),$$

where we write

$$\beth_\epsilon := \forall i^0 \exists \underline{z}^\beta \forall k \leq_0 i (\exists \underline{z}^* \forall m \leq_0 k \psi_0(\underline{z}^*, m) \rightarrow \psi_0(\underline{z}, k)).$$

Now, we have that $\mathcal{M}^{\omega,\Omega,S} \models \beth_\epsilon$. To see this, let i be given and let us take $k^* \leq i$ maximal such that $\exists \underline{z} \forall m \leq k^* \psi_0(\underline{z}, m)$. We may suppose that such a k^* exists, as if not, then we would in particular have $\forall \underline{z} \neg \psi_0(\underline{z}, 0)$, which yields that the premise of \beth_ϵ is always false, so that \beth_ϵ itself is satisfied and the claim already holds true. Given such a maximal k^* , take a corresponding \underline{z} with $\forall m \leq k^* \psi_0(\underline{z}, m)$. Let $k \leq i$ be arbitrary. If $k \leq k^*$, then $\psi_0(\underline{z}, k)$ and so the conclusion of \beth_ϵ is satisfied, so that \beth_ϵ itself is satisfied. If $k > k^*$, then by maximality of k^* , we have $\forall \underline{z} \exists m \leq k \neg \psi_0(\underline{z}, m)$, which yields that the premise of \beth_ϵ is false, so that again \beth_ϵ itself is satisfied. Thus, we have $\mathcal{M}^{\omega,\Omega,S} \models \exists n \leq_0 \Phi(x^*) \varphi_\exists(x, n)$ and the result follows as the types are low. \square

While the above can potentially be extended to principles $\text{QF-CC}^{\Omega,1}$ and $\text{QF-CC}^{S,1}$ of contraction at type 1, we do not pursue this here further and remain with the above result.

3. A SYSTEMATIC APPROACH TO REPRESENTING PROBABILISTIC STATEMENTS

In this section, we present a general approach for formally representing almost sure statements in the language of the formal systems presented in Section 2. To that end, we will rely on the notions of outer (and inner) content (or measure) of a formula in this language, inspired by the corresponding notion(s) from ordinary probability theory.

As we will illustrate in the following Section 5, this representation can then be used, in combination with a monotone variant of Gödel's functional interpretation due to Kohlenbach, to systematically assign (and hence extract) the computational content of measure-theoretic statements from semi-constructive or even classical proofs (if chained with a negative translation), resulting ultimately in dedicated logical metatheorems for such statements in the style of proof mining.

To motivate our formal representation of probabilistic statements, consider a content space (Ω, S, \mathbb{P}) and suppose that we are given a property $\varphi(\varpi)$ with a variable ϖ over Ω which induces a measurable property

$$A_\varphi := \{\varpi \in \Omega \mid \varphi(\varpi)\} \in S.$$

In that case, we can formally assign a probability to the statement φ by setting

$$\mathbb{P}(\varphi) := \mathbb{P}(A_\varphi)$$

so that φ being true almost surely is precisely captured by the statement that $\mathbb{P}(\varphi) = 1$. If we are interested in capturing this approach formally, we encounter an immediate logical problem. The approach requires both a formal notion of a “measurable formula” and a comprehension principle to construct the corresponding measurable sets. However, this comprehension principle is logically highly complicated due to hidden quantifiers, and their inclusion would “taint” the translation with auxiliary complexity that stems only from this arbitrary representation.

Instead, we opt for a representation that does not attempt to eliminate the quantifiers via such comprehension, but rather leaves the formula (and, as we will see later, also its computational content) intact in an appropriate way. The device that facilitates this lift is essentially the outer measure (see e.g. [23]), or more appropriately the outer content since we are working over content spaces (see e.g. [56]), a measure-theoretic device which assigns a probability to a generic set, measurable or not, only by putting it into relation with other measurable sets. Concretely, associated to the probability content \mathbb{P} is the outer content \mathbb{P}^* of \mathbb{P} that assigns to each set $A \subseteq \Omega$, measurable or not, the value

$$\mathbb{P}^*(A) := \inf\{\mathbb{P}(B) \mid A \subseteq B \text{ and } B \in S\},$$

which coincides with the content $\mathbb{P}(A)$ when A is measurable. Instead of considering $\mathbb{P}(A_\varphi) = 1$ as a representation of φ being true almost surely, we may thus consider $\mathbb{P}^*(A_\varphi) = 1$. While mathematically the same, this notion $\mathbb{P}^*(A_\varphi) = 1$, based on the outer content, can, due to its simpler definition, be formally captured rather immediately by the statement

$$\forall A^S (\forall \varpi^\Omega (\varphi(\varpi) \rightarrow \varpi \in A) \rightarrow \mathbb{P}(A) =_{\mathbb{R}} 1).$$

Now, as with the outer content in general, it is actually irrelevant for this formulation whether A_φ is measurable or not. In the same way, as also immediately apparent from its phrasing, the above property is suitable for arbitrary statements $\varphi(\varpi^\Omega)$. It is precisely this formulation that underlies our representation of probabilistic statements.

However, to arrive at a general working notion, we will first require some technical considerations. First, to be able to quantify over the degree of probability, we do not want to restrict to a representation of almost sure notions only. Therefore, we will consider the more general abbreviation

$$\mathbb{P}[\varphi] \geq \lambda := \forall A^S (\forall \varpi^\Omega (\varphi(\varpi) \rightarrow \varpi \in A) \rightarrow \mathbb{P}(A) \geq_{\mathbb{R}} \lambda)$$

where $\varphi(\varpi^\Omega)$ is a given, arbitrary, formula as before and λ^1 is a free variable of type 1. This notion, of which “ φ being true almost surely” is a special case, by setting $\lambda = 1$, comes naturally equipped with a dual notion of low probability,

$$\mathbb{P}[\varphi] \leq \lambda := \mathbb{P}[\neg\varphi] \geq 1 - \lambda.$$

Further, as we will later see, this dual notion can be equivalently captured via

$$\mathbb{P}[\varphi] \leq \lambda \leftrightarrow \forall A^S (\forall \varpi^\Omega (\varpi \in A \rightarrow \varphi(\varpi)) \rightarrow \mathbb{P}(A) \leq_{\mathbb{R}} \lambda)$$

and, similarly to the abbreviation $\mathbb{P}[\varphi] \geq \lambda$, which is motivated using the outer content, this dual notion can be understood as being motivated by the inner content \mathbb{P}_* associated with \mathbb{P}

(see e.g. [56]), defined for generic sets $A \subseteq \Omega$ via

$$\mathbb{P}_*(A) := \sup\{\mathbb{P}(B) \mid A \supseteq B \text{ and } B \in S\}.$$

As these two abbreviations are dually linked via the negation of the formula in that way, it is in particular justified to drop the super- and subscripts used in the above notation of the outer and inner content as they yield a single underlying concept of probability, high or low, for a formula.

Now, for the purpose of dealing with classical reasoning, these formal representations of probabilistic formulas provide the precise internal notions needed to represent the associated probability-theoretic statements in a suitable way, which allows for applications in proof mining. However, in this paper, we will also be concerned with semi-constructive reasoning, in which the above property proves to be wholly unsuitable for this purpose. Indeed, the above formulation of the outer/inner content of a formula is inherently classical as φ , appearing in a double negative position, is treated by essentially all proof interpretations as being inside a double negation. This has the effect that, for example, the modified realizability interpretation of $\mathbb{P}[\varphi] \geq 1$ (i.e. our representation of φ being true almost surely) just extracts the empty realizer, while the functional interpretation extracts a realizer similar to that for the double negation of φ .

Thus, to obtain a notion that uniformly represents such probabilistic properties, and moreover preserves the computational meaning of φ while suitably lifting it to the probabilistic level in both classical and constructive settings, we consider the following abbreviation as our formal representation:

Definition 3.1. Let $\varphi(\varpi^\Omega)$ be a formula and λ^1 be a free variable. We define the abbreviation

$$(P) \quad \mathbb{P}[\varphi] \geq \lambda \quad :\equiv \quad \forall A^S (\mathbb{P}(A) <_{\mathbb{R}} \lambda \rightarrow \exists \varpi^\Omega \in A^c \varphi(\varpi)).$$

Dually, we define

$$(D) \quad \mathbb{P}[\varphi] \leq \lambda \quad :\equiv \quad \mathbb{P}[\bar{\varphi}] \geq 1 - \lambda$$

where $\bar{\varphi}$ is the De Morgan dual of φ .

Classically, the above notion is equivalent to our previous outer/inner content, so that in that context, no expressivity is lost. Constructively, however, our “official” definition is much more productive as φ , and by duality also $\bar{\varphi}$ in the low probability variant, appear positively. It should be noted that the motivation for utilizing the De Morgan dual $\bar{\varphi}$ instead of $\neg\varphi$ is again due to constructive considerations, as $\bar{\varphi}$ provides a constructively much more productive notion than the object $\neg\varphi$, for which, for example, the modified realizability interpretation again just extracts the empty realizer.

Following standard probability theory notation, we generally omit parameters of type Ω in expressions of the form $\mathbb{P}[\varphi] \geq \lambda$. E.g., we write $\mathbb{P}[Qx^\tau \varphi(x)] \geq \lambda$ for $\mathbb{P}[Qx^\tau \varphi(x, \varpi)] \geq \lambda$, where $Q \in \{\forall, \exists\}$ is some quantifier, and we write $\mathbb{P}[A] \geq \lambda$ for $\mathbb{P}[\varpi \in A] \geq \lambda$, where A is of type S .

4. THE ALGEBRA OF PROBABILISTIC STATEMENTS

In this section, we study the basic algebra of abbreviations (P) and (D). To that end, we begin by noting some trivial monotonicity properties and fixing some further notation in that context: Clearly, if $\lambda \geq_{\mathbb{R}} \mu$ and $\mathbb{P}[\varphi] \geq \lambda$, then also $\mathbb{P}[\varphi] \geq \mu$. Likewise, we have that $\lambda \leq_{\mathbb{R}} \mu$ and $\mathbb{P}[\varphi] \leq \lambda$ imply $\mathbb{P}[\varphi] \leq \mu$. Further, we write $\mathbb{P}[\varphi] > \lambda$ or $\mathbb{P}[\varphi] < \lambda$ for the respective non-strict variants resulting from replacing $\geq_{\mathbb{R}}$ or $\leq_{\mathbb{R}}$ in the above by $>_{\mathbb{R}}$ or $<_{\mathbb{R}}$, respectively. It should be noted, however, that these only serve a theoretical purpose later on as $>_{\mathbb{R}}$ and $<_{\mathbb{R}}$

are existential, which will be convenient in some places. The mathematically more appropriate phrasing of these strict relations of the outer and inner content would rather be given by

$$\exists \mu^{\mathbb{Q}} > 0 (\mathbb{P}[\varphi] \geq \lambda + \mu) \text{ and } \exists \mu^{\mathbb{Q}} > 0 (\mathbb{P}[\varphi] \leq \lambda - \mu),$$

but if such a relation is required in our context, we will always use the above formulas explicitly.

Now as mentioned briefly before, the definition (D) of low probability, which is simply connected to the definition (P) of high probability by duality, can be given a formal definition inspired by inner contents analogous to the official definition of (P) inspired by outer contents:

Proposition 4.1. *For any formula $\varphi(\varpi^{\Omega})$:*

$$\mathcal{F}_i^{\omega}[\mathbb{P}] \vdash \mathbb{P}[\varphi] \leq \lambda \leftrightarrow \forall A^S(\mathbb{P}(A) >_{\mathbb{R}} \lambda \rightarrow \exists \varpi^{\Omega} \in A \bar{\varphi}(\varpi)).$$

Proof. Working over $\mathcal{F}_i^{\omega}[\mathbb{P}]$ and fixing λ^1 , suppose that $\mathbb{P}[\varphi] \leq \lambda$. We have by definition that

$$\forall A^S(\mathbb{P}(A) <_{\mathbb{R}} 1 - \lambda \rightarrow \exists \varpi^{\Omega} \in A^c \bar{\varphi}(\varpi)).$$

Instantiating A with A^c , this implies

$$\forall A^S(\mathbb{P}(A^c) <_{\mathbb{R}} 1 - \lambda \rightarrow \exists \varpi^{\Omega} \in (A^c)^c \bar{\varphi}(\varpi)).$$

Since $\mathcal{F}_i^{\omega}[\mathbb{P}]$ proves that $(A^c)^c =_S A$ and $1 =_{\mathbb{R}} \mathbb{P}(A^c) + \mathbb{P}(A)$ we obtain

$$\forall A^S(\mathbb{P}(A) >_{\mathbb{R}} \lambda \rightarrow \exists \varpi^{\Omega} \in A \bar{\varphi}(\varpi)).$$

Similarly, we can show that $\forall A^S(\mathbb{P}(A) >_{\mathbb{R}} \lambda \rightarrow \exists \varpi^{\Omega} \in A \bar{\varphi}(\varpi))$ implies $\mathbb{P}[\varphi] \leq \lambda$. \square

This, in particular, also justifies dropping the sub- and superscripts as discussed in the previous section.

Another key observation is that measurable properties cause the outer and inner interpretations to collapse to the underlying content. For that, we introduce the abbreviation

$$\mathbb{P}[\varphi] = \lambda \text{ :} \equiv \mathbb{P}[\varphi] \leq \lambda \wedge \mathbb{P}[\varphi] \geq \lambda.$$

Proposition 4.2. *Over $\mathcal{F}_i^{\omega}[\mathbb{P}]$, suppose that φ is measurable in the sense that for some A_0^S :*

$$\forall \varpi^{\Omega} (\varphi(\varpi) \leftrightarrow \varpi \in A_0).$$

Then for any free variable λ^1 :

$$\mathbb{P}[\varphi] \geq \lambda \leftrightarrow \mathbb{P}(A_0) \geq_{\mathbb{R}} \lambda \quad \text{and} \quad \mathbb{P}[\varphi] \leq \lambda \leftrightarrow \mathbb{P}(A_0) \leq_{\mathbb{R}} \lambda.$$

In particular $\mathbb{P}[\varphi] = \mathbb{P}(A_0)$.

Proof. First, note that we have

$$\begin{aligned} \mathbb{P}[\varphi] \geq \lambda &\equiv \forall A^S (\mathbb{P}(A) <_{\mathbb{R}} \lambda \rightarrow \exists \varpi \in A^c \varphi(\varpi)) \\ &\rightarrow \mathbb{P}(A_0) <_{\mathbb{R}} \lambda \rightarrow \exists \varpi \in A_0^c \varphi(\varpi) \\ &\leftrightarrow \mathbb{P}(A_0) <_{\mathbb{R}} \lambda \rightarrow \exists \varpi \in A_0^c (\varpi \in A_0) \\ &\leftrightarrow \neg(\mathbb{P}(A_0) <_{\mathbb{R}} \lambda) \end{aligned}$$

over $\mathcal{F}_i^{\omega}[\mathbb{P}]$, as $A^c \cap A =_S \emptyset$. Conversely, assume $\neg(\mathbb{P}(A_0) <_{\mathbb{R}} \lambda)$ and let $\mathbb{P}(A) <_{\mathbb{R}} \lambda$. Now, we have that $\neg(\mathbb{P}(A_0) <_{\mathbb{R}} \lambda) \leftrightarrow \mathbb{P}(A_0) \geq_{\mathbb{R}} \lambda$. Then $A_0 \not\subseteq A$, by monotonicity of \mathbb{P} , and so there is an $\varpi \in A^c$ with $\varpi \in A_0$, that is $\varphi(\varpi)$. As A was arbitrary, we have shown

$$\neg(\mathbb{P}(A_0) <_{\mathbb{R}} \lambda) \rightarrow \forall A^S (\mathbb{P}(A) <_{\mathbb{R}} \lambda \rightarrow \exists \varpi \in A^c \varphi(\varpi)) \leftrightarrow \mathbb{P}[\varphi] \geq \lambda$$

over $\mathcal{F}_i^{\omega}[\mathbb{P}]$. So, we have shown $\mathbb{P}[\varphi] \geq \lambda \leftrightarrow \mathbb{P}(A_0) \geq_{\mathbb{R}} \lambda$ over $\mathcal{F}_i^{\omega}[\mathbb{P}]$. For the dual claim, simply note that $\forall \varpi^{\Omega} (\bar{\varphi}(\varpi) \leftrightarrow \varpi \notin A_0)$ so that using (D), we get

$$\mathbb{P}[\varphi] \leq \lambda \leftrightarrow \mathbb{P}[\bar{\varphi}] \geq 1 - \lambda \leftrightarrow \mathbb{P}(A_0^c) \geq_{\mathbb{R}} 1 - \lambda \leftrightarrow \mathbb{P}(A_0) \leq_{\mathbb{R}} \lambda,$$

using again that $1 =_{\mathbb{R}} \mathbb{P}(A^c) + \mathbb{P}(A)$, for any A^S . \square

Remark 4.3. In Proposition 4.2, note that already $\forall \varpi^\Omega (\varphi(\varpi) \rightarrow \varpi \in A_0)$ implies

$$\mathbb{P}[\varphi] \geq \lambda \rightarrow \mathbb{P}(A_0) \geq_{\mathbb{R}} \lambda \text{ and } \mathbb{P}(A_0) \leq_{\mathbb{R}} \lambda \rightarrow \mathbb{P}[\varphi] \leq \lambda$$

and $\forall \varpi^\Omega (\varpi \in A_0 \rightarrow \varphi(\varpi))$ implies

$$\mathbb{P}(A_0) \geq_{\mathbb{R}} \lambda \rightarrow \mathbb{P}[\varphi] \geq \lambda \text{ and } \mathbb{P}[\varphi] \leq \lambda \rightarrow \mathbb{P}(A_0) \leq_{\mathbb{R}} \lambda.$$

The outer content considered in this paper is the outer content \mathbb{P}^* relative to the content \mathbb{P} , and in that way a special instance of the general notion of an outer content (or measure) studied in measure theory. In the probabilistic case, this class of general outer contents on Ω comprises all set functions $\nu : 2^\Omega \rightarrow [0, 1]$ that satisfy the following three properties:

- (1) $\nu(\emptyset) = 0$,
- (2) for all $A, B \in 2^\Omega$: if $A \subseteq B$, then $\nu(A) \leq \nu(B)$,
- (3) for all $A_0, \dots, A_m \in 2^\Omega$: $\nu(\bigcup_{n=0}^m A_n) \leq \sum_{n=0}^m \nu(A_n)$.

