Forum of Mathematics, Sigma (2025), **Vol. 13**:e187 1–47
doi:10.1017/fms.2025.10138

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Proof mining and probability theory

Morenikeji Neri [1] and Nicholas Pischke [2]

[1]University of Bath, United Kingdom; E-mail: neri@mathematik.tu-darmstadt.de (Corresponding author).
[2]TU Darmstadt, Germany; E-mail: nnp39@bath.ac.uk.

Received: 12 September 2024; **Revised:** 3 June 2025; **Accepted:** 15 October 2025

2020 Mathematics Subject Classification: *Primary* – 03F10; *Secondary* – 03F35, 28A12, 28A25, 60A10

Abstract

We extend the theoretical framework of proof mining by establishing general logical metatheorems that allow for the extraction of the computational content of theorems with prima facie "noncomputational" proofs from probability theory, thereby unlocking a major branch of mathematics as a new area of application for these methods. Concretely, we first devise proof-theoretically tame logical systems that allow for the formalization of proofs involving algebras of sets together with probability contents, that is probability measures which are only assumed to be finitely additive. Based on these systems, we provide extensions for the tame treatment of Lebesgue integrals on probability contents as well as σ-algebras and associated probability measures, all via intensional approaches. All these systems are then shown to be amenable to proof-theoretic metatheorems in the style of proof mining which guarantee the extractability of effective and tame bounds from large classes of ineffective existence proofs in probability theory. Moreover, these extractable bounds are guaranteed to be highly uniform in the sense that they will be independent of all parameters relating to the underlying probability space, particularly regarding events or measures of them. As such, these results in particular provide the first logical explanation for the success and the observed uniformities of the previous ad hoc case studies of proof mining in these areas and further illustrate their extent. Lastly, we establish a general proof-theoretic transfer principle that allows for the lift of quantitative information on a relationship between different modes of convergence for sequences of real numbers to sequences of random variables.

1. Introduction

One of the fundamental driving questions of proof theory is the following: What is the computational content of a mathematical theorem, and how can it be exhibited? Proof mining, which emerged as a subfield of mathematical logic in the 1990s through the work of Ulrich Kohlenbach and his collaborators,[1] aims at answering that question by extracting this computational content from theorems with proofs as they are found in the mainstream mathematical literature. This is a nontrivial task, in particular as such proofs are prima facie noneffective, involving both classical logic as well as various noncomputational (set-theoretic) principles. However, backed by a logical apparatus relying on the utilization of various methods from proof theory like functional interpretations and majorizability, the program of proof mining has had great success in various areas of mathematics, in particular regarding (nonlinear) analysis and optimization (see in particular the recent surveys [27, 28]).

Two areas that proof mining has previously only touched upon briefly are the fields of measure theory in general and probability theory in particular. Concretely, we refer to the works [1, 2, 3] which

[1]Historically, proof mining has its roots in Kreisel's program of the "unwinding of proofs" [33, 34]. We refer to [26] for a comprehensive monograph on proof mining and its applications until 2008, and to the survey article [31] for details on the earlier development of proof mining.

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

are essentially the only proof mining case studies in these areas preceding this work. From a practical perspective, this diffidence of proof mining regarding these areas is at least partially due to the fact that they are so far not substantiated by underlying logical methods as other areas of applications for proof mining are. This absence of a firm proof-theoretic foundation is largely due to a range of logical difficulties inherently present in the context of areas strongly reliant on set-theoretic methods, as is precisely the case with measure and probability theory. In any case, this lack of such a formal approach to these areas also renders the previous applications, to a certain degree, ad hoc.

It is the aim of this paper to extend the current logical methods used in proof mining as to render them applicable to large classes of proofs from probability theory, in particular so that they allow for a logical explanation of the success of the previous case studies mentioned before (as well as of the various properties of the extracted content).

1.1. The logical foundations of proof mining

The fundamental logical "substrates" of the proof mining program are the so-called general logical metatheorems.[2] These use well-known proof interpretations like Gödel's functional interpretation [13], negative translations (see, e.g., [36]), and their extensions[3] to provide a general result that quantifies and allows for the extraction of the computational content of large classes of theorems from their proofs. Furthermore, these proofs may involve classical logic and various noncomputational principles. In that way, proof mining, as substantiated by these metatheorems, has led to hundreds of applications in the last decades.