It should be noted that outer measures, compared to outer contents, are generally required to be sub- σ -additive (see e.g. [18]) instead of just subadditive, but this property does not hold true already for the outer content \mathbb{P}^* relative to a content \mathbb{P} if the latter is only formulated over a field (see e.g. Remark 4.1.5 in [56]).

For our particular outer content \mathbb{P}^* formalized using $\mathbb{P}[\cdot]$, we can replicate these key properties formally in the system $\mathcal{F}_i^\omega[\mathbb{P}]$, identifying general sets $A \in 2^\Omega$ with arbitrary formulas $\varphi(\varpi)$ in the language of $\mathcal{F}_i^\omega[\mathbb{P}]$.

In particular, a formal variant of the first property (1) already follows from the preceding Lemma 4.2, by which any set specified by $\varphi(\varpi)$ that equals \emptyset , that is $\forall \varpi^\Omega \neg \varphi(\varpi)$, already satisfies $\mathbb{P}[\varphi] = 0$.

The next lemma now gives a formal account of the second property, formalizing inclusions $A \subseteq B$ among general sets $A, B \in 2^\Omega$ via implications $\forall \varpi^\Omega (\varphi(\varpi) \rightarrow \psi(\varpi))$ for general formulas $\varphi(\varpi), \psi(\varpi)$ in the language of $\mathcal{F}_i^\omega[\mathbb{P}]$. To model a comparison between outer contents, which in our system are not numbers but actually formal expressions involving comparisons, we introduce the notation

$$\mathbb{P}[\varphi] \leq \mathbb{P}[\psi] := \forall \lambda^1 (\mathbb{P}[\varphi] \geq \lambda \rightarrow \mathbb{P}[\psi] \geq \lambda).$$

Proposition 4.4. *Over $\mathcal{F}_i^\omega[\mathbb{P}]$, assume $\varphi(\varpi^\Omega)$ and $\psi(\varpi^\Omega)$ are formulas such that $\varphi(\varpi) \rightarrow \psi(\varpi)$ outside of a null set A_0 , i.e.*

$$\mathbb{P}(A_0) =_{\mathbb{R}} 0 \wedge \forall \varpi^\Omega \in A_0^c (\varphi(\varpi) \rightarrow \psi(\varpi)).$$

Then $\mathbb{P}[\varphi] \leq \mathbb{P}[\psi]$.

Proof. Let λ^1 be arbitrary and suppose $\mathbb{P}[\varphi] \geq \lambda$, i.e.

$$\forall A^S (\mathbb{P}(A) <_{\mathbb{R}} \lambda \rightarrow \exists \varpi^\Omega \in A^c \varphi(\varpi)).$$

Let A^S be arbitrary and suppose $\mathbb{P}(A) <_{\mathbb{R}} \lambda$. Then we have

$$\mathbb{P}(A \cup A_0) \leq_{\mathbb{R}} \mathbb{P}(A) + \mathbb{P}(A_0) <_{\mathbb{R}} \lambda$$

so that the above implies $\exists \varpi^\Omega \in (A \cup A_0)^c \varphi(\varpi)$. Provably in $\mathcal{F}_i^\omega[\mathbb{P}]$, we have $(A \cup A_0)^c =_S A^c \cap A_0^c$, and hence, we get that $\varphi(\varpi)$ implies $\psi(\varpi)$ as $\varpi \in A_0^c$. Since $\varpi \in A^c$ and A was arbitrary, we have shown that $\mathbb{P}[\psi] \geq \lambda$, i.e.

$$\forall A^S (\mathbb{P}(A) <_{\mathbb{R}} \lambda \rightarrow \exists \varpi^\Omega \in A^c \psi(\varpi)).$$

As λ was arbitrary, it follows that $\mathbb{P}[\varphi] \leq \mathbb{P}[\psi]$. \square

In particular, the above Proposition 4.4 establishes a type of *almost-sure* extensionality of the expression $\mathbb{P}[\varphi] \geq \lambda$ in φ .

Let us now consider the last property, $\nu(\bigcup_{n=0}^m A_n) \leq \sum_{n=0}^m \nu(A_n)$, for arbitrary $A_0, \dots, A_m \in 2^\Omega$. In the following, we will show how we can formally recognize this property, phrased using $\mathbb{P}[\cdot]$, in our system $\mathcal{F}_i^\omega[\mathbb{P}]$, using a formula $\varphi(n, \varpi)$ to represent the sets A_n and $\exists n \leq_0 m \varphi(n, \varpi)$ to represent their finite union $\bigcup_{n=0}^m A_n$.

To formally encompass this property, we require a representation of a finite sum of outer contents, which is again not completely immediate as outer contents are not numbers but formal comparisons. Given a formula $\varphi(n, \varpi)$, we in that vein write

$$\sum_{n=0}^m \mathbb{P}[\varphi(n)] \geq \lambda \equiv \forall \tilde{\lambda}_{(\cdot)}^{1(0)} \left(\lambda \geq_{\mathbb{R}} \sum_{n=0}^m \tilde{\lambda}_n \rightarrow \exists n \leq_0 m (\mathbb{P}[\varphi(n)] \geq \tilde{\lambda}_n) \right).$$

We now obtain the subadditivity of the outer content:

Lemma 4.5. *Over $\mathcal{F}^\omega[\mathbb{P}]$, let $\varphi(n^0, \varpi^\Omega)$ be a formula. Then*

$$\mathbb{P}[\exists n \leq_0 m \varphi(n)] \leq \sum_{n=0}^m \mathbb{P}[\varphi(n)],$$

that is

$$\forall \lambda^1 \left(\mathbb{P}[\exists n \leq_0 m \varphi(n)] \geq \lambda \rightarrow \sum_{n=0}^m \mathbb{P}[\varphi(n)] \geq \lambda \right).$$

Proof. Suppose for a contradiction that $\mathbb{P}[\exists n \leq_0 m \varphi(n)] \geq \lambda$ but $\sum_{n=0}^m \mathbb{P}[\varphi(n)] \geq \lambda$ fails. Then there is a $\tilde{\lambda}_{(\cdot)}^{1(0)}$ such that $\lambda \geq_{\mathbb{R}} \sum_{n=0}^m \tilde{\lambda}_n$ but $\neg \mathbb{P}[\varphi(n)] \geq \tilde{\lambda}_n$, for all $n \leq_0 m$. In particular, for any $n \leq_0 m$ there exists an A_n such that $\mathbb{P}(A_n) <_{\mathbb{R}} \tilde{\lambda}_n$ but $\neg \varphi(n, \varpi)$ for all $\varpi \in A_n^c$. However we then have

$$\mathbb{P} \left(\bigcup_{n=0}^m A_n \right) \leq_{\mathbb{R}} \sum_{n=0}^m \mathbb{P}(A_n) <_{\mathbb{R}} \sum_{n=0}^m \tilde{\lambda}_n \leq_{\mathbb{R}} \lambda.$$

As $\mathbb{P}[\exists n \leq_0 m \varphi(n)] \geq \lambda$ holds, we get

$$\exists \varpi \in \left(\bigcup_{n=0}^m A_n \right)^c \exists n \leq_0 m \varphi(n, \varpi).$$

In particular, we have $\varpi \in A_n^c$ and $\varphi(n, \varpi)$. This contradicts that $\neg \varphi(n, \varpi)$ for all $\varpi \in A_n^c$ as established previously. \square

We now move to the quantifier structure of a formula. Here, we begin with the following partial prenexation results:

Proposition 4.6. *For any type τ , any formula $\varphi(x^\tau, \varpi^\Omega)$, and any free variable λ^1 , $\mathcal{F}_i^\omega[\mathbb{P}]$ proves:*

- (1) $\mathbb{P}[\forall x^\tau \varphi(x)] \geq \lambda \rightarrow \forall x^\tau (\mathbb{P}[\varphi(x)] \geq \lambda)$,
- (2) $\forall \mu^{\mathbb{Q}} > 0 \exists x^\tau (\mathbb{P}[\varphi(x)] \leq \lambda + \mu) \rightarrow \mathbb{P}[\forall x^\tau \varphi(x)] \leq \lambda$.

Further, $\mathcal{F}_i^\omega[\mathbb{P}]$ also proves:

- (3) $\mathbb{P}[\exists x^\tau \varphi(x)] \leq \lambda \rightarrow \forall x^\tau (\mathbb{P}[\varphi(x)] \leq \lambda)$,
- (4) $\forall \mu^{\mathbb{Q}} > 0 \exists x^\tau (\mathbb{P}[\varphi(x)] \geq \lambda - \mu) \rightarrow \mathbb{P}[\exists x^\tau \varphi(x)] \geq \lambda$.

Proof. We are working over $\mathcal{F}_i^\omega[\mathbb{P}]$ and we only show (1) and (2). Items (3) and (4) follow from these by (D).

(1) We have

$$\begin{aligned} \mathbb{P}[\forall x^\tau \varphi(x)] \geq \lambda &\equiv \forall A^S(\mathbb{P}(A) <_{\mathbb{R}} \lambda \rightarrow \exists \varpi^\Omega \in A^c \forall x^\tau \varphi(x, \varpi)) \\ &\rightarrow \forall A^S(\mathbb{P}(A) <_{\mathbb{R}} \lambda \rightarrow \forall x^\tau \exists \varpi^\Omega \in A^c \varphi(x, \varpi)) \\ &\leftrightarrow \forall x^\tau \forall A^S(\mathbb{P}(A) <_{\mathbb{R}} \lambda \rightarrow \exists \varpi^\Omega \in A^c \varphi(x, \varpi)) \\ &\equiv \forall x^\tau (\mathbb{P}[\varphi(x)] \geq \lambda), \end{aligned}$$

where all the (bi-)implications are constructively true.

(2) We have, by Proposition 4.1, that

$$\begin{aligned} \forall \mu^{\mathbb{Q}} > 0 \exists x^\tau (\mathbb{P}[\varphi(x)] \leq \lambda + \mu) &\equiv \forall \mu^{\mathbb{Q}} > 0 \exists x^\tau \forall A^S(\mathbb{P}(A) >_{\mathbb{R}} \lambda + \mu \rightarrow \exists \varpi^\Omega \in A \overline{\varphi}(x, \varpi)) \\ &\rightarrow \forall \mu^{\mathbb{Q}} > 0 \forall A^S(\mathbb{P}(A) >_{\mathbb{R}} \lambda + \mu \rightarrow \exists \varpi^\Omega \in A \exists x^\tau \overline{\varphi}(x, \varpi)) \\ &\leftrightarrow \forall A^S(\exists \mu^{\mathbb{Q}} > 0 (\mathbb{P}(A) >_{\mathbb{R}} \lambda + \mu) \rightarrow \exists \varpi^\Omega \in A \exists x^\tau \overline{\varphi}(x, \varpi)) \\ &\leftrightarrow \forall A^S(\mathbb{P}(A) >_{\mathbb{R}} \lambda \rightarrow \exists \varpi^\Omega \in A \overline{\forall x^\tau \varphi(x, \varpi)}) \\ &\equiv \mathbb{P}[\forall x^\tau \varphi(x)] \leq \lambda, \end{aligned}$$

where again all the (bi-)implications are constructively true. \square

While the above principles are generally valid, the converse variants of these results are only valid under a monotonicity assumption, and in particular only provable in the presence of a uniform boundedness principle. Moreover, in the intuitionistic context, the converse statements subtly rely either on uniform boundedness or contra-collection, depending on polarity, which are conflated in the classical context. As the picture here starts to become more complicated, we hence first focus on the classical case. As only a fragment of the full uniform boundedness principle is permissible in a classical context while aiming for computable bounds, we have to make restrictions on the quantifier complexity of the formulas in that context.

Proposition 4.7. *Over $\mathcal{F}^\omega[\mathbb{P}]$, let τ be a pure type, let $\varphi(x^\tau, \varpi^\Omega)$ be a formula and let λ^1 be an arbitrary free variable. Then:*

(1) *If φ is a \forall -formula and if $\varphi(x, \varpi)$ is anti-monotone in x , i.e.*

$$\forall \varpi^\Omega \forall x^\tau, y^\tau (y \geq_\tau x \wedge \varphi(y, \varpi) \rightarrow \varphi(x, \varpi)),$$

then $\mathcal{F}^\omega[\mathbb{P}] + \exists\text{-UB}^{\Omega, \tau}$ proves

$$\mathbb{P}[\forall x^\tau \varphi(x)] \geq \lambda \leftrightarrow \forall x^\tau (\mathbb{P}[\varphi(x)] \geq \lambda)$$

and $\mathcal{F}^\omega[\mathbb{P}] + \exists\text{-UB}^{S, \tau}$ proves

$$\forall \mu^{\mathbb{Q}} > 0 \exists x^\tau (\mathbb{P}[\varphi(x)] < \lambda + \mu) \leftrightarrow \mathbb{P}[\forall x^\tau \varphi(x)] \leq \lambda.$$

(2) *If φ is an \exists -formula and if $\varphi(x, \varpi)$ is monotone in x , i.e.*

$$\forall \varpi^\Omega \forall x^\tau, y^\tau (y \geq_\tau x \wedge \varphi(x, \varpi) \rightarrow \varphi(y, \varpi)),$$

then $\mathcal{F}^\omega[\mathbb{P}] + \exists\text{-UB}^{\Omega, \tau}$ proves

$$\mathbb{P}[\exists x^\tau \varphi(x)] \leq \lambda \leftrightarrow \forall x^\tau (\mathbb{P}[\varphi(x)] \leq \lambda)$$

and $\mathcal{F}^\omega[\mathbb{P}] + \exists\text{-UB}^{S, \tau}$ proves

$$\forall \mu^{\mathbb{Q}} > 0 \exists x^\tau (\mathbb{P}[\varphi(x)] > \lambda - \mu) \leftrightarrow \mathbb{P}[\exists x^\tau \varphi(x)] \geq \lambda.$$

Proof. We only prove (1). Item (2) again follows using (D). For the first part of item (1), by Proposition 4.6 we are left with showing

$$\forall x^\tau (\mathbb{P}[\varphi(x)] \geq \lambda) \rightarrow \mathbb{P}[\forall x^\tau \varphi(x)] \geq \lambda.$$

Suppose $\forall x^\tau (\mathbb{P}[\varphi(x)] \geq \lambda)$, i.e.

$$\forall x^\tau \forall A^S (\mathbb{P}(A) <_{\mathbb{R}} \lambda \rightarrow \exists \varpi^\Omega \in A^c \varphi(x, \varpi)).$$

Using classical logic, this is equivalent to

$$\forall A^S (\exists x^\tau \forall \varpi^\Omega (\varphi(x, \varpi) \rightarrow \varpi \in A) \rightarrow \mathbb{P}(A) \geq_{\mathbb{R}} \lambda).$$

Now fix A and assume $\forall \varpi^\Omega \exists x^\tau (\varphi(x, \varpi) \rightarrow \varpi \in A)$. Using $\exists\text{-UB}^{\Omega, \tau}$, we get

$$\exists \chi^\tau \forall \varpi^\Omega \exists x \leq_\tau \chi (\varphi(x, \varpi) \rightarrow \varpi \in A)$$

as φ is a \forall -formula. As φ is anti-monotone, we get

$$\exists x^\tau \forall \varpi^\Omega (\varphi(x, \varpi) \rightarrow \varpi \in A).$$

By the above, we get $\mathbb{P}(A) \geq_{\mathbb{R}} \lambda$. Hence, we have shown

$$\forall A^S (\forall \varpi^\Omega (\forall x^\tau \varphi(x, \varpi) \rightarrow \varpi \in A) \rightarrow \mathbb{P}(A) \geq_{\mathbb{R}} \lambda)$$

which is classically equivalent to $\mathbb{P}[\forall x^\tau \varphi(x)] \geq \lambda$.

For the second part of item (1), by Proposition 4.6, we are similarly left with showing

$$\mathbb{P}[\forall x^\tau \varphi(x)] \leq \lambda \rightarrow \forall \mu^{\mathbb{Q}} > 0 \exists x^\tau (\mathbb{P}[\varphi(x)] < \lambda + \mu).$$

So suppose $\mathbb{P}[\forall x^\tau \varphi(x)] \leq \lambda$. Using classical logic and Proposition 4.1, this is equivalent to

$$\forall A^S (\forall \varpi^\Omega (\varpi \in A \rightarrow \forall x^\tau \varphi(x, \varpi)) \rightarrow \mathbb{P}(A) \leq_{\mathbb{R}} \lambda)$$

and hence

$$\forall \mu^{\mathbb{Q}} > 0 \forall A^S \exists x^\tau (\forall \varpi^\Omega (\varpi \in A \rightarrow \varphi(x, \varpi)) \rightarrow \mathbb{P}(A) <_{\mathbb{R}} \lambda + \mu).$$

As φ is a \forall -formula, $\exists\text{-UB}^{S, \tau}$ yields

$$\forall \mu^{\mathbb{Q}} > 0 \exists \chi^\tau \forall A^S \exists x \leq_\tau \chi (\forall \varpi^\Omega (\varpi \in A \rightarrow \varphi(x, \varpi)) \rightarrow \mathbb{P}(A) <_{\mathbb{R}} \lambda + \mu)$$

and anti-monotonicity of φ yields

$$\forall \mu^{\mathbb{Q}} > 0 \exists x^\tau \forall A^S (\forall \varpi^\Omega (\varpi \in A \rightarrow \varphi(x, \varpi)) \rightarrow \mathbb{P}(A) <_{\mathbb{R}} \lambda + \mu)$$

which is classically equivalent, using Proposition 4.1, to $\forall \mu^{\mathbb{Q}} > 0 \exists x^\tau (\mathbb{P}[\varphi(x)] < \lambda + \mu)$. \square

We now consider the constructive case. There, the second equivalence of both items of Proposition 4.7 remains admissible in the presence of the constructively (at least as far as the Dialectica interpretation is concerned) acceptable principle of independence of premise for universal formulas. Further, the restrictions previously in place on the uniform boundedness principle can now be lifted as uniform boundedness is admissible without restrictions on the formula in a semi-constructive context while aiming for the extraction of computable bounds.

Proposition 4.8. *Over $\mathcal{F}_i^\omega[\mathbb{P}]$, let τ be a pure type, let $\varphi(x^\tau, \varpi^\Omega)$ be any formula and let λ^1 be an arbitrary free variable. Then:*

(1) *If φ is anti-monotone in x , then $\mathcal{F}_i^\omega[\mathbb{P}] + \text{UB}^{S, \tau}$ proves*

$$\forall \mu^{\mathbb{Q}} > 0 \exists x^\tau (\mathbb{P}[\varphi(x)] < \lambda + \mu) \leftrightarrow \mathbb{P}[\forall x^\tau \varphi(x)] \leq \lambda.$$

(2) *If φ is monotone in x , then $\mathcal{F}_i^\omega[\mathbb{P}] + \text{UB}^{S, \tau}$ proves*

$$\forall \mu^{\mathbb{Q}} > 0 \exists x^\tau (\mathbb{P}[\varphi(x)] > \lambda - \mu) \leftrightarrow \mathbb{P}[\exists x^\tau \varphi(x)] \geq \lambda.$$

Proof. As before, we only need to prove one of the items as the other follows using (D). Here, we only show (2). By Proposition 4.6, also similar to before, we are only left to show

$$\mathbb{P}[\exists x^\tau \varphi(x)] \geq \lambda \rightarrow \forall \mu >_{\mathbb{R}} 0 \exists x^\tau (\mathbb{P}[\varphi(x)] > \lambda - \mu).$$

For that, note that we have

$$\begin{aligned} \mathbb{P}[\exists x^\tau \varphi(x)] \geq \lambda &\equiv \forall A^S (\mathbb{P}(A) <_{\mathbb{R}} \lambda \rightarrow \exists \varpi^\Omega \in A^c \exists x^\tau \varphi(x, \varpi)) \\ &\leftrightarrow \forall A^S (\exists \mu^{\mathbb{Q}} > 0 (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - \mu) \rightarrow \exists x^\tau \exists \varpi^\Omega \in A^c \varphi(x, \varpi)) \\ &\leftrightarrow \forall \mu^{\mathbb{Q}} > 0 \forall A^S (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - \mu \rightarrow \exists x^\tau \exists \varpi^\Omega \in A^c \varphi(x, \varpi)) \\ &\leftrightarrow \forall \mu^{\mathbb{Q}} > 0 \forall A^S \exists x^\tau (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - \mu \rightarrow \exists \varpi^\Omega \in A^c \varphi(x, \varpi)), \end{aligned}$$

where all equivalences are constructive but the last, which follows using $\text{IP}_{\forall}^\omega$. Using $\exists\text{-UB}^{S,\tau}$, we get

$$\forall \mu^{\mathbb{Q}} > 0 \exists \chi^\tau \forall A^S \exists x \leq_\tau \chi (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - \mu \rightarrow \exists \varpi^\Omega \in A^c \varphi(x, \varpi))$$

and as φ is monotone, we have

$$\forall \mu^{\mathbb{Q}} > 0 \exists x^\tau \forall A^S (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - \mu \rightarrow \exists \varpi^\Omega \in A^c \varphi(x, \varpi))$$

which is $\forall \mu^{\mathbb{Q}} > 0 \exists x^\tau (\mathbb{P}[\varphi(x)] > \lambda - \mu)$. \square

Now, the first equivalence of both items of Proposition 4.7 was previously derived using classical logic, essentially invoking the converse of the uniform boundedness principle. Intuitionistically, this essentially just amounts to an application of the principle of contra-collection, which in the context of the present paper is, however, restricted not only in terms of the permissible formulas but also regarding the type, resulting in the following rather restricted result:

Proposition 4.9. *Over $\mathcal{F}_i^\omega[\mathbb{P}]$, let $\varphi_0(x^0, \varpi^\Omega)$ be a quantifier-free formula and let λ^1 be an arbitrary free variable. Then:*

(1) *If φ_0 is anti-monotone in x , then $\mathcal{F}_i^\omega[\mathbb{P}] \oplus \text{QF-CC}^{\Omega,0}$ proves*

$$\mathbb{P}[\forall x^0 \varphi_0(x)] \geq \lambda \leftrightarrow \forall x^0 (\mathbb{P}[\varphi_0(x)] \geq \lambda).$$

(2) *If φ_0 is monotone in x , then $\mathcal{F}_i^\omega[\mathbb{P}] \oplus \text{QF-CC}^{\Omega,0}$ proves*

$$\mathbb{P}[\exists x^0 \varphi_0(x)] \leq \lambda \leftrightarrow \forall x^0 (\mathbb{P}[\varphi_0(x)] \leq \lambda).$$

Proof. As before, we only need to prove one of the items as the other follows using (D). Here, we only show (1). By Proposition 4.6, also similar to before, we are only left to show

$$\forall x^0 (\mathbb{P}[\varphi_0(x)] \geq \lambda) \rightarrow \mathbb{P}[\forall x^0 \varphi_0(x)] \geq \lambda.$$

For that, note that we have

$$\begin{aligned} \forall x^0 (\mathbb{P}[\varphi_0(x)] \geq \lambda) &\equiv \forall x^0 \forall A^S (\mathbb{P}(A) <_{\mathbb{R}} \lambda \rightarrow \exists \varpi^\Omega \in A^c \varphi_0(x, \varpi)) \\ &\leftrightarrow \forall A^S (\mathbb{P}(A) <_{\mathbb{R}} \lambda \rightarrow \forall x^0 \exists \varpi^\Omega \in A^c \varphi_0(x, \varpi)) \\ &\leftrightarrow \forall A^S (\mathbb{P}(A) <_{\mathbb{R}} \lambda \rightarrow \forall x^0 \exists \varpi^\Omega \in A^c \forall y \leq_0 x \varphi_0(y, \varpi)), \end{aligned}$$

where the last line follows from the anti-monotonicity of φ_0 . Using $\text{QF-CC}^{\Omega,0}$, we get

$$\forall x^0 \exists \varpi^\Omega \in A^c \forall y \leq_0 x \varphi_0(y, \varpi) \rightarrow \exists \varpi^\Omega \in A^c \forall x^0 \varphi_0(x, \varpi),$$

so that $\forall x^0 (\mathbb{P}[\varphi_0(x)] \geq \lambda)$ implies

$$\forall A^S (\mathbb{P}(A) <_{\mathbb{R}} \lambda \rightarrow \exists \varpi^\Omega \in A^c \forall x^0 \varphi_0(x, \varpi)),$$

i.e. $\mathbb{P}[\forall x^0 \varphi_0(x)] \geq \lambda$, as claimed. \square

While the restriction of the above result to quantifier-free formulas seems quite essential, the restriction to type 0 is entirely dependent on the fact that we only treated $\text{QF-CC}^{\Omega,0}$ and $\text{QF-CC}^{S,0}$. Indeed, it is easy to see that the above proof immediately generalizes whenever contra-collection of higher types is available. This would be an advantage of the bounded functional interpretation [12], where an analogous contra-collection principle is immediately available for arbitrary types without arduous pre- and post-processing, at the expense of using an intensional majorizability relation of course.

Remark 4.10. Such prenexation laws as the above are formulated mathematically, for general (measurable) formulas, for example in Lemma 3.1 in [50]. Concretely, let (Ω, S, \mathbb{P}) be a probability space and write $\forall n \varphi(n) = \bigcap_{n \in \mathbb{N}} \varphi(n)$ as well as $\exists n \varphi(n) = \bigcup_{n \in \mathbb{N}} \varphi(n)$ for any measurable φ , that is $\varphi(n) \in S$ for all $n \in \mathbb{N}$. Then, Lemma 3.1 in [50] proves semantically that

- (1) $\mathbb{P}(\forall n \varphi(n)) \geq p$ if and only if $\forall n (\mathbb{P}(\varphi(n)) \geq p)$,
- (2) $\mathbb{P}(\forall n \varphi(n)) \leq p$ if and only if $\forall \lambda > 0 \exists n (\mathbb{P}(\varphi(n)) < p + \lambda)$,

for any measurable anti-monotone φ and probability $p \in [0, 1]$, and

- (1) $\mathbb{P}(\exists n \varphi(n)) \leq p$ if and only if $\forall n (\mathbb{P}(\varphi(n)) \leq p)$,
- (2) $\mathbb{P}(\exists n \varphi(n)) \geq p$ if and only if $\forall \lambda > 0 \exists n (\mathbb{P}(\varphi(n)) > p - \lambda)$,

for any measurable and monotone φ and probability $p \in [0, 1]$. However, this result only holds when \mathbb{P} is a measure and is indeed established in [50] using limit theorems for \mathbb{P} , so that these principles are generally not true for contents. Even further, the laws formulated above are, in their generality, not even true for the outer content associated with a probability content. Strikingly, the previous propositions hence show that variants of such prenexation laws are admissible in proofs for the purpose of extracting computable information. As our systems are based on contents, the resulting consequences will, further, be valid for general contents. So, uniform boundedness allows one to falsely treat contents, to some degree, like a measure in proofs, all the while retaining meaningful bound extraction theorems with valid conclusions over content spaces. In these contexts, we are in particular not restricted to the type 0 in our quantifiers, and can even allow the use of uncountable quantifiers.

Remark 4.11. Note that out of the semi-constructive results above, only Proposition 4.8 requires the use of $\text{IP}_{\forall}^{\omega}$. Further, none of the results use AC or M^{ω} in any form. Also, none of the classical results use DC .

As a last result of the present section, we now provide a type of conservativity result for the systems considered here regarding the principle of σ -additivity by relating it, in a classical context, to the uniform boundedness principles. Here, σ -additivity concretely takes the form of

$$\mathbb{P} \left(\bigcup_{n \in \mathbb{N}} A_n \right) = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n),$$

where A_n are pairwise disjoint events. First note that, using the monotonicity of the series $\sum_{n=0}^l \mathbb{P}(A_n)$, the above statement is equivalent to the conjunction of the two principles

$$\forall l \in \mathbb{N} \left(\mathbb{P} \left(\bigcup_{n \in \mathbb{N}} A_n \right) \geq \sum_{n=0}^l \mathbb{P}(A_n) \right) \text{ and } \forall k \in \mathbb{N} \exists l \in \mathbb{N} \left(\mathbb{P} \left(\bigcup_{n \in \mathbb{N}} A_n \right) \leq \sum_{n=0}^l \mathbb{P}(A_n) + 2^{-k} \right).$$

To formulate this property formally, we first recall from [46] the following operation on terms of type $S(0)$ that allows for the implicit quantification over a disjoint countable family of sets:

Given $A^{S(0)}$, we set $(A\uparrow)_0 := A_0$ and

$$(A\uparrow)_{n+1} := A_{n+1} \cap \left(\bigcup_{i=0}^n A_i \right)^c.$$

This operation thus turns A , representing an arbitrary sequence of measurable sets, into a sequence of disjoint sets $A\uparrow$ with the same (partial) union(s). Further, if A was already a disjoint family, then it is left unchanged by the operation (see also the discussion in [46]).

The fact that ϖ is included in the union $\bigcup_{n \in \mathbb{N}} A_n$ can now be expressed using the outer content via the formula $\exists n^0 (\varpi \in A_n)$. The first inequality above then corresponds to the statement

$$\forall A^{S(0)} \forall l^0 \left(\mathbb{P} [\exists n^0 A_n] \geq \sum_{n=0}^l \mathbb{P}((A\uparrow)_n) \right)$$

while the second can be formally recognized via

$$\forall A^{S(0)} \forall k^0 \exists l^0 \left(\mathbb{P} [\exists n^0 A_n] \leq \sum_{n=0}^l \mathbb{P}((A\uparrow)_n) + 2^{-k} \right).$$

We denote the conjunction of these two principles by (σ -**additivity**). The key result regarding this principle is then the following:

Proposition 4.12. *The system $\mathcal{F}^\omega[\mathbb{P}] + \exists\text{-UB}^{\Omega,0}$ proves (σ -**additivity**).*

Proof. Abusing notation slightly, note that

$$\sum_{n=0}^l \mathbb{P}((A\uparrow)_n) =_{\mathbb{R}} \mathbb{P} \left(\bigcup_{n=0}^l A_n \right) \leq \mathbb{P}[\exists n \leq_0 l A_n]$$

is provable in $\mathcal{F}^\omega[\mathbb{P}]$ (recall Proposition 4.2), so that the first part of (σ -**additivity**) is similarly provable in $\mathcal{F}^\omega[\mathbb{P}]$, using Proposition 4.4 and that $\exists n \leq_0 l (\varpi \in A_n)$ implies $\exists n^0 (\varpi \in A_n)$. Hence, we turn to the second part of (σ -**additivity**). Let $A^{S(0)}$ and k^0 be given. The partial sums $\sum_{n=0}^l \mathbb{P}((A\uparrow)_n) =_{\mathbb{R}} \mathbb{P}(\bigcup_{n=0}^l A_n)$ now form a monotone sequence of nonnegative real numbers, bounded above by 1. Since $\mathcal{F}^\omega[\mathbb{P}]$ (already without DC) proves the Cauchy formulation of the monotone convergence theorem (see e.g. Proposition 4.5 in [46]), there exists an l^0 such that

$$\forall m^0 \left(\mathbb{P} \left(\bigcup_{n=0}^{l+m} A_n \right) \leq_{\mathbb{R}} \sum_{n=0}^l \mathbb{P}((A\uparrow)_n) + 2^{-k} \right).$$

Using Proposition 4.2, we in particular get

$$\forall m^0 \left(\mathbb{P}[\exists n \leq_0 (l+m) A_n] \leq \sum_{n=0}^l \mathbb{P}((A\uparrow)_n) + 2^{-k} \right).$$

Now part (2) of Proposition 4.7 yields (observing the required monotonicity property in m), that with $\exists\text{-UB}^{\Omega,0}$ we have

$$\mathbb{P}[\exists m^0 \exists n \leq_0 (l+m) A_n] \leq \sum_{n=0}^l \mathbb{P}((A\uparrow)_n) + 2^{-k}$$

which, combined with Proposition 4.4, since $\exists m^0 \exists n \leq_0 (l + m) (\varpi \in A_n)$ is equivalent to $\exists n^0 (\varpi \in A_n)$, yields

$$\mathbb{P}[\exists n^0 A_n] \leq \sum_{n=0}^l \mathbb{P}((A \uparrow)_n) + 2^{-k},$$

as required. \square

Proposition 4.12 in particular shows that any abstract use of σ -additivity, that is σ -additivity formulated for measurable sets via the abstract type S , is eliminated via a proof mining analysis at the expense of a simple functional iteration akin to the monotone convergence theorem. Naturally, for more complex measurable sets which are not immediately represented via terms of type S (i.e. via quantifier-free formulas), the resulting principle of σ -additivity might of course still carry a lot of computational strength, akin to the limit theorems utilized previously to interchange quantifiers with the (outer) content.

Nevertheless, this crucial observation that an *abstract* use as above remains computationally tame, finally provides a formal explanation and justification for the (up to this point only empirically observed) phenomenon that proof mining seems to eliminate such uses of this axiom in the course of an analysis. As such, it in particular formally illustrates that probability contents indeed seem to be the finitary core of probability theory through the lens of proof mining, as conjectured in [46] based on the case studies available at the time. In particular, in light of Theorem 2.9, we have that probabilistic theorems provably equivalent over $\mathcal{F}^\omega[\mathbb{P}] + (\sigma\text{-additivity})$ to $\forall\exists$ -statements remain true over finitely additive spaces.

5. LOGICAL METATHEOREMS FOR PROBABILISTIC THEOREMS

In this section, we illustrate how our formal translation based on the outer measure can be used to systematically associate computational content to probabilistic statements, and how this association can be used in combination with the logical machinery underlying the logical metatheorems for proof mining in probability theory due to the first and third author to extract that content from proofs.

Our discussion will primarily be focused on probabilistic $\forall\exists$ -theorems, that is theorems of the form

$$(*) \quad \text{“}\forall x^\tau \exists n^0 \varphi(x, n) \text{ with probability } \geq \lambda\text{”},$$

where, at first, we want to think of τ as a generic admissible type and φ as an arbitrary formula, although we later also discuss various special cases.

There are three different, but a priori equally natural, ways to express this property formally in our system, and we now want to motivate them in what follows. The most naive way is to simply wrap the $\forall\exists$ -statement in the outer content introduced earlier, leading to

$$(*)_1 \quad \mathbb{P}[\forall x^\tau \exists n^0 \varphi(x, n)] \geq \lambda.$$

From a purely mathematical perspective, at least over measure spaces when $\tau = 0$, and if the statement $\varphi(x, n)$ is measurable and anti-monotone in x , the leading universal quantifier of that statement can be prenexed out of the probability (recall Remark 4.10), leading to a statement of the form

$$(*)_2 \quad \forall x^\tau (\mathbb{P}[\exists n^0 \varphi(x, n)] \geq \lambda).$$

If, moreover, $\varphi(x, n)$ is assumed to be monotone in n , then this statement can be further prenexed at the expense of a probabilistic error, leading to a statement of the form

$$(*)_3 \quad \forall m^0 \forall x^\tau \exists n^0 (\mathbb{P}[\varphi(x, n)] \geq \lambda - 2^{-m}).$$

While thereby naturally related in the mathematical context of measures, for suitably monotone and measurable properties and sufficiently low types, these statements all provide, in one way or another, suitable formal representations of the previously mentioned theorems of the type (*).

Presented with three choices for a formal access towards probabilistic $\forall\exists$ -theorems, we are now faced with three immediate questions:

- (1) What is the computational content of theorems of the form $(*)_1 - (*)_3$ and when can it be extracted from corresponding (classical or constructive) proofs?
- (2) How are the statements $(*)_1 - (*)_3$ and their computational interpretations related, potentially irrespective of monotonicity assumptions as used above?
- (3) Which formulations naturally occur in practical case studies from probability theory and statistics?