A crucial innovation in the techniques underlying these logical metatheorems was introduced by Kohlenbach in [25], marking the "modern age" of proof mining: The metatheorems for proof mining preceding [25] were based on "pure" systems for arithmetic in all finite types (see, e.g., [26, 60]) and as such were restricted in their expressivity to dealing with spaces and structures that were representable as Polish metric spaces in the sense of Baire space (and thus separable). The paradigm first proposed in [25] was to extend the language of the underlying systems with additional abstract base types which, together with additional constants and governing axioms, could be used to talk about much larger and broader classes of spaces and objects on them beyond merely representable ones.[4] The class of spaces and objects treated in this fashion has grown since then to a rather sizable amount, ranging from fundamental examples like general (nonseparable) metric, hyperbolic, CAT(0), Banach and Hilbert spaces to much more involved objects like $\mathbb{R}$-trees, L^p-spaces, the dual of a (nonseparable) Banach space as well as monotone operators and nonlinear semigroups, among many others.

Further, the approach to represent various classes of spaces via abstract types has, together with an ingenious combination due to Kohlenbach (see the discussions in Section 8 later on for further details and references on this) of Gödel's functional interpretation with (a suitable extension of) Howard's notion of majorizability [19], resulted in logical metatheorems that, beyond the extractability of effective bounds from noneffective existence proofs, guarantee a high degree of uniformity of the respective data. These perspectives of using abstract types to represent general spaces, together with the use of notions of majorizability to induce uniformities, are also fundamental to the present paper.

1.2. An abstract approach to probability theory

In this paper, we follow a novel abstract approach towards a treatment of various fundamental notions from probability theory to avoid the range of issues which are present a priori in that context, as briefly

[2]Examples of such logical metatheorems for proof mining can, for example, be found in [10, 12, 16, 25, 26, 30, 37, 38, 49, 50, 51, 52, 56].

[3]For example, other proof interpretations used in proof mining include Kreisel's modified realizability interpretation [33] for the treatment of semi-constructive proofs, monotone variants of the Dialectica interpretation [23] to deal with the extraction of computable bounds from proofs or the related bounded functional interpretation [11].

[4]Whether a class of spaces or objects can be treated in such a context ultimately depends on the uniformity and complexity of the axioms describing said class. This will be discussed in more detail later in this paper.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

mentioned before. Namely, already the most fundamental notions from probability like countable unions of measurable sets or σ-additive measures require the use of proof-theoretically strong comprehension principles to deal with the high quantifier complexity inherent in their defining axioms. For example, the existence of countable unions can be rather immediately recognized as a type of Σ_1-comprehension principle while the σ-additivity of a measure is likewise a strong existence property as it immediately implies various limit theorems for probabilities. Now, while this strength might be realized in extreme cases, the practice of probability theory suggests that in many situations, in particular where these objects are only discussed in an "abstract" way, theorems from probability theory can be given computational solutions with low complexity. An approach towards extractive proof theory based on a direct specification of these objects and notions would hence distort the complexity of bounds extracted from proofs in these situations.

We here provide a proof-theoretic approach which characterizes situations where these notions are indeed inherently "tame" in the sense that, although in principle being subject to well-known Gödelian phenomena, their mere presence and abstract use do not contribute to the strength of extractable bounds.[5] For that, we first focus on so-called probability contents (also called charges), that is, evaluations of sets which behave like a measure but are only finitely additive. This absence of the strong requirement of σ-additivity then also allows us to, at first, lift the restriction of the closure under countable unions of the space of events, so that we only consider these mappings to operate on (Boolean) algebras of sets (also called fields). In fact, these contents on algebras form a well-studied part of modern probability and measure theory with a rich theory and, as, for example, highlighted in the seminal book by K.P.S. Bhaskara Rao and M. Bhaskara Rao [6] on general (that is not necessarily $[0, 1]$-valued) contents, they are in a way "more interesting, more difficult to handle, and perhaps more important than countably additive ones" (which, as a statement, is attributed to Bochner in the foreword of [6]).

To formally approach probability contents on algebras, we follow a so-called intensional approach, that is, we fundamentally employ new base types and constants to provide an abstract access to the involved objects instead of relying on particular representations. Concretely, over a base theory of arithmetic in all finite types, we use new abstract types to provide a quantifier-free access to the base set of the content space and the algebra of events over this set and employ constants utilizing these types to provide an abstract access to the related set-theoretic operations and the content. These new types and constants are then governed by admissible axioms, that is axioms of low computational strength, which simply describe the fundamental properties of these objects instead of extensionally specifying their precise structure through any kind of coding.

We then utilize this base system for contents on algebras to define extensions which allow us to treat countably infinite unions and hence to provide access to the theory of σ-algebras and associated σ-additive probability measures. Again, our treatment is intensional here, to avoid the inherent difficulties with infinite unions laid out before. Concretely, we approach infinite unions by adding a novel constant to the underlying system that provides a direct and abstract access to them as an operator associating a new measurable set, representing the union, with a sequence of measurable sets. However, instead of specifying that the resulting values indeed represent the unions in question through the strong associated (comprehension-type) axiom, which prompted us to consider probability contents on algebras in the first place, we only specify that it "behaves like" a union through a combination of admissible axioms and rules. Further, we provide a tame approach to the space of bounded and Borel-measurable functions to introduce the Lebesgue integral (already in the context of probability contents), similarly an object that relies on strong comprehension principles in its classical formulation. Also here, we opt for an intensional specification which, instead of describing said objects explicitly and completely, axiomatizes a general structure just adhering to some essential properties.