The next three parts of the present section provide a comprehensive answer to the first question by presenting corresponding metatheorems on the extraction of bounds tailored to theorems of the form $(*)_1 - (*)_3$. In that way, we extend the existing arsenal of proof mining metatheorems with dedicated tools to handle all the different kinds of probabilistic theorems illustrated above. As in the work [46], and as already highlighted before, the abstract treatment of the sample and event spaces, and in particular their uniform majorizability, guarantees that the resulting bounds will be highly uniform, not depending on any parameters related to the underlying probability content space, such as its measure. In the classical case, φ will be restricted to be purely existential, while there are no such formula restrictions in the constructive case. In general, the type τ of x will be allowed to exceed 0, as long as it remains reasonably low (as is typical for proof mining metatheorems, recall Section 2). Throughout, we only consider the case of the outer content and high probability, as the case of the inner content and low probability can be treated using duality. Strikingly, the statements $(*)_1 - (*)_3$ exhibit very different quantitative behavior, illustrating how the transformations (based on the continuity of the measure) used above to motivate the various formulations carry heavy computational strength in general. As such, the corresponding metatheorems get progressively more difficult to establish, requiring more and more theory beyond [46]. In particular the second question on the relation between $(*)_1 - (*)_3$ in that way, turns out to be rather subtle and complex, so that the last part of the present section is dedicated solely to this discussion.

The third question will be addressed in the following and final Section 6, where various exemplary applications from mathematical practice are discussed, and where in particular these formulations and their computational content will be linked with the by now quite considerable range of quantitative theorems from probability theory and statistics derived via the proof mining program, as already briefly mentioned in the introduction.

5.1. Probabilistic theorems of the form $\forall x^\tau (\mathbb{P}[\exists n^0 \varphi(x, n)] \geq \lambda)$. We begin with probabilistic bound extraction metatheorems for statements of the form $(*)_2$. In both the classical and semi-constructive context, the results are rather immediate applications of the resulting classical and semi-constructive metatheorems presented in Section 2.

Let us consider first the classical setting, where $\varphi(x, n)$ is assumed to be an existential formula $\varphi_\exists(x, n)$.

Theorem 5.1. *In the context of the assumptions of Theorem 2.9, suppose that*

$$\mathcal{F}^\omega[\mathbb{P}] + \exists\text{-UB}^{\Omega, \omega} + \exists\text{-UB}^{S, \omega} \vdash \forall x^\tau (\mathbb{P}[\exists n^0 \varphi_\exists(x, n)] \geq \lambda),$$

where τ is admissible and $\varphi_\exists(x, n, \varpi)$ is an existential formula such that all internal quantifiers have admissible types. Then, from this proof one can extract a partial functional $\Phi : \mathbb{N} \times \mathcal{S}_\tau \rightarrow \mathbb{N}$

which is total and bar-recursively computable on $\mathbb{N} \times \mathcal{M}_{\hat{\tau}}$ and such that

$$\mathcal{S}^{\omega, \Omega, S} \models \forall m^0 \forall x^* \hat{\tau} \forall x \lesssim_{\tau} x^* (\mathbb{P}[\exists n \leq_0 \Phi(m, x^*) \varphi_{\exists}(x, n)] \geq \lambda - 2^{-m})$$

holds for $\mathcal{S}^{\omega, \Omega, S}$ defined via any non-empty set Ω , algebra $S \subseteq 2^{\Omega}$, and probability content \mathbb{P} on S (and with suitable interpretations of the additional constants).

The same remarks (1) – (5) as in Theorem 2.9 apply.

Proof. Expanding the abbreviation, we have that $\forall x^{\tau} (\mathbb{P}[\exists n^0 \varphi_{\exists}(x, n)] \geq \lambda)$ is equivalent to

$$\forall x^{\tau} \forall m^0 \forall A^S \exists n^0 (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists \varpi \in A^c \varphi_{\exists}(x, n, \varpi)).$$

As the inner formula is existential with suitable type restrictions, we can then apply Theorem 2.9 to extract a suitable partial functional $\Phi : \mathbb{N} \times \mathcal{S}_{\hat{\tau}} \rightarrow \mathbb{N}$ such that

$$\mathcal{S}^{\omega, \Omega, S} \models \forall m^0 \forall x^* \hat{\tau} \forall x \lesssim_{\tau} x^* \forall A^S \exists n \leq_0 \Phi(m, x^*) (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists \varpi \in A^c \varphi_{\exists}(x, n, \varpi))$$

for all suitable models $\mathcal{S}^{\omega, \Omega, S}$. This in particular implies

$$\mathcal{S}^{\omega, \Omega, S} \models \forall m^0 \forall x^* \hat{\tau} \forall x \lesssim_{\tau} x^* (\mathbb{P}[\exists n \leq_0 \Phi(m, x^*) \varphi_{\exists}(x, n)] \geq \lambda - 2^{-m})$$

as claimed. \square

Note that Theorem 5.1 also applies to statements of the form $(*)_1$, that is of the form $\mathbb{P}[\forall x^{\tau} \exists n^0 \varphi_{\exists}(x, n)] \geq \lambda$, since, by Proposition 4.6, this implies $\forall x^{\tau} (\mathbb{P}[\exists n^0 \varphi_{\exists}(x, n)] \geq \lambda)$ already over $\mathcal{F}^{\omega}[\mathbb{P}]$.

Let us now consider the case when statements of the form $(*)_2$ have been proved in a semi-constructive setting. As in Theorem 2.11, the semi-constructive context allows for greater flexibility in the assumptions.

Theorem 5.2. *In the context of the assumptions of Theorem 2.11, suppose that*

$$\mathcal{F}_i^{\omega}[\mathbb{P}] + \mathbf{UB}^{\Omega, \omega} + \mathbf{UB}^{S, \omega} \vdash \forall x^{\tau} (\mathbb{P}[\exists n^0 \varphi(x, n)] \geq \lambda),$$

where τ is admissible and $\varphi(x, n, \varpi)$ is any formula such that all positively occurring universal quantifiers and all negatively occurring existential quantifiers have small types and all other types are admissible.

Then one can extract a primitive-recursive functional $\Phi : \mathbb{N} \times \mathcal{S}_{\hat{\tau}} \rightarrow \mathbb{N}$ such that

$$\mathcal{S}^{\omega, \Omega, S} \models \forall m^0 \forall x^* \hat{\tau} \forall x \lesssim_{\tau} x^* (\mathbb{P}[\exists n \leq_0 \Phi(m, x^*) \varphi(x, n)] \geq \lambda - 2^{-m})$$

holds for $\mathcal{S}^{\omega, \Omega, S}$ defined via any non-empty set Ω , algebra $S \subseteq 2^{\Omega}$, and probability content \mathbb{P} on S (and with suitable interpretations of the additional constants).

If the theorem is established over $\mathcal{F}_i^{\omega}[\mathbb{P}] + \mathbf{UB}^{\Omega, \omega} + \mathbf{UB}^{S, \omega} \oplus \mathbf{QF}\text{-CC}^{\Omega, 0} \oplus \mathbf{QF}\text{-CC}^{S, 0}$ instead, then the result remains true if $\varphi(x, n, \varpi)$ is a quantifier-free formula.

We have the same remarks (1) – (3) as in Theorem 2.11.

Proof. As in the proof of Theorem 5.1, note that, in $\mathcal{F}_i^{\omega}[\mathbb{P}] + \mathbf{UB}^{\Omega, \omega} + \mathbf{UB}^{S, \omega}$, the statement $\forall x^{\tau} (\mathbb{P}[\exists n^0 \varphi(x, n)] \geq \lambda)$ is equivalent to

$$\forall x^{\tau} \forall m^0 \forall A^S \exists n^0 (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists \varpi \in A^c \varphi(x, n, \varpi)).$$

The first result now follows from Theorem 2.11 and the second result, regarding contraction, follows from Theorem 2.13. \square

As before, Theorem 5.2 also applies to statements of the form $(*)_1$, i.e. $\mathbb{P}[\forall x^{\tau} \exists n^0 \varphi(x, n)] \geq \lambda$, since, by Proposition 4.6, this provably implies $\forall x^{\tau} (\mathbb{P}[\exists n^0 \varphi(x, n)] \geq \lambda)$ even in a constructive setting.

5.2. Probabilistic theorems of the form $\forall m^0 \forall x^\tau \exists n^0 (\mathbb{P}[\varphi(x, n)] \geq \lambda - 2^{-m})$. We now move to statements of the form $(*)_3$, which, in general (without any monotonicity assumptions), are stronger than theorems of the form $(*)_2$ (recall Proposition 4.6). Nevertheless, they obey similarly nice bound extraction theorems, which are also readily established. However, in the classical case we do rely on epsilon-terms in the style of [17], and here provide a bit more detail than before, were they were only mentioned in passing. In that classical context, we then in particular assume that φ is quantifier-free, although this requirement can be somewhat relaxed.

Theorem 5.3. *In the context of the assumptions of Theorem 2.9, suppose that*

$$\mathcal{F}^\omega[\mathbb{P}] + \exists\text{-UB}^{\Omega, \omega} + \exists\text{-UB}^{S, \omega} \vdash \forall m^0 \forall x^\tau \exists n^0 (\mathbb{P}[\varphi_0(x, n)] \geq \lambda - 2^{-m}),$$

where τ is admissible and $\varphi_0(x, n, \varpi)$ is quantifier-free.

Then one can extract a partial functional $\Phi : \mathbb{N} \times \mathcal{S}_{\hat{\tau}} \rightarrow \mathbb{N}$ which is total and bar-recursively computable on $\mathbb{N} \times \mathcal{M}_{\hat{\tau}}$ and such that

$$\mathcal{S}^{\omega, \Omega, S} \models \forall m^0 \forall x^{*\hat{\tau}} \forall x \lesssim_\tau x^* \exists n \leq_0 \Phi(m, x^*) (\mathbb{P}[\varphi_0(x, n)] \geq \lambda - 2^{-m}),$$

holds for $\mathcal{S}^{\omega, \Omega, S}$ defined via any non-empty set Ω , algebra $S \subseteq 2^\Omega$, and probability content \mathbb{P} on S (and with suitable interpretations of the additional constants).

The same remarks (1) – (5) as in Theorem 2.9 apply.

Proof. The abbreviation $\forall m^0 \forall x^\tau \exists n^0 (\mathbb{P}[\varphi_0(x, n)] \geq \lambda - 2^{-m})$ is equivalent to

$$\forall x^\tau \forall m^0 \exists n^0 \forall A^S (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists \varpi \in A^c \varphi_0(x, n, \varpi)).$$

Using additional epsilon-terms (see e.g. [17]), we can “remove” the bounded quantifier of type S and regard the formula

$$\forall A^S (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists \varpi \in A^c \varphi_0(x, n, \varpi))$$

as purely existential, with respective type restrictions. Concretely, we add a constant $\chi^{S(0)(\tau)(0)(1)}$ to the language that is governed by the following axiom of type Δ :

$$(\chi) \quad \left\{ \begin{array}{l} \forall \lambda^1, m^0, x^\tau, n^0 \left((\mathbb{P}(\chi_{\lambda, m, x, n}) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists \varpi \in (\chi_{\lambda, m, x, n})^c \varphi_0(x, n, \varpi)) \right. \\ \left. \rightarrow \forall A^S (\mathbb{P}(A) <_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists \varpi \in A^c \varphi_0(x, n, \varpi)) \right) \end{array} \right.$$

The intended semantics of that constant is that $\chi_{\lambda, m, x, n} := A$ for A such that $\mathbb{P}(A) \leq \lambda - 2^{-m}$ and $\forall \varpi \in A^c \neg \varphi_0(x, n, \varpi)$, if existent, and $\chi_{\lambda, m, x, n} := \emptyset$ otherwise. In that way, the constant immediately satisfies the above axiom (for suitably large m), and as χ maps into the type S , it is immediately majorizable. In the theory extended with the constant χ and the axiom (χ) , we can then in particular derive

$$\forall x^\tau \forall m^0 \exists n^0 (\mathbb{P}(\chi_{\lambda, m, x, n}) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists \varpi \in (\chi_{\lambda, m, x, n})^c \varphi_0(x, n, \varpi)).$$

Applying Theorem 2.9 to that statement in this extension, we can extract a suitable partial functional $\Phi : \mathbb{N} \times \mathcal{S}_{\hat{\tau}} \rightarrow \mathbb{N}$ such that, using (χ) , we have

$$\mathcal{S}^{\omega, \Omega, S} \models \forall m^0 \forall x^{*\hat{\tau}} \forall x \lesssim_\tau x^* \exists n \leq_0 \Phi(m, x^*) \forall A^S (\mathbb{P}(A) <_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists \varpi \in A^c \varphi_0(x, n, \varpi))$$

for all suitable models $\mathcal{S}^{\omega, \Omega, S}$, which yields the claim. \square

Remark 5.4. In the context of Theorem 5.1, if $\varphi_\exists(x, n)$ is monotone in n then (by Proposition 4.4) the functional Φ would in fact be a witness for n , i.e.

$$\mathcal{S}^{\omega, \Omega, S} \models \forall m^0 \forall x^{*\hat{\tau}} \forall x \lesssim_\tau x^* (\mathbb{P}[\varphi_\exists(x, \Phi(m, x^*))] \geq \lambda - 2^{-m}),$$

which is even stronger than the conclusion of Theorem 5.3. In the context of Theorem 5.1, this result could then also have been established by first pulling the existential quantifier out of the outer content using uniform boundedness and Proposition 4.7.

We now move to the semi-constructive case. Here, the result is generally a bit more immediate and we have greater flexibility in the assumptions, as before.

Theorem 5.5. *In the context of the assumptions of Theorem 2.11, suppose that*

$$\mathcal{F}_i^\omega[\mathbb{P}] + \mathbf{UB}^{\Omega,\omega} + \mathbf{UB}^{S,\omega} \vdash \forall m^0 \forall x^\tau \exists n^0 (\mathbb{P}[\varphi(x, n)] \geq \lambda - 2^{-m}),$$

where τ is admissible and $\varphi(x, n, \varpi)$ is any formula such that all positively occurring universal quantifiers and all negatively occurring existential quantifiers have small types and all other types are admissible.

Then one can extract a primitive-recursive functional $\Phi : \mathbb{N} \times \mathcal{S}_\tau \rightarrow \mathbb{N}$ such that

$$\mathcal{S}^{\omega,\Omega,S} \models \forall m^0 \forall x^{*\hat{\tau}} \forall x \lesssim_\tau x^* \exists n \leq_0 \Phi(m, x^*) (\mathbb{P}[\varphi(x, n)] \geq \lambda - 2^{-m})$$

holds for $\mathcal{S}^{\omega,\Omega,S}$ defined via any non-empty set Ω , algebra $S \subseteq 2^\Omega$, and probability content \mathbb{P} on S (and with suitable interpretations of the additional constants).

If the theorem is established over $\mathcal{F}_i^\omega[\mathbb{P}] + \mathbf{UB}^{\Omega,\omega} + \mathbf{UB}^{S,\omega} \oplus \mathbf{QF}\text{-CC}^{\Omega,0} \oplus \mathbf{QF}\text{-CC}^{S,0}$ instead, then the result remains true if $\varphi(x, n, \varpi)$ is an existential formula with the respective type restrictions.

We have the same remarks (1) – (3) as in Theorem 2.11.

Proof. As in the proof of Theorem 5.3, note that, in $\mathcal{F}_i^\omega[\mathbb{P}] + \mathbf{UB}^{\Omega,\omega} + \mathbf{UB}^{S,\omega}$, the statement $\forall m^0 \forall x^\tau \exists n^0 (\mathbb{P}[\varphi(x, n)] \geq \lambda - 2^{-m})$ is equivalent to

$$\forall x^\tau \forall m^0 \exists n^0 \forall A^S (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists \varpi \in A^c \varphi(x, n, \varpi)).$$

The first result then immediately follows from Theorem 2.11. For the second result, we instead first require similar considerations on epsilon-terms as in the proof of Theorem 5.3, after which Theorem 2.13 then applies. \square

Remark 5.6. As in Remark 5.4, suppose in the context of Theorem 5.2 additionally that φ is monotone in n . Then, by Proposition 4.4, for that same functional Φ we in fact have

$$\mathcal{S}^{\omega,\Omega,S} \models \forall m^0 \forall x^{*\hat{\tau}} \forall x \lesssim_\tau x^* (\mathbb{P}[\varphi(x, \Phi(m, x^*))] \geq \lambda - 2^{-m}),$$

so that also here the monotonicity allows us to derive an even stronger result. As before, in the context of Theorem 5.2, this result could then also have been established by first pulling the existential quantifier out of the outer content using uniform boundedness, now as in Proposition 4.8.

5.3. Probabilistic theorems of the form $\mathbb{P}[\forall x^\tau \exists n^0 \varphi(x, n)] \geq \lambda$. We now consider theorems of the form $(*)_1$. This formulation of a probabilistic $\forall\exists$ -theorem actually proves to be quite subtle in various ways. Indeed, both the classical and semi-constructive metatheorems will prove to be considerably more difficult to establish than the previous results, with greatest difficulty in the classical case. In particular, we rely rather extensively on the use of uniform boundedness as well as epsilon-terms in the style of [17] and as the result seems to be rather sensitive, we here in particular are very explicit on the use of epsilon-terms, compared to some of the previous results, where they were only mentioned in passing or briefly sketched. With these difficulties also come various slightly more severe restrictions on the reach of the metatheorem. However, while discussed and exemplified in Section 6 later on, we want to already remark here that, currently, apart from one notable recent case study [49], no other use of formulations of the form $(*)_1$ has ever been observed in proof mining practice, with essentially all theorems

naturally confining to the forms $(*)_2$ and $(*)_3$. In that way, the resulting complications so far had a rather minimal impact on proof mining practice.

We now begin with the classical case. To simplify the presentation, we here restrict ourselves to $\tau \in \{0, 1\}$ and we assume that φ is quantifier-free, although both of these requirements can be relaxed in some ways.

Theorem 5.7. *In the context of the assumptions of Theorem 2.9, assume that*

$$\mathcal{F}^\omega[\mathbb{P}] + \exists\text{-UB}^{\Omega, \omega} + \exists\text{-UB}^{S, \omega} \vdash \mathbb{P}[\forall x^\tau \exists n^0 \varphi_0(x, n)] \geq \lambda,$$

where τ is of degree 1 and $\varphi_0(x, n, \varpi)$ is quantifier-free.

Then one can extract a total bar-recursive functional $\Phi : \mathbb{N} \times \mathcal{M}_{\tau(0(\tau))} \rightarrow \mathcal{M}_{0(\tau)}$ such that

$$\mathcal{M}^{\omega, \Omega, S} \models \forall m \forall x^* \forall \tilde{x} \lesssim x^* (\mathbb{P}[\exists N \lesssim \Phi(m, x^*) \forall x \leq \tilde{x}(N) \varphi_0(x, N(x))] \geq \lambda - 2^{-m})$$

holds for $\mathcal{M}^{\omega, \Omega, S}$ defined via any non-empty set Ω , algebra $S \subseteq 2^\Omega$, and probability content \mathbb{P} on S (and with suitable interpretations of the additional constants).

If $\tau = 0$, then the conclusion moreover holds in a model based on $\mathcal{S}^{\omega, \Omega, S}$, with Φ then being a partial function $\mathbb{N} \times \mathcal{S}_2 \rightarrow \mathcal{M}_1$.

We have the same remarks (2) – (4) as in Theorem 2.9.

Proof. We focus on the more complex case $\tau = 1$. Expanding the abbreviation, we have that $\mathcal{F}^\omega[\mathbb{P}] + \exists\text{-UB}^{\Omega, \omega} + \exists\text{-UB}^{S, \omega}$ proves

$$\forall m^0 \forall A^S (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists \varpi \in A^c \forall x^\tau \exists n^0 \varphi_0(x, n, \varpi)).$$

In particular, using QF-AC, the above is equivalent to

$$\forall m^0 \forall A^S (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists N^{0(\tau)} \exists \varpi \in A^c \forall x^\tau \varphi_0(x, Nx, \varpi)).$$

Using $\exists\text{-UB}^{\Omega, \tau}$ and classical logic, this is further equivalent to

$$\forall m^0 \forall A^S (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists N^{0(\tau)} \forall x^{*\tau} \exists \varpi \in A^c \forall x \leq_\tau x^* \varphi_0(x, Nx, \varpi)).$$

This yields

$$(\%) \quad \forall m^0 \forall A^S \exists N^{0(\tau)} \forall x^{*\tau} (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists \varpi \in A^c \forall x^\tau \varphi_0(\min_\tau(x, x^*), N \min_\tau(x, x^*), \varpi)).$$

We now add an epsilon-term $\phi^{1(0(1))^{(1)}}$ in the style of [17] to the language with the intended semantics that $\phi(y^1, z^{0(1)}) := x \leq_1 y$ for x such that $z(x) =_0 0$, if existent, and $\phi(y^1, z^{0(1)}) := \lambda n^0.0$ otherwise. This constant is governed by the following universal axiom:

$$(\phi) \quad \forall x^1, y^1, z^{0(1)} \left(z(\min_1(x, y)) =_0 0 \rightarrow z(\min_1(\phi(y, z), y)) =_0 0 \right).$$

The formal semantics, which later assures majorizability of that constant, is now as follows: over $\mathcal{S}^{\omega, \Omega, S}$, it is interpreted as

$$\phi(y^1, z^{0(1)}) :=_1 \begin{cases} \min_1(x, y) & \text{for } x^1 \text{ with } z(\min_1(x, y)) =_0 0 \text{ if existent,} \\ \lambda n^0.0 & \text{otherwise.} \end{cases}$$

As $\min_1(\min_1(x, y), y) = \min_1(x, y)$, this satisfies the axiom. Let now $t_{\neg\varphi}$ be the term such that

$$t_{\neg\varphi}(x, N, \varpi) =_0 0 \leftrightarrow \neg\varphi_0(x, Nx, \varpi).$$

Applying (ϕ) to $\lambda x^1.t_{\neg\varphi}(x, N, \varpi)$ and x^* yields

$$\begin{aligned} \varphi_0(\min_1(\phi(x^*, \lambda x^1.t_{\neg\varphi}(x, N, \varpi)), x^*), N \min_1(\phi(x^*, \lambda x^1.t_{\neg\varphi}(x, N, \varpi)), x^*), \varpi) \\ \rightarrow \varphi_0(\min_1(x, x^*), N \min_1(x, x^*), \varpi) \end{aligned}$$

so that

$$\begin{aligned} & \varphi_0(\min_1(\phi(x^*, \lambda x^1.t_{-\theta}(x, N, \varpi)), x^*), N \min_1(\phi(x^*, \lambda x^1.t_{-\theta}(x, N, \varpi)), x^*), \varpi) \\ & \leftrightarrow \forall x^1 \varphi_0(\min_1(x, x^*), N \min_1(x, x^*), \varpi). \end{aligned}$$

Using another “epsilon”-term, we hence get a quantifier-free formula $\theta_0(x^*, N)$ such that

$$\theta_0(x^*, N) \leftrightarrow \exists \varpi \in A^c \forall x^\tau \varphi_0(\min_\tau(x, x^*), N \min_\tau(x, x^*), \varpi).$$

The above formula ($\%_0$) is hence equivalent

$$\forall m^0 \forall A^S \exists N^{0(\tau)} \forall x^{*\tau} (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \theta_0(x^*, N))$$

and using QF-AC to

$$\forall m \forall A \forall \tilde{x}^* \exists N (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \theta_0(\tilde{x}^*(N), N)).$$

The (negative) functional interpretation (see Lemma 8.3 in [46]) combined with majorization in the model $\mathcal{M}^{\omega, \Omega, S}$ (see Lemma 8.8 in [46] and recall the proof of Theorem 2.9) now yields a functional Φ such that

$$\mathcal{M}^{\omega, \Omega, S} \models \forall m \forall A \forall x^* \forall \tilde{x} \lesssim x^* \exists N \lesssim \Phi(m, x^*) (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \theta_0(\tilde{x}(N), N))$$

or, expanded,

$$\begin{aligned} & \forall m \forall A \forall x^* \forall \tilde{x} \lesssim x^* \exists N \lesssim \Phi(m, x^*) (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \\ & \rightarrow \exists \varpi \in A^c \forall x \varphi_0(\min(x, \tilde{x}(N)), N(\min(x, \tilde{x}(N))), \varpi)) \end{aligned}$$

which is equivalent to

$$\forall m \forall x^* \forall \tilde{x} \lesssim x^* \mathbb{P}[\exists N \lesssim \Phi(m, x^*) \forall x \varphi_0(\min(x, \tilde{x}(N)), N(\min(x, \tilde{x}(N))))] \geq \lambda - 2^{-m}.$$

As we are working over the model $\mathcal{M}^{\omega, \Omega, S}$, we now get (using extensionality) that the above is equivalent to

$$\mathcal{M}^{\omega, \Omega, S} \models \forall m \forall x^* \forall \tilde{x} \lesssim x^* \mathbb{P}[\exists N \lesssim \Phi(m, x^*) \forall x \leq \tilde{x}(N) \varphi_0(x, N(x))] \geq \lambda - 2^{-m},$$

as claimed. That this reduces to truth in a model based on $\mathcal{S}^{\omega, \Omega, S}$ if $\tau = 0$ follows immediately as the types are then low-enough (see in particular the discussion in the proof of Theorem 17.52 in [31]). \square

Note that already for $\tau = 0$, the above modulus Φ is a functional of type 3, which is quite high. Indeed, such a modulus has never been observed in practice, as will also be discussed in more details in the next section.

As before, we now also consider the special case of semi-constructively proven theorems of the form $\mathbb{P}[\forall x^\tau \exists n^0 \varphi(x, n)] \geq \lambda$. Again, the corresponding metatheorem is harder to establish than that given in Theorem 5.2, though only slightly more difficult. In particular, we here rely on a model-theoretic result due to Kohlenbach [32].

Theorem 5.8. *In the context of the assumptions of Theorem 2.11, assume that*

$$\mathcal{F}_i^\omega[\mathbb{P}] + \text{UB}^{\Omega, \omega} + \text{UB}^{S, \omega} \vdash \mathbb{P}[\forall x^\tau \exists n^0 \varphi(x, n)] \geq \lambda,$$

where τ is admissible and $\varphi(x, n, \varpi)$ is any formula such that all positively occurring universal quantifiers and all negatively occurring existential quantifiers have small types and all other types are admissible.

Then one can extract a primitive-recursive functional $\Phi : \mathbb{N} \times \mathcal{S}_\tau \rightarrow \mathbb{N}$ such that

$$\mathcal{S}^{\omega, \Omega, S} \models \forall m^0 (\mathbb{P}[\forall x^* \forall x \lesssim_\tau x^* \exists n \leq_0 \Phi(m, x^*) \varphi(x, n)] \geq \lambda - 2^{-m})$$

holds for $\mathcal{S}^{\omega, \Omega, S}$ defined via any non-empty set Ω , algebra $S \subseteq 2^\Omega$, and probability content \mathbb{P} on S (and with suitable interpretations of the additional constants).

We have the same remarks (1) – (3) as in Theorem 2.11.

Proof. Expanding the abbreviation, we have that $\mathcal{F}_i^\omega[\mathbb{P}] + \mathbf{UB}^{\Omega, \omega} + \mathbf{UB}^{S, \omega}$ proves

$$\forall m^0 \forall A^S (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists \varpi \in A^c \forall x^\tau \exists n^0 \varphi(x, n, \varpi)).$$

Using AC, we get

$$\forall m^0 \forall A^S (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists N^{0(\tau)} \exists \varpi \in A^c \forall x^\tau \varphi(x, Nx, \varpi)).$$

The principle $\mathbf{IP}_{\forall}^\omega$ yields

$$\forall m^0 \forall A^S \exists N^{0(\tau)} (\mathbb{P}(A) \leq_{\mathbb{R}} \lambda - 2^{-m} \rightarrow \exists \varpi \in A^c \forall x^\tau \varphi(x, Nx, \varpi)).$$

Using the (negative) functional interpretation (see again Lemma 8.3 in [46]) combined with majorization in the model $\mathcal{M}^{\omega, \Omega, S}$ (see again Lemma 8.8 in [46] and Theorem 2.9), we get a majorizable functional $\Phi : \mathbb{N} \rightarrow \mathcal{M}_{0(\hat{\tau})}$ such that:

$$\mathcal{M}^{\omega, \Omega, S} \models \forall m^0 (\mathbb{P}[\exists N \lesssim_{0(\tau)} \Phi(m) \forall x^\tau \varphi(x, Nx, \varpi)] \geq \lambda - 2^{-m}).$$

By Theorem 2.5 from [32], the above is equivalent to

$$\mathcal{M}^{\omega, \Omega, S} \models \forall m^0 \left(\mathbb{P}[\forall x^{*\hat{\tau}} \forall x \lesssim_\tau x^* \exists n \leq_0 \Phi(m, x^*) \varphi(x, n, \varpi)] \geq \lambda - 2^{-m} \right).$$

As τ is admissible, we get that the resulting claim is also true in $\mathcal{S}^{\omega, \Omega, S}$ (recall again the discussion in the proof of Theorem 17.52 in [31]). \square

5.4. Relationships between $(*)_1$, $(*)_2$ and $(*)_3$. We now want to comment on the relationship between the statements $(*)_1 - (*)_3$, and their respective computational interpretations, under various natural assumptions on the types and associated problem matrices. First, let us recall the various types of computational content that these different statements generate:

The computational content of statements of the form $\mathbb{P}[\forall x^\tau \exists n^0 \varphi(x, n)] \geq \lambda$, that is of the form $(*)_1$, was dependent in form already on whether this statement is considered in a classical or semi-constructive context, even when the internal matrix φ is quantifier-free. Concretely, in a semi-constructive context this content takes the form of a functional Φ such that

$$(+)_1^i \quad \forall m^0 \left(\mathbb{P}[\forall x^* \forall x \lesssim_\tau x^* \exists n \leq_0 \Phi(m, x^*) \varphi(x, n)] \geq \lambda - 2^{-m} \right),$$

while, in a classical context, this content takes the form of a functional Φ such that

$$(+)_1^c \quad \forall m \forall x^* \forall \tilde{x} \lesssim x^* \left(\mathbb{P}[\exists N \lesssim \Phi(m, x^*) \forall x \leq \tilde{x}(N) \varphi_0(x, N(x))] \geq \lambda - 2^{-m} \right).$$

The associated statement of the form $(*)_2$ where the universal quantifier is prenexed, that is $\forall x^\tau (\mathbb{P}[\exists n^0 \varphi(x, n)] \geq \lambda)$, was not so sensitive in form, resulting generally (that is classically or semi-constructively) in functionals Φ such that

$$(+)_2 \quad \forall m^0 \forall x^* \forall x \lesssim_\tau x^* \left(\mathbb{P}[\exists n \leq_0 \Phi(m, x^*) \varphi(x, n)] \geq \lambda - 2^{-m} \right).$$

Lastly, if further prenexed appropriately by also drawing out the existential quantifier, at the expense of a probabilistic error, that is moving to the statement $\forall m^0 \forall x^\tau \exists n^0 (\mathbb{P}[\varphi(x, n)] \geq \lambda - 2^{-m})$ of the form $(*)_3$, we obtain a functional Φ such that

$$(+)_3 \quad \forall m^0 \forall x^* \forall x \lesssim_\tau x^* \exists n \leq_0 \Phi(m, x^*) \left(\mathbb{P}[\varphi(x, n)] \geq \lambda - 2^{-m} \right).$$

The relation between the different problem formulations and their resulting computational interpretations proves to be quite subtle, so that instead of providing a comprehensive answer, we are restricted to a collection of specialized observations.

First, note that by Proposition 4.6 both $(*)_1$ and $(*)_3$ imply $(*)_2$, without any additional assumptions on the matrix φ or the type τ . In particular, the resulting interpretations $(+)_1^c$ and $(+)_3$ primitive recursively imply $(+)_2$ (as is perhaps apparent from their formulation already). Further, intuitionistically $(*)_1$ implies its own negative translation, so that $(+)_1^i$ primitive recursively implies $(+)_1^c$. Putting these relations into a picture, we get the following overview:

$$\begin{array}{ccc}
 (*_3) & \searrow & (+)_3 \\
 & & \searrow \\
 & (*_2) & \searrow \\
 (*_1) & \nearrow & (+)_1^i \\
 & & \nearrow \\
 & & (+)_1^c
 \end{array}
 \quad \text{viz.} \quad
 \begin{array}{ccc}
 (+)_3 & \searrow & (+)_2 \\
 & & \searrow \\
 & & \searrow \\
 (+)_1^i & \nearrow & (+)_1^c
 \end{array}$$

In other words, pulling the first universal quantifier out of the (outer) measure makes the problem considerably weaker, while pulling the second existential quantifier out too, with a probabilistic error, strengthens the problem again. In particular, we currently know of *no* relation between $(*)_1$ and $(*)_3$ (viz. $(+)_1^c$, $(+)_1^i$ and $(+)_3$) and it seems that no further relations can be directly inferred without further assumptions.

Now, if $\varphi(x, n, \varpi)$ is both monotone in n and quantifier-free, then Propositions 4.7 and 4.8 guarantee that $(*)_2$ implies back $(*)_3$, a formal version of their mathematical equivalence for measures (recall Remark 4.10), established, however, only under the set-theoretically false principle of uniform boundedness, which nevertheless has no impact on bound extractions so that in that context, both statements can be used interchangeably.

An immediate question is whether this equivalence somehow extends to more complex matrices. Here, we can give an answer at least for the nearest case of (measurable and suitably monotone) Π_1^0 -matrices, by generalizing a mathematically motivated construction of Avigad, Dean and Rute [2] to a more general context. In [2], the authors show how specific modes of finitary probabilistic convergence relate to each other, corresponding to specific instantiations of $(*)_2$ and $(*)_3$ if seen in the context of the present paper, and by that give a quantitative variant of Egorov's theorem [9] on the equivalence between pointwise and uniform almost-sure convergence of random variables, as will also be discussed later in Section 6. This result was formally recognized in the context of the present systems in [46], in which it was in particular observed that the corresponding result already holds over contents. While all of this was tailored to specific modes of convergence, we can now actually recognize this formal representation of the results from [2] given in [46] as the blueprint of a general construction for the equivalence of $(*)_2$ and $(*)_3$ for Π_1^0 -matrices, beyond convergence.

As hinted on before, the key assumption for this will be the measurability of the matrix. Concretely, we will in the following assume that $\varphi(x^0, n^0, \varpi^\Omega) := \forall k^0 \varphi_0(x, n, k, \varpi)$ and that $\varphi_0(x, n, k, \varpi)$ is quantifier-free and measurable in the sense that there is a term $P_{x,n,k}$ of type $S(0)(0)(0)$ such that

$$(M) \quad \forall x^0 \forall n^0 \forall k^0 \forall \varpi^\Omega (\varpi \in P_{x,n,k} \leftrightarrow \varphi_0(x, n, k, \varpi))$$

is provable in the system.⁶ Then, instead of working with the formula φ_0 , the construction given in [2] and further abstracted in [46] relies on working with the sets $P_{x,n,m}$ and their probabilities more abstractly. The key result for that is the following combinatorial result established in [2] and then formalized over $\mathcal{F}^\omega[\mathbb{P}]$ in [46]:

⁶As such, this can immediately be ensured formally by adding a corresponding (majorizable) constant P of the respective type together with the above principle as an additional (universal) axiom.

Lemma 5.9 (Theorem 9.7 in [46]). *The system $\mathcal{F}^\omega[\mathbb{P}]$ proves:*

$$\forall A^{S(0)}, N^{0(1)}, u^0, v^0 >_0 \ u \exists i^0 \left(\forall g^1 \left(\mathbb{P} \left(\bigcap_{l=0}^{N(g)} \bigcup_{j=l}^{g(l)} A_j \right) \leq_{\mathbb{R}} 2^{-v} \right) \rightarrow \mathbb{P}(A_i) <_{\mathbb{R}} 2^{-u} \right).$$

This result is far from trivial and is in particular the source of the computational complexity of the resulting quantitative variant of our result, as it crucially relies on $\Pi_1^0\text{-AC}$, which is contained in DC and hence in $\mathcal{F}^\omega[\mathbb{P}]$.

The result we then get is the following, which closely follows the proof strategy of Theorem 9.9 in [46].