In that way, we arrive at suitable systems for probability contents on algebras of sets, and at various extensions of those systems for these other fundamental notions from probability theory. All of these we

[5]This phenomenon is referred to as "proof-theoretic tameness" in the work of Kohlenbach [29], see also [40, 41] for related discussions of such phenomena.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

then endow with corresponding metatheorems on the extraction of computational bounds in the style of proof mining, based on (a monotone variant of) Gödel's functional interpretation.

These metatheorems thereby provide a formal proof-theoretic perspective on how an abstract use of the central notions of probability theory, now formally characterized through our intensional approach, does not (artificially) contribute to the strength of extractable bounds. Concretely, the metatheorems formally guarantee that the complexity of the extractable information depends only on (and can be a priori bounded in terms of) the complexity of the principles used in the corresponding proof. In that way, while the main base system taken in this paper is actually one that does contain large amounts of comprehension (to illustrate the potential strength of systems which are amenable to proof mining methods), the approach to the various objects from probability theory taken here does not rely on these strong principles at all, as motivated before, and thus can also immediately be developed over suitably weak subsystems (as, e.g., the collection of systems introduced in [24] based on the Grzegorczyk hierarchy [15]) where most of the mathematics discussed here can be similarly carried out but where then bounds of correspondingly low complexity could be guaranteed. In that way, the metatheorems established here indeed provide the right background to formally elucidate the extent of the phenomenon of proof-theoretic tameness for probability theory, which as discussed in detail before, was one of the main motivations for their design.

Beyond this tameness, one of the most crucial features of the new metatheorems presented in this paper is that they guarantee a high uniformity of the extracted content, in particular that it will be independent of the measure, the underlying set, and the algebra. As in the case of the first modern metatheorems mentioned before, this relies on a specific extension of the notion of majorizability due to Howard, which in our case utilizes the probability content (or measure) to provide a corresponding notion of majorizability for the new abstract types. In particular, we want to note that this is the first time in proof mining that an extension of Howard's majorizability notion to abstract spaces is utilized that does not rely on any metric structure of the underlying spaces. In that way, the present metatheorems provide the first concrete logical explanation of the uniformities of extractable bounds observed in the previously mentioned proof mining case studies, as will be discussed further later on (see in particular Section 1.4).

1.3. Related work

The present approach to proof mining and probability theory is, in particular, to be distinguished from previous work on logical aspects of (quantitative) probability theory. On the proof-theoretic side, the main preceding work is that by Kreuzer [35] on extracting computational content from proofs in measure theory. This approach, however, relies on strong forms of comprehension for treating measure spaces (represented via a specific coding), resulting in a rather restricted formal theory with a more limited scope of analyzable theorems, and which further does not guarantee any of the uniformity features for the quantitative information extracted thereby, in contrast to the present results. Outside of proof theory, finitizations of concepts from probability theory that enjoy similar uniformities as the ones considered here have also been obtained using tools from model theory, particularly ultraproducts [4, 8, 14]. However, these works differ from the present one in both method and scope. The first crucial difference is that the proof-theoretic metatheorems presented here actually provide a method for extracting the uniform quantitative information which the model-theoretic approach can only infer the existence of. Beyond that, as already commented on before, in the present approach, the complexity of that information can be gauged beforehand based on the principles used in the proof, and the proof-theoretic tameness of that information can be guaranteed thereby. Lastly, the model-theoretic approach is essentially fixed to focus on convergence statements and uniformities relating to their so-called metastable formulations. In contrast, the proof-theoretic approach presented here does not have that limitation (which has a crucial impact on the range of possible applications as will also be discussed again in Section 1.4). These facts crucially separate our work from the model-theoretic approach. Nevertheless, the model-theoretic approaches are certainly not subsumed by our work but rather complementary. For one, the

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

model-theoretic approaches to uniformities of rates of metastability in probability theory only rely on the truth of a statement, while our proof-theoretic results rely on the provability of the underlying theorem in an (albeit very strong) underlying theory. Also, we want to highlight that in particular the approach by Dueñez and Iovino [8] seems to rely, upon a closer inspection, on rather similar ideas for approaching some of the initial objects in question (e.g., by effectively treating probability contents on algebras instead of measures on σ-algebras and treating these underlying spaces with their operations as abstract entities), a fact that illustrates the apparent similarity of the problems that both the proof-theoretic and the model-theoretic approaches to these types of questions face and which in particular further highlights the naturalness of the approach followed