Proposition 5.10. *Let $\varphi_0(x^0, n^0, k^0, \varpi^\Omega)$ be measurable and quantifier-free, and monotone in n as well as anti-monotone in k . Over $\mathcal{F}^\omega[\mathbb{P}] + \exists\text{-UB}^{\Omega,0} + \exists\text{-UB}^{S,0}$, the following are equivalent:*

- (1) $\forall x^0 (\mathbb{P}[\exists n^0 \forall k^0 \varphi_0(x, n, k)] \geq 1)$,
- (2) $\forall m^0 \forall x^0 \exists n^0 (\mathbb{P}[\forall k^0 \varphi_0(x, n, k)] \geq 1 - 2^{-m})$.

Proof. By Proposition 4.6, (2) implies (1). To see the converse, consider (1), which by QF-AC is equivalent to $\forall x^0 (\mathbb{P}[\forall g^1 \exists n^0 \varphi_0(x, n, g(n))] \geq 1)$ and so in particular, by Proposition 4.6, implies

$$\forall x^0 \forall g^1 (\mathbb{P}[\exists n^0 \varphi_0(x, n, g(n))] \geq 1).$$

By $\exists\text{-UB}^{S,0}$, we get

$$\forall m^0 \forall x^0 \forall g^1 \exists n^0 (\mathbb{P}[\exists i \leq_0 n \varphi_0(x, i, g(i))] \geq 1 - 2^{-m}).$$

As φ_0 is measurable by (M), so is $\exists i \leq_0 n \varphi_0(x, i, g(i))$ and so we get

$$\forall m^0 \forall x^0 \forall g^1 \exists n^0 \left(\mathbb{P} \left(\bigcup_{i=0}^n P_{x,i,g(i)} \right) \geq_{\mathbb{R}} 1 - 2^{-m} \right)$$

by Proposition 4.2. Written in its contrapositive, we get we get

$$\forall m^0 \forall x^0 \forall g^1 \exists n^0 \left(\mathbb{P} \left(\bigcap_{i=0}^n P_{x,i,g(i)}^c \right) \leq_{\mathbb{R}} 2^{-m} \right).$$

Hence, using $\Pi_1^0\text{-AC}$, we get that there exists a function N with

$$(\dagger) \quad \forall m^0 \forall x^0 \forall g^1 \left(\mathbb{P} \left(\bigcap_{i=0}^{N_{m,x}(g)} P_{x,i,g(i)}^c \right) \leq_{\mathbb{R}} 2^{-m} \right).$$

Now consider (2). By Proposition 4.7 and using $\exists\text{-UB}^{\Omega,0}$, this is equivalent to

$$\forall m^0 \forall x^0 \exists n^0 \forall k^0 (\mathbb{P}[\varphi_0(x, n, k)] \geq 1 - 2^{-m}).$$

For a contradiction, suppose that (2) fails, so that in particular, also using (M) and Proposition 4.2, there are m_0 and x_0 such that $\forall n^0 \exists k^0 (\mathbb{P}(P_{x_0,n,k}^c) >_{\mathbb{R}} 2^{-m_0})$. Using $\Pi_1^0\text{-AC}$, we get a function g_0 such that

$$\forall n^0 (\mathbb{P}(P_{x_0,n,g_0(n)}^c) >_{\mathbb{R}} 2^{-m_0}).$$

Now, define $A_i := P_{x_0,i,g_0(i)}^c$. Then, using the monotonicity of φ_0 , and hence of P^c , we have

$$\bigcup_{j=l}^{g(l)} A_j = \bigcup_{j=l}^{g(l)} P_{x_0,j,g_0(j)}^c \subseteq P_{x_0,l,\tilde{g}_0(g(l))}^c$$

for all l and g , where $\tilde{g}_0(i) := \max_{j \leq i} g_0(j)$. In particular, this implies

$$\forall g^1 \left(\bigcap_{l=0}^{N'_{m_0+1, x_0}(g)} \bigcup_{j=l}^{N'_{m_0+1, x_0}(g)} A_j \subseteq \bigcap_{l=0}^{N'_{m_0+1, x_0}(g)} P_{x_0, l, \tilde{g}_0(g(l))}^c \right),$$

where $N'_{m_0+1, x_0}(g) := N_{m_0+1, x_0}(\lambda n \cdot \tilde{g}_0(g(n)))$. By (\dagger) and the monotonicity of \mathbb{P} , we get

$$\forall g^1 \left(\mathbb{P} \left(\bigcap_{l=0}^{N'_{m_0+1, x_0}(g)} \bigcup_{j=l}^{N'_{m_0+1, x_0}(g)} A_j \right) \leq_{\mathbb{R}} 2^{-(m_0+1)} \right),$$

so that Lemma 5.9 implies that there exists an i such that $\mathbb{P}(P_{x_0, i, g_0(i)}^c) =_{\mathbb{R}} \mathbb{P}(A_i) <_{\mathbb{R}} 2^{-m_0}$, which is a contradiction. \square

Note by inspection of the above proof that we actually have established that the implication

$$\left\{ \begin{array}{l} \forall m^0 \forall x^0 \forall g^1 \exists n^0 (\mathbb{P}[\exists i \leq_0 n \varphi_0(x, i, g(i))] \geq 1 - 2^{-m}) \\ \rightarrow \forall m^0 \forall x^0 \exists n^0 \forall k^0 (\mathbb{P}[\varphi_0(x, n, k)] \geq 1 - 2^{-m}) \end{array} \right.$$

already holds over $\mathcal{F}^\omega[\mathbb{P}]$, that is without any uniform boundedness principles. This is essentially what was established in [46] for the special case of a convergence statement and what was presented in a quantitative form in the paper [2]. In particular, even this result crucially relies on Π_1^0 -AC, so that we can here only guarantee that there is a bar-recursive witness for the above implication (and in that way, essentially, mapping solutions of $(+)_2$ to solutions of $(+)_3$) for measurable and suitably monotone Π_1^0 -matrices. Moreover, such a witness, for this generalized result, can be explicitly constructed by more or less directly following the arguments of [2], although we do not explore this here any further. We know of no formal relationship beyond Π_1^0 -matrices.

Moving to the other branch of the picture, we know a bit less. Recall that for quantifier-free and anti-monotone formulas $\psi(x^0, \varpi^\Omega)$ (not necessarily measurable), we provably have that $\mathbb{P}[\forall x^0 \psi(x)] \geq \lambda$ is equivalent to $\forall x^0 \mathbb{P}[\psi(x)] \geq \lambda$, which (in the classical context) even extends to x of higher type (recall Propositions 4.7 and 4.9).

The question on the formal relationship between $(*)_2$ and $(*)_1$ thereby reduces to the question if, and how, this formal equivalence extends to formulas which are not decidable anymore but have the next highest complexity, that is Σ_1^0 -matrices. The key result in that direction is the following, relating the respective interpretations $(+)_2$ and $(+)_1^i$ (and by that also $(+)_1^c$) with each other.

Lemma 5.11. *Let $\varphi_0(x^0, n^0, \varpi^\Omega)$ be measurable and quantifier-free. Over $\mathcal{F}^\omega[\mathbb{P}] + (\sigma\text{-additivity})$, we have the following implication: For any functional $\Phi(m, x)$ with*

$$\forall m^0 \forall x^0 (\mathbb{P}[\exists n \leq_0 \Phi(m, x) \varphi_0(x, n)] \geq 1 - 2^{-m}),$$

the functional $\Phi'(m, x) := \Phi(m + x + 1, x)$ satisfies

$$\forall m^0 (\mathbb{P}[\forall x^0 \exists n \leq_0 \Phi'(m, x) \varphi_0(x, n)] \geq 1 - 2^{-m}).$$

Proof. Fix m^0 . Note that $\forall n \leq_0 \Phi(m + x + 1, x) \neg \varphi_0(x, n)$ is provably measurable in $\mathcal{F}^\omega[\mathbb{P}]$, and (as an abuse of notation) we in the following simply write the above expression for that set.

Using (σ -additivity), for any k^0 there exists an l^0 such that (again, slightly abusing notation):

$$\begin{aligned} \mathbb{P} [\exists x^0 \forall n \leq_0 \Phi(m+x+1, x) \neg \varphi_0(x, n)] &\leq \sum_{x=0}^{l^0} \mathbb{P} (\forall n \leq_0 \Phi(m+x+1, x) \neg \varphi_0(x, n)) + 2^{-k} \\ &\leq_{\mathbb{R}} \sum_{x=0}^{l^0} 2^{-m-(x+1)} + 2^{-k} \leq_{\mathbb{R}} 2^{-m} + 2^{-k}. \end{aligned}$$

Since k^0 was arbitrary, we have

$$\mathbb{P} [\exists x^0 \forall n \leq_0 \Phi(m+x+1, x) \neg \varphi_0(x, n)] \leq 2^{-m}$$

which yields the result using duality. \square

Note that through Proposition 4.12, the above result in particular holds over $\mathcal{F}^\omega[\mathbb{P}] + \exists\text{-UB}^{\Omega,0}$ and so can be used in the context of bound extraction theorems without affecting extracted bounds. Regardless of that, we want to further emphasize that Lemma 5.11 is actually mathematically true for probability spaces (Ω, S, \mathbb{P}) .

Further, as statements of the form $(*)_2$ can be formally related to their quantitative variants $(+)_2$ using uniform boundedness, we obtain the following equivalence:

Proposition 5.12. *Let $\varphi_0(x^0, n^0, \varpi^\Omega)$ be measurable and quantifier-free. Over $\mathcal{F}^\omega[\mathbb{P}] + \exists\text{-UB}^{\Omega,0} + \exists\text{-UB}^{S,0}$, the following are equivalent:*

- (1) $\mathbb{P}[\forall x^0 \exists n^0 \varphi_0(x, n)] \geq 1$,
- (2) $\forall x^0 (\mathbb{P}[\exists n^0 \varphi_0(x, n)] \geq 1)$.

Proof. By Proposition 4.6, we only have to establish the direction from (2) to (1). We first show that over $\mathcal{F}^\omega[\mathbb{P}] + \exists\text{-UB}^{S,0}$, (2) is equivalent to

$$\exists \Phi^{0(0)(0)} \forall m^0 \forall x^0 (\mathbb{P}[\exists n \leq_0 \Phi(m, x) \varphi_0(x, n)] \geq 1 - 2^{-m}).$$

To see this, note that $\forall x^0 (\mathbb{P}[\exists n^0 \varphi_0(x, n)] \geq 1)$ is equivalent to

$$\forall m^0 \forall x^0 \forall A^S \exists n^0 (\mathbb{P}(A) \leq_{\mathbb{R}} 1 - 2^{-m} \rightarrow \exists \varpi \in A^c \varphi_0(x, n, \varpi)).$$

Using $\exists\text{-UB}^{S,0}$, we get that

$$\exists \Phi^{0(0)(0)} \forall m^0 \forall x^0 \forall A^S \exists n \leq_0 \Phi(m, x) (\mathbb{P}(A) \leq_{\mathbb{R}} 1 - 2^{-m} \rightarrow \exists \varpi \in A^c \varphi_0(x, n, \varpi))$$

which amounts to the above. Hence, over $\mathcal{F}^\omega[\mathbb{P}] + \exists\text{-UB}^{\Omega,0} + \exists\text{-UB}^{S,0}$, we can now proceed as follows: Using Lemma 5.11 (and Proposition 4.12), (2) now implies

$$\exists \Phi^{0(0)(0)} \forall m^0 (\mathbb{P}[\forall x^0 \exists n \leq_0 \Phi'(m, x) \varphi_0(x, n)] \geq 1 - 2^{-m}).$$

Dropping the quantitative information, we obtain $\forall m^0 (\mathbb{P}[\forall x^0 \exists n^0 \varphi_0(x, n)] \geq 1 - 2^{-m})$, which in turn yields (1). \square

Based on the use of the above mathematical result from Lemma 5.11, we are here restricted to quantifiers of type 0. It would be very interesting to see whether the above can be extended to e.g. type 1 using continuity arguments, but we leave the investigation of this for future work.

In general, we see already that the formula structure and its interaction with \mathbb{P} has a subtle impact on the computational strength of a statement, and that realizing the previously mentioned mathematical equivalences over measures can be far from computationally trivial. However, perhaps surprisingly, in the context of proof mining practice, these issues rarely surface, if at all. Indeed, as will be discussed in detail in the following Section 6, essentially all statements ever considered in the context of proof mining case studies are naturally formalized

via statements of the form $(*)_2$ or $(*)_3$, by virtue of most proofs of theorems of the form $(*)$ actually being of a sufficiently “local” nature that keeps universal parameters fixed throughout the argument.

As such, essentially all moduli ever considered in the context of proof mining practice have been of the form $(+)_2$ or $(+)_3$. That means that neither are these stronger functionals extracted in case studies nor are they commonly required to resolve an implication where a statement like $(*)_1$ features in the premise. Indeed, only one recent case study [49] involved such a modulus, as will be discussed in detail later.

In any case, if such formulations of probabilistic theorems will be required in the future, the present section provided the necessary logical tools and remarks on their relation to deal with them in practice.

6. EXAMPLES FROM MATHEMATICAL PRACTICE

We now illustrate the rather wide applicability of our probabilistic logical metatheorems by giving a range of examples from mathematical practice where they yield exactly the type of computational content expected and observed in applications of proof mining to probability theory. We focus on some key types of theorems, covering central examples. Nevertheless, also essentially all other (perhaps more minor) quantitative notions considered in the context of proof mining case studies in probability can similarly be systematically justified from a proof-theoretic perspective via the present paper, including in particular notions like moduli of almost sure finiteness or tightness as considered in [50], moduli of absolute continuity as considered in [54] and the notion and related theory of finitary martingales developed in [49]. It should be noted that for all the quantitative notions discussed in this section, the present paper in particular illustrates for the first time that these notions arise naturally through a systematic proof-theoretic approach towards assigning computational content to (classical or semi-constructive) almost sure statements, and in that way provides the first formal proof-theoretic justification for their extraction in the various case studies mentioned throughout. In this section, if not stated otherwise, all mathematical results are formulated over a respective probability content space $(\Omega, \mathcal{S}, \mathbb{P})$.

6.1. A_n infinitely often almost surely. Let $(A_n)_{n \in \mathbb{N}}$ be a sequence of events. The fact that ϖ belongs to the event $\limsup_{n \rightarrow \infty} A_n := \bigcap_{k \in \mathbb{N}} \bigcup_{n \geq k} A_n$ can be expressed via the formula

$$\forall k^0 \exists n^0 \exists i^0 \in [k; n](\varpi \in A_i).$$

We intentionally introduced the bounded quantification $\exists i^0 \in [k; n]$ here in order to make the formula $\exists i^0 \in [k; n](\varpi \in A_i)$ monotone in n and anti-monotone in k . The statement that A_n occurs *infinitely often almost surely*, that is $\mathbb{P}(\limsup_{n \rightarrow \infty} A_n) = 1$, can now be represented formally in different ways in our setting.

The representation with the weakest computational interpretation is given by the statement

$$\forall k^0 (\mathbb{P}[\exists n^0 \exists i^0 \in [k; n] A_i] \geq 1),$$

that is a statement of the form $(*)_2$. By Theorems 5.1 and 5.2, from a (classical or semi-constructive) proof of this, we are able to extract a modulus Φ such that

$$\forall m \in \mathbb{N} \forall k \in \mathbb{N} \left(\mathbb{P} \left(\bigcup_{i \in [k; \Phi(m, k)]} A_i \right) \geq 1 - 2^{-m} \right).$$

We call such Φ a *pointwise modulus of A_n infinitely often almost surely*. Such a modulus has been used in [1, Theorem 2.2] in connection with the quantitative interpretation of the second

Borel-Cantelli lemma or in the context of stochastic optimization [48, 49, 51]. Moreover, note that since $\exists i^0 \in [k; n](\varpi \in A_i)$ is anti-monotone in k , this just coincides with the quantitative interpretation of the statement

$$\forall m^0 \forall k^0 \exists n^0 (\mathbb{P}[\exists i^0 \in [k; n] A_i] \geq 1 - 2^{-m}),$$

that is a corresponding statement of the form $(*)_3$.

Now, if we instead were to consider a representation via

$$\mathbb{P}[\forall k^0 \exists n^0 \exists i^0 \in [k; n] A_i] \geq 1,$$

that is a statement of the form $(*)_1$, then we get quite different computational interpretations based on whether we are in a classical or semi-constructive context, as illustrated in the previous Section 5 (recall Theorems 5.7 and 5.8). In particular, in a semi-constructive context, as illustrated in Theorem 5.8, the resulting computational interpretation is given via a functional Φ satisfying

$$\forall m \in \mathbb{N} \left(\mathbb{P} \left[\bigcap_{k \in \mathbb{N}} \bigcup_{i \in [k; \Phi(m, k)]} A_i \right] \geq 1 - 2^{-m} \right).$$

Compared to the above modulus, which is of a more pointwise nature, we will call such a modulus Φ a *uniform modulus of A_n infinitely often almost surely*, which by the aforementioned theorem can be extracted from large classes of semi-constructive proofs of the above property. Such a modulus seems to be needed only rarely. Indeed, there is so far only one case study that actually relied on a modulus with this particularly strong property, given recently in [49] (for a specific choice of sets A_i arising in the context of stochastic optimization, see Definition 4.3 therein, that we do not discuss here any further). Indeed, therein we find a critical mismatch between two parts of a proof usually bridged by corresponding limit theorems for the measure, with the first part of the proof naturally establishing a theorem of the form $\forall k^0 (\mathbb{P}[\exists n^0 \exists i^0 \in [k; n] A_i] \geq 1)$ and the latter part of the proof crucially relying on formulating this theorem in the form of $\mathbb{P}[\forall k^0 \exists n^0 \exists i^0 \in [k; n] A_i] \geq 1$. Even though there generally might be no immediate quantitative relation between statements of the form $(*)_2$ and $(*)_1$ as discussed in the previous Section 5, as the types here were low enough, this gap in the particular case study [49] could be bridged by Lemma 5.11, by which, over a measure space, a solution to the latter is definable from a solution to the former.

We know of no other case where such an argument was ever needed. Ultimately, this might also be the reason why the classical interpretation of the above as given in Theorem 5.7, which is actually represented via a functional of type 3, might never have been observed in mathematical practice before, as in cases such as the above it is simply subsumed by this stronger property over measure spaces.

6.2. A_n infinitely often almost never. Dually to the above, we can also consider the statement that A_n occurs *infinitely often almost never*, that is $\mathbb{P}(\limsup_{n \rightarrow \infty} A_n) = 0$. By following the above approach, we can represent this statement via duality through

$$\mathbb{P}[\forall k^0 \exists n^0 \exists i^0 \in [k; n] A_i] \leq 0 \quad \equiv \quad \mathbb{P}[\exists k^0 \forall n^0 \forall i^0 \in [k; n] A_i^c] \geq 1.$$

As the inner matrix of the outer measure above is Σ_2^0 , it will make a difference whether this statement is considered in a classical or a semi-constructive context.

In a semi-constructive context, as illustrated by Theorem 5.2, the computational content of the above statement is given by a function Ψ such that

$$\forall m \in \mathbb{N} \left(\mathbb{P} \left[\bigcup_{i \geq \Psi(m)} A_i \right] \leq 2^{-m} \right).$$

Over measure spaces, rewriting the infinite union as a sequence of finite unions, we in particular get (recall Remark 4.10) that the above is equivalent to

$$\forall m \in \mathbb{N} \forall n \in \mathbb{N} \left(\mathbb{P} \left(\bigcup_{i \in [\Psi(m); n]} A_i \right) \leq 2^{-m} \right),$$

which has the benefit of being well-defined and suitable also for contents. We will call such a Ψ a *modulus of A_n infinitely often almost never*. This modulus has been used in [1, Theorem 2.1] in connection with the quantitative interpretation of the first Borel-Cantelli lemma, but generally, by the aforementioned theorem, can be extracted from large classes of semi-constructive proofs of the above property.

In a classical context, however, one would have to first transform this statement into a $\forall\exists$ -form in the style of the negative translation of the respective statement. This turns $\exists k^0 \forall n^0 \forall i^0 \in [k; n] (\varpi \in A_i^c)$ into $\forall g^1 \exists k^0 \forall i^0 \in [k; g(k)] (\varpi \in A_i^c)$. Therefore, there are various options for representing the resulting probabilistic statement formally. Here, we now just focus on the representation via

$$\forall m^0 \forall g^1 \exists k^0 (\mathbb{P}[\forall i^0 \in [k; g(k)] A_i^c] \geq 1 - 2^{-m}),$$

that is in the form of $(*)_3$, as the other two immediate representations of the associated property, that is $\forall g^1 (\mathbb{P}[\exists k^0 \forall i^0 \in [k; g(k)] A_i^c] \geq 1)$ in the form of $(*)_2$ and $\mathbb{P}[\forall g^1 \exists k^0 \forall i^0 \in [k; g(k)] A_i^c] \geq 1$ in the form of $(*)_1$, both classically and constructively, have never been observed in practice. Note that here, the above property is not even mathematically equivalent to the formulation given by $\forall g^1 (\mathbb{P}[\exists k^0 \forall i^0 \in [k; g(k)] A_i^c] \geq 1)$ as the inner statement, by the presence of the term $g(k)$, is not monotone in k .

Now, by Theorem 5.1, from a classical proof of the above, we would be guaranteed to obtain a modulus Ψ such that

$$\forall m \in \mathbb{N} \forall g : \mathbb{N} \rightarrow \mathbb{N} \exists k \leq \Psi(m, g) \left(\mathbb{P} \left(\bigcup_{i \in [k; g(k)]} A_i \right) \leq 2^{-m} \right).$$

Such a modulus, which we call a *uniform metastable modulus of A_n infinitely often almost never*, features prominently in the recent work by Powell and Wan on quantitative aspects of zero-one laws in probability theory [55], but generally, by the aforementioned theorem, can be extracted from large classes of classical proofs of the above property.

Also, such a modulus could have already featured in the interpretation of the first Borel-Cantelli lemma as given by Arthan and the second author in [1], if the result would have been analysed classically instead of constructively. To make this more clear, recall that the first Borel-Cantelli lemma states that for a sequence of events $(A_n)_{n \in \mathbb{N}}$ in a probability space (Ω, S, \mathbb{P}) , if $\sum_{n \in \mathbb{N}} \mathbb{P}(A_n) < \infty$, then $\mathbb{P}(\limsup_{n \rightarrow \infty} A_n) = 0$. The paper [1] presents a constructive interpretation of this implication by extracting from its (constructive) proof a transformation that

maps any *rate of convergence* $\phi(m)$ realizing the premise, that is such that

$$\forall m \in \mathbb{N} \left(\sum_{n \geq \phi(m)} \mathbb{P}(A_n) \leq 2^{-m} \right),$$

into a modulus of A_n infinitely often almost never Ψ as defined above.

If we, however, were to analyse the proof classically, that is by first considering the negative translations, we are actually able to provide a transformation that maps any so-called rate of metastability $\phi(m, g)$ realizing the premise, a weaker computational interpretation given by

$$\forall m \in \mathbb{N} \forall g : \mathbb{N} \rightarrow \mathbb{N} \exists k \leq \phi(m, g) \left(\sum_{n=k}^{g(k)} \mathbb{P}(A_n) \leq 2^{-m} \right),$$

into a uniform metastable modulus of A_n infinitely often almost never Ψ as defined above. Indeed, the argument is essentially trivial. Simply note that

$$\mathbb{P} \left(\bigcup_{n \in [k; g(k)]} A_n \right) \leq \sum_{n=k}^{g(k)} \mathbb{P}(A_n)$$

holds using (finite) subadditivity, so that one can immediately set $\Psi = \phi$. Note that this result actually has the benefit that it immediately and elementarily, that is without any additional proof-specific arguments, implies the previous result on transformations of rates of convergence to moduli of A_n infinitely often almost never: If ϕ is a rate of convergence for $\sum_{n \in \mathbb{N}} \mathbb{P}(A_n) < \infty$, it is also a rate of metastability for $\sum_{n \in \mathbb{N}} \mathbb{P}(A_n) < \infty$, so that we obtain

$$\forall m \in \mathbb{N} \forall g : \mathbb{N} \rightarrow \mathbb{N} \exists k \leq \phi(m) \left(\sum_{n=k}^{g(k)} \mathbb{P}(A_n) \leq 2^{-m} \right).$$

A simple argument via contradiction shows that this is elementarily equivalent to ϕ being a modulus of A_n infinitely often almost never, as constructed in [1].⁷

6.3. Almost sure convergence. Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of random variables. One of the most fundamental statements about such a stochastic process is that these variables converge almost surely, where we here focus on the corresponding Cauchy formulation. That $(X_n(\varpi))_{n \in \mathbb{N}}$ is Cauchy can be expressed via the formula

$$\forall k^0 \exists n^0 \forall l^0 \forall i^0, j^0 \in [n; n+l] (|X_i(\varpi) - X_j(\varpi)| \leq_{\mathbb{R}} 2^{-k}).$$

We are now interested in different formal representations of the statement that X_n *converges almost surely*, together with the computational content that they entail.

Indeed, such convergence results so far are the most common type of theorems analysed in proof mining case studies, as aside from various infinitely often statements already discussed in the previous parts of this section, essentially all other major quantitative results in proof mining and probability theory so far focus on such a probabilistic convergence theorem in one way or the other.

Similarly to the previous case, with events occurring infinitely often almost never, we are here confronted with a Π_3^0 -property and so it will generally make a difference whether the statement

⁷While such an argument may seem somewhat elaborate for a result with such a trivial proof, we nonetheless presented it in some detail above in order to emphasize a recurring theme in proof mining: The classical analysis of a constructive proof often yields a more general computational result that, in particular, contains the constructive interpretation as a special case.

is proven constructively or not, in addition to the different interpretations already induced by the concrete formalizations of the associated properties in the style of $(*)_1 - (*)_3$.

Let us again begin with the formulation with the weakest computational content in the sense of Section 5, that is with

$$\forall k^0 (\mathbb{P}[\exists n^0 \forall l^0 \forall i^0, j^0 \in [n; n+l] (|X_i - X_j| \leq_{\mathbb{R}} 2^{-k})] \geq 1).$$

In a semi-constructive context, as illustrated by Theorem 5.2, the computational content of the above statement is given by a modulus $\Phi : \mathbb{N}^2 \rightarrow \mathbb{N}$ such that⁸

$$\forall m \in \mathbb{N} \forall k \in \mathbb{N} \exists n \leq \Phi(m, k) \forall l \in \mathbb{N} (\mathbb{P}(\forall i, j \in [n; n+l] (|X_i - X_j| \leq 2^{-k})) \geq 1 - 2^{-m}).$$

We call such a modulus Φ a *rate of almost sure convergence for X_n* , which, by the aforementioned theorem, can be extracted from large classes of semi-constructive proofs of the previous property. Such rates are abundant in the probability theory literature, and they feature crucially in recent case studies of proof mining in that area, such as laws of large numbers [42, 43, 44] and in particular stochastic optimization [48, 52, 53, 54].

In particular, as also highlighted before, these theorems will guarantee that the resulting rates will hold true for all (suitable) finitely additive probability spaces, which is a rather surprising result as in general, almost sure convergence does not imply the existence of a rate of almost sure convergence over finitely additive probability space.

In a classical context, one would as before first have to translate the respective statement into a $\forall\exists$ -theorem in the style of the negative translation of the respective statement, turning the previous formulation of convergence into

$$\forall k^0 \forall g^1 \exists n^0 \forall i^0, j^0 \in [n; n+g(n)] (|X_i(\varpi) - X_j(\varpi)| \leq_{\mathbb{R}} 2^{-k}).$$

Hence, there are again various options for representing the resulting probabilistic statement formally.

Let us first consider the statement

$$\forall k^0 \forall g^1 \mathbb{P}[\exists n^0 \forall i^0, j^0 \in [n; n+g(n)] (|X_i - X_j| \leq_{\mathbb{R}} 2^{-k})] \geq 1,$$

that is a statement of the form $(*)_2$. By Theorem 5.1, from a proof of this we would be able to extract a modulus $\Phi : \mathbb{N}^2 \times \mathbb{N}^{\mathbb{N}} \rightarrow \mathbb{N}$ such that

$$\forall m \in \mathbb{N} \forall k \in \mathbb{N} \forall g : \mathbb{N} \rightarrow \mathbb{N} (\mathbb{P}[\exists n \leq \Phi(m, k, g) \forall i, j \in [n; n+g(n)] (|X_i - X_j| \leq 2^{-k})] \geq 1 - 2^{-m}).$$

⁸Note that such a bound would also depend on a majorant of the sequence of random variables (X_n) . In the context of the system $\mathcal{F}^{\omega}[\mathbb{P}]$, random variables are represented as objects of type $1(\Omega)$ and hence, due to the uniform majorizability of the type Ω , are all assumed to be bounded (see [46] for further discussions of this). In that way, all extracted moduli concerned with (sequences) of random variables actually retain such a dependence on a (sequence of) almost-sure bounds for the variable(s). This creates a mismatch with some applications, where it is possible to derive bounds which are even more uniform on the variables in the sense that they only depend on their L_1 -norm or other measures related to integrability. These uniform bounds can be systematically accounted for via a slight extension of the systems presented in [46] and used here, based on the use of an additional abstract type for random variables and a point-free approach to the content space. While this will be fully explored in upcoming work with Thomas Powell, we want to highlight that this merely amounts to a modular extension of the present paper, tailored to a specific, more narrow notion of point-free proofs and corresponding uniform bounds, and in particular builds on the general translation developed in this paper. The resulting arguments for extracting such uniform bounds remain almost exactly the same as the ones given here, so that for the rest of this section, we reserve the right to be vague regarding the quantitative dependence of the bounds on the random variables, as we merely want to illustrate their structure. Regardless, we want to stress that the bounds are still guaranteed to be independent of all parameters relating to the underlying probability content space.

We call such a modulus Φ a *rate of metastable pointwise convergence for X_n* . This notion, as far as we know, first appeared in the work of Avigad, Dean and Rute [2] (extending a notion of Tao [60]). In recent proof mining case studies, it appears prominently in [49] for a quantitative version of a central convergence theorem in stochastic optimization.

The other immediate representation of the associated property is given by

$$\forall m^0 \forall k^0 \forall g^1 \exists n^0 (\mathbb{P}[\forall i^0, j^0 \in [n; n + g(n)] (|X_i - X_j| \leq_{\mathbb{R}} 2^{-k})] \geq 1 - 2^{-m})$$

in the form of $(*)_3$. Note that also here, similar to before in the case of infinitely often almost never statements, the above property is not immediately mathematically equivalent to the formulation given by $\forall k^0 \forall g^1 \mathbb{P}[\exists n^0 (|X_{n+g(n)}(\varpi) - X_n(\varpi)| \leq_{\mathbb{R}} 2^{-k})] \geq 1$ as the inner statement, by the presence of the term $g(n)$, is not monotone in n . Again, by Theorem 5.1, from a proof of this we would be able to extract a modulus $\Phi : \mathbb{N}^2 \times \mathbb{N}^{\mathbb{N}} \rightarrow \mathbb{N}$ such that

$$\begin{aligned} &\forall m \in \mathbb{N} \forall k \in \mathbb{N} \forall g : \mathbb{N} \rightarrow \mathbb{N} \exists n \leq \Phi(m, k, g) \\ &(\mathbb{P}[\forall i, j \in [n; n + g(n)] (|X_i - X_j| \leq 2^{-k})] \geq 1 - 2^{-m}). \end{aligned}$$

We call such a modulus Φ a *rate of metastable uniform convergence for X_n* , which again, by the aforementioned theorem, can be extracted from large classes of proofs. This notion, as far as we know, first appears in the work of Avigad, Gerhardy and Towsner [4] on applications of proof mining to ergodic theory and was later more systematically synthesized in the work of Avigad, Dean and Rute [2].

In particular, the work [2] is concerned with the relation that rates of metastable uniform convergence have with rates of metastable pointwise convergence. Now, as already discussed in the preceding Section 5 more generally, but as also immediately apparent from their formulation here, it is clear that a rate of metastable uniform convergence is a rate of metastable pointwise convergence, so only the converse direction is of interest. Now, in the context of probability measures, having a rate of metastable uniform convergence is immediately equivalent to the well-known notion of almost uniform convergence, while having a rate of metastable pointwise convergence is equivalent to pointwise convergence. However, almost sure and almost uniform convergence are equivalent over probability measures by the well-known theorem of Egorov (originally phrased in [9] and later extended by Riesz and Sierpiński), but simple counterexamples show that this equivalence does not extend to probability contents.

Nevertheless, these two quantitative notions of metastable pointwise and metastable uniform convergence remain equivalent over probability contents, as essentially already implicitly shown in [2], but only fully highlighted through a proof-theoretic lens in [46]. In particular, the work [2] provides a quantitative version of Egorov’s theorem on the level of these two metastable notions by means of an explicitly constructed functional which transforms a modulus of metastable pointwise convergence into a modulus of metastable uniform convergence. However, this functional has high computational complexity (bar-recursive in this case), reflecting the complexity of the proof of Egorov’s theorem, as also justified from a logical perspective in [46]. So, the proof-theoretic methodology furnishes the two notions of almost sure and almost uniform convergence with uniform quantitative variants for which a variant of Egorov’s theorem can be retained for probability contents. It is exactly this result of [2] that we abstracted into a general relationship between problems of the form $(*)_2$ and $(*)_3$ in the context of measurable and suitably monotone Π_1^0 -matrices in Section 5.

Lastly, we want to stress that a rate of almost sure convergence, as described above, is also a direct realiser of the formulation of almost uniform convergence as given above. Therefore, although almost sure convergence and almost uniform convergence are distinct notions for probability contents, if one can establish that a sequence of random variables converges almost

surely via a semi-constructive proof as formulated by the systems used in this paper, then one can not only extract a rate but also derive the qualitative result that the sequence converges almost uniformly in the context of probability contents.

Despite the presence of this high-complexity translation relating metastable uniform convergence and metastable pointwise convergence, essentially all quantitative convergence results for stochastic processes obtained via proof mining that do not yield a rate of almost sure convergence already natively provide the stronger rate of metastable uniform convergence. In particular, these results give rise to such rates directly and with low complexity, rather than first establishing a rate of metastable pointwise convergence and then appealing to the translation from [2]. This seems to be essentially because the respective proofs of the convergence results could be immediately formalized with the convergence explicitly phrased in the form

$$\forall m^0 \forall k^0 \forall g^1 \exists n^0 (\mathbb{P}[\forall i^0, j^0 \in [n; n + g(n)] (|X_i(\varpi) - X_j(\varpi)| \leq_{\mathbb{R}} 2^{-k})] \geq 1 - 2^{-m}).$$

Examples of this arise in the context of general stochastic processes, martingales and stochastic optimization [50, 51], laws of large numbers [43, 44], and ergodic theory in finitely additive probability spaces [45]. In the latter case, the general considerations on quantitative modes of convergence for random variables, as obtained via proof mining in the present paper, have been instrumental in deriving a new notion of almost sure convergence for probability contents. Indeed, there has so far been only one proof-mining application to stochastic processes that directly yields the weaker rate of metastable pointwise convergence, which is given in the recent work [49] on stochastic optimization already mentioned previously for its peculiar features regarding almost sure infinitely often statements.

In all these other cases, not only are these rates Φ of metastable uniform convergence of low complexity but they are in particular often of the special form

$$\Phi(m, k, g) := \tilde{g}^{(\phi(m, k))}(0)$$

for $\tilde{g}(n) := n + g(n)$ and another function $\phi(m, k)$ (commonly itself also a simple function), that is they are *learnable*, following the terminology of the work of Kohlenbach and Safarik [35] (which actually studies a much broader concept, in a deterministic setting however).⁹ Indeed, it is shown in [35] that under a restriction on the amount of law-of-excluded-middle in a proof, one can guarantee the extractability of learnable rates of metastability for deterministic sequences, and we deem it highly likely that this result extends to the present probabilistic context, so that results such as the above can be a priori guaranteed by means of our present probabilistic metatheorems together with a restriction on the amount of law-of-excluded-middle in a proof. We, however, leave this for future work.

Lastly, we just want to remark that statements such as

$$\forall k^0 (\mathbb{P}[\forall g^1 \exists n^0 (|X_{n+g(n)}(\varpi) - X_n(\varpi)| \leq_{\mathbb{R}} 2^{-k})] \geq 1)$$

or even $\mathbb{P}[\forall k^0 \forall g^1 \exists n^0 (|X_{n+g(n)}(\varpi) - X_n(\varpi)| \leq_{\mathbb{R}} 2^{-k})] \geq 1$, in the form of $(*)_1$, have never been observed in practice, so that we do not expand on them here.

6.4. Finite fluctuations. The final important property we want to discuss is the property that a stochastic process $(X_n)_{n \in \mathbb{N}}$ has *finite fluctuations almost surely*, that is that

$$\forall k \in \mathbb{N} (\mathbb{P}(J_k(X_n) < \infty) = 1),$$

where $J_k(X_n)$ is the total number of 2^{-k} -fluctuations that the process $(X_n)_{n \in \mathbb{N}}$ experiences, that is $J_k(X_n(\varpi)) := \lim_{N \rightarrow \infty} J_{N, k}(X_n(\varpi))$ where $J_{N, k}(X_n)$ is the number of 2^{-k} -fluctuations

⁹In general, even the notion of a rate of metastable uniform/pointwise convergence can be related to as a notion of generalized learnability as discussed in [47].

that the process $(X_n)_{n \in \mathbb{N}}$ experiences during the first N elements, i.e. the maximal $n \in \mathbb{N}$ such that there are

$$i_1 < j_1 \leq i_2 < j_2 \leq \dots \leq i_n < j_n < N \text{ with } |X_{i_l}(\varpi) - X_{j_l}(\varpi)| > 2^{-k},$$

for all $l = 1, \dots, n$. Fluctuations are a central tool in the study of stochastic processes and in particular their quantitative behaviour. In particular, they have featured in quite a few case studies on proof mining and probability theory, notably [49, 50] (see also [6] and [45]). In fact, having finite fluctuations is (non-effectively) equivalent to the process converging almost-surely (see e.g. the discussion in [50]). However, the quantitative property of having a bound on the number of fluctuations, that is a so-called *modulus of finite fluctuations almost surely* $\phi : \mathbb{N}^2 \rightarrow \mathbb{N}$ with the property that

$$\forall m \in \mathbb{N} \forall k \in \mathbb{N} (\mathbb{P}(J_k(X_n) \leq \phi(m, k)) \geq 1 - 2^{-m}),$$

falls, as a quantitative property, strictly between having a rate of almost-sure convergence and having a learnable rate of metastable pointwise convergence, which is in fact already true for deterministic sequences (see [35]). We refer to [35] for further discussions on this and related notions in the deterministic case and to [50] for the stochastic case.

The question when the extractability of moduli of finite fluctuations almost surely from convergence proofs can be guaranteed is quite subtle, as already discussed at length in the work of Kohlenbach and Safarik [35] for the deterministic case. Indeed, it seems difficult to find any a priori guarantee for this, and [35] only provides a type of a posteriori condition where, if one obtains a learnable rate of metastability that satisfies a certain so-called gap condition, then such a rate can be converted to a modulus of finite fluctuations. However, as far as is known, this gap condition cannot be guaranteed a priori based on additional structure on the proof, such as restricted uses of the law-of-excluded-middle.

The circumstances in the case studies mentioned above are, however, more specific. On a closer look, we find that these case studies actually all analyse proofs where the property of having finite fluctuations is *directly* derived from an associated strong inequality bounding the number of crossings for the stochastic process in question (see e.g. Theorem 5.1 in [50] and generally the exposition therein for further information on crossing inequalities). While the resulting proof of the almost-sure convergence of the process is not constructive, as having finite fluctuations almost surely only implies almost sure convergence non-effectively, this derivation of the property of finite fluctuations is in these cases always semi-constructive and can be analysed on its own, which provides a priori guarantees for the existence of modulus of finite fluctuations almost surely in these cases.

Indeed, we can formally recognize these cases as follows: Adapting Definition 2.2 from [35], we write

$$J_{k, \varpi}^X(n, i, j) \equiv \begin{cases} \text{lh}(i) =_0 \text{lh}(j) =_0 n \\ \wedge \forall l <_0 n (i_l <_0 j_l) \\ \wedge \forall l <_0 n - 1 (j_l \leq_0 i_{l+1}) \\ \wedge \forall l <_0 n (|X_{i_l}(\varpi) - X_{j_l}(\varpi)| >_{\mathbb{R}} 2^{-k}), \end{cases}$$

where $\text{lh}(i)$ and $\text{lh}(j)$ represent the length of the finite sequence coded by i and j (under some canonical coding of finite sequences) and i_l, j_l access their respective elements at index l . Therefore, $J_{k, \varpi}^X(n, i, j)$ formally expresses the property that $(X_n(\varpi))_{n \in \mathbb{N}}$ has n -many 2^{-k} -fluctuations with indices coded by i and j . We write

$$J_{k, \varpi}^X \leq b \equiv \forall n >_0 b \forall i^0 \forall j^0 \neg J_{k, \varpi}^X(n, i, j),$$

expressing that $(X_n(\varpi))_{n \in \mathbb{N}}$ has $\leq b$ -many 2^{-k} -fluctuations. That $(X_n(\varpi))_{n \in \mathbb{N}}$ has finite fluctuations can now be formally expressed via

$$\forall k^0 \exists b^0 (J_{k, \varpi}^X \leq b),$$

so that the resulting property of having finite fluctuations almost surely translates to

$$\forall k^0 (\mathbb{P}[\exists b^0 (J_k^X \leq b)] \geq 1),$$

that is a statement of type $(*)_2$. By Theorem 5.5, from a semi-constructive proof of this (which is a necessary requirement as $J_k^X \leq b$ is, essentially, universal), we are then able to extract a modulus $\phi : \mathbb{N}^2 \rightarrow \mathbb{N}$ such that

$$\forall m \in \mathbb{N} \forall k \in \mathbb{N} \left(\mathbb{P} \left[\bigcup_{b \leq \phi(m, k)} (J_k^X \leq b) \right] \geq 1 - 2^{-m} \right).$$

Noting that using the monotonicity of $J_k^X \leq b$ the above is equivalent to

$$\forall m \in \mathbb{N} \forall k \in \mathbb{N} (\mathbb{P} [J_k^X \leq \phi(m, k)] \geq 1 - 2^{-m}),$$

this directly represents that ϕ is a modulus of finite fluctuations almost surely, as defined above. It is exactly such analyses of semi-constructive proofs of this property that give rise to the moduli of finite fluctuations almost surely constructed in the case studies mentioned previously, giving the first systematic proof-theoretic explanation thereof.

Acknowledgments: The authors are grateful to Ulrich Kohlenbach for numerous helpful comments on various drafts of this paper. In particular, he provided crucial insights on the treatment of uniform boundedness principles in higher types and of contra-collection principles in the context of the monotone functional interpretation. The authors also thank Thomas Powell for valuable remarks on a draft of the paper.

REFERENCES

- [1] R. Arthan and P. Oliva. On the Borel–Cantelli Lemmas, the Erdős–Rényi Theorem, and the Kochen–Stone Theorem. *Journal of Logic and Analysis*, 13(6):1–23, 2021.
- [2] J. Avigad, E. Dean, and J. Rute. A metastable dominated convergence theorem. *Journal of Logic and Analysis*, 4(3), 2012. 19pp.
- [3] J. Avigad and S. Feferman. Gödel’s functional (“Dialectica”) interpretation. In S.R. Buss, editor, *Handbook of Proof Theory*, volume 137 of *Studies in Logic and the Foundations of Mathematics*, pages 337–405. Elsevier, Amsterdam, 1998.
- [4] J. Avigad, P. Gerhardy, and H. Towsner. Local stability of ergodic averages. *Transactions of the American Mathematical Society*, 362(1):261–288, 2010.
- [5] J. Avigad and J. Iovino. Ultraproducts and metastability. *New York Journal of Mathematics*, 19:713–727, 2013.
- [6] J. Avigad and J. Rute. Oscillation and the mean ergodic theorem for uniformly convex Banach spaces. *Ergodic Theory and Dynamical Systems*, 35(4):1009–1027, 2015.
- [7] M. Bezem. Strongly majorizable functionals of finite type: a model for bar recursion containing discontinuous functionals. *The Journal of Symbolic Logic*, 50:652–660, 1985.
- [8] E. Dueñez and J. Iovino. Model theory and metric convergence I: Metastability and dominated convergence. In J. Iovino, editor, *Beyond first order model theory*, volume 1, pages 131–187. Chapman and Hall/CRC, New York, 2017.
- [9] D.Th. Egoroff. Sur les suites des fonctions mesurables. *Comptes rendus hebdomadaires des séances de l’Académie des sciences*, 152:244–246, 1911.
- [10] P. Engrácia and F. Ferreira. Bounded functional interpretation with an abstract type. In A. Rezus, editor, *Contemporary Logic and Computing*, volume 1 of *Landscapes in Logic*, pages 87–112. College Publications, London, 2020.

- [11] F. Ferreira. Injecting uniformities into Peano Arithmetic. *Annals of Pure and Applied Logic*, 157(2–3):122–129, 2009.
- [12] F. Ferreira and P. Oliva. Bounded functional interpretation. *Annals of Pure and Applied Logic*, 135:73–112, 2005.
- [13] P. Gerhardy and U. Kohlenbach. Strongly uniform bounds from semi-constructive proofs. *Annals of Pure and Applied Logic*, 141:89–107, 2006.
- [14] P. Gerhardy and U. Kohlenbach. General logical metatheorems for functional analysis. *Transactions of the American Mathematical Society*, 360:2615–2660, 2008.
- [15] K. Gödel. Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *Dialectica*, 12:280–287, 1958.
- [16] A. Grzegorzcyk, editor. *Some classes of recursive functions*, volume IV of *Rozprawy Matematyczne*. Instytut Matematyczny Polskiej Akademi Nauk, Warsaw, 1953.
- [17] D. Günzel and U. Kohlenbach. Logical metatheorems for abstract spaces axiomatized in positive bounded logic. *Advances in Mathematics*, 290:503–551, 2016.
- [18] P.R. Halmos. *Measure Theory*, volume 18 of *Graduate Texts in Mathematics*. Springer New York, NY, 1950.
- [19] D. Hilbert. Über das Unendliche. *Mathematische Annalen*, 95(1):161–190, 1926.
- [20] W.A. Howard. Hereditarily majorizable functionals of finite type. In A.S. Troelstra, editor, *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, volume 344 of *Lecture Notes in Mathematics*, pages 454–461. Springer, New York, 1973.
- [21] H.J. Keisler. Probability Quantifiers. In J. Barwise and S. Feferman, editors, *Model-Theoretic Logics*, Perspectives in Mathematical Logic, pages 509–556. Springer Berlin, Heidelberg, 1985.
- [22] H.J. Keisler. Measures and forking. *Annals of Pure and Applied Logic*, 34(2):119–169, 1987.
- [23] A. Klenke. *Probability Theory: A Comprehensive Course*. Universitext. Springer London, 2008.
- [24] U. Kohlenbach. Effective bounds from ineffective proofs in analysis: an application of functional interpretation and majorization. *The Journal of Symbolic Logic*, 57:1239–1273, 1992.
- [25] U. Kohlenbach. Analysing proofs in analysis. In W. Hodges, M. Hyland, C. Steinhorn, and J. Truss, editors, *Logic: From Foundations to Applications. European Logic Colloquium (Keele, 1993)*, pages 225–260. Oxford University Press, Oxford, 1996.
- [26] U. Kohlenbach. Mathematically strong subsystems of analysis with low rate of growth of provably recursive functionals. *Archive for Mathematical Logic*, 36:31–71, 1996.
- [27] U. Kohlenbach. Relative constructivity. *The Journal of Symbolic Logic*, 63:1218–1238, 1998.
- [28] U. Kohlenbach. The use of a logical principle of uniform boundedness in analysis. In A. Cantini, E. Casari, and P. Minari, editors, *Logic and Foundations of Mathematics*, volume 280 of *Synthese Library*, pages 93–106. Kluwer Academic, Dordrecht, 1999.
- [29] U. Kohlenbach. Some logical metatheorems with applications in functional analysis. *Transactions of the American Mathematical Society*, 357(1):89–128, 2005.
- [30] U. Kohlenbach. A logical uniform boundedness principle for abstract metric and hyperbolic spaces. *Electronic Notes in Theoretical Computer Science*, 165:81–93, 2006.
- [31] U. Kohlenbach. *Applied Proof Theory: Proof Interpretations and their Use in Mathematics*. Springer Monographs in Mathematics. Springer Berlin, Heidelberg, 2008.
- [32] U. Kohlenbach. A note on the monotone functional interpretation. *Mathematical Logic Quarterly*, 57(6):611–614, 2011.
- [33] U. Kohlenbach. Proof-theoretic Methods in Nonlinear Analysis. In B. Sirakov, P. Ney de Souza, and M. Viana, editors, *Proceedings ICM 2018*, volume 2, pages 61–82. World Scientific, Singapore, 2019.
- [34] U. Kohlenbach and P. Oliva. Proof Mining: A Systematic Way of Analysing Proofs in Mathematics. *Proceedings of the Steklov Institute of Mathematics*, 242:136–164, 2003.
- [35] U. Kohlenbach and P. Safarik. Fluctuations, effective learnability and metastability in analysis. *Annals of Pure and Applied Logic*, 165(1):266–304, 2014.
- [36] G. Kreisel. On the Interpretation of Non-Finitist Proofs – Part I. *The Journal of Symbolic Logic*, 16(4):241–267, 1951.
- [37] G. Kreisel. On the Interpretation of Non-Finitist Proofs – Part II. Interpretation of Number Theory. Applications. *The Journal of Symbolic Logic*, 17(1):43–58, 1952.
- [38] G. Kreisel. Interpretation of analysis by means of constructive functionals of finite types. In A. Heyting, editor, *Constructivity in Mathematics*, Lecture Notes in Mathematics, pages 101–128. North-Holland Publishing Company, Amsterdam, 1959.

- [39] G. Kreisel. On weak completeness of intuitionistic predicate logic. *The Journal of Symbolic Logic*, 27:139–158, 1962.
- [40] S. Kuroda. Intuitionistische Untersuchungen der formalistischen Logik. *Nagoya Mathematical Journal*, 2:35–47, 1951.
- [41] P.A. Loeb. Conversion from nonstandard to standard measure spaces and applications in probability theory. *Transactions of the American Mathematical Society*, 211:113–122, 1975.
- [42] M. Neri. Quantitative Strong Laws of Large Numbers. *Electronic Journal of Probability*, 30(20), 2024. 22pp.
- [43] M. Neri. A finitary Kronecker’s lemma and large deviations in the strong law of large numbers. *Annals of Pure and Applied Logic*, 176(6), 2025. 103569, 31pp.
- [44] M. Neri. Oscillations in the Strong Law of Large Numbers, 2025. preprint available at <https://kejineri.github.io>.
- [45] M. Neri. The pointwise ergodic theorem on finitely additive spaces, 2025. preprint available at <https://arxiv.org/abs/2511.01676>.
- [46] M. Neri and N. Pischke. Proof mining and probability theory. *Forum of Mathematics, Sigma*, 13, 2025. e187, 47pp.
- [47] M. Neri, N. Pischke, and T. Powell. Generalized learnability of stochastic principles. In A. Beckmann, I. Oitavem, and F. Manea, editors, *Proceedings of the 21st Conference on Computability in Europe*, volume 15764 of *Lecture Notes in Computer Science*, pages 333–348. Springer, Berlin Heidelberg, 2025.
- [48] M. Neri, N. Pischke, and T. Powell. On the asymptotic behaviour of stochastic processes, with applications to supermartingale convergence, Dvoretzky’s approximation theorem, and stochastic quasi-Fejér monotonicity, 2025. preprint available at <https://arxiv.org/abs/2504.12922>.
- [49] M. Neri, N. Pischke, and T. Powell. Generalized fluctuation bounds for stochastic algorithms in the presence of compactness, 2026. preprint available at <https://arxiv.org/abs/2602.22741>.
- [50] M. Neri and T. Powell. On quantitative convergence for stochastic processes: Crossings, fluctuations and martingales. *Transactions of the American Mathematical Society, Series B*, 12:974–1019, 2025.
- [51] M. Neri and T. Powell. A quantitative Robbins-Siegmund theorem. *Annals of Applied Probability*, 36(1):636–651, 2026.
- [52] N. Pischke. Mean-square and linear convergence of a stochastic proximal point algorithm in metric spaces of nonpositive curvature, 2025. preprint available at <https://arxiv.org/abs/2510.10697>.
- [53] N. Pischke. On Busemann subgradient methods for stochastic minimization in Hadamard spaces, 2026. preprint available at <https://arxiv.org/abs/2602.08127>.
- [54] N. Pischke and T. Powell. Asymptotic regularity of a generalised stochastic Halpern scheme, 2024. preprint available at <https://arxiv.org/abs/2411.04845>.
- [55] T. Powell and A. Wan. An approximate zero-one law via the Dialectica interpretation, 2025. preprint available at <https://arxiv.org/abs/2508.20849>.
- [56] K.P.S. Bhaskara Rao and M. Bhaskara Rao. *Theory of Charges*, volume 109 of *Pure and Applied Mathematics*. Academic Press, London, 1983.
- [57] M. Schönfinkel. Über die Bausteine der mathematischen Logik. *Mathematische Annalen*, 92:305–316, 1924.
- [58] D. Scott and P. Krauss. Assigning probabilities to logical formulas. In J. Hintikka and P. Suppes, editors, *Aspects of Inductive Logic*, volume 43 of *Studies in Logic and the Foundations of Mathematics*, pages 219–264. North-Holland, Amsterdam, 1966.
- [59] C. Spector. Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics. In J.C.E. Dekker, editor, *Recursive Function Theory*, volume 5 of *Proceedings of Symposia in Pure Mathematics*, pages 1–27. American Mathematical Society, Providence, RI, 1962.
- [60] T. Tao. Norm convergence of multiple ergodic averages for commuting transformations. *Ergodic Theory and Dynamical Systems*, 28:657–688, 2008.
- [61] T. Tao. *Structure and Randomness: Pages from Year One of a Mathematical Blog*, chapter Soft Analysis, Hard Analysis, and the Finite Convergence Principle, pages 17–29. American Mathematical Society, Providence, RI, 2008.
- [62] A.S. Troelstra, editor. *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, volume 344 of *Lecture Notes in Mathematics*. Springer, Berlin, 1973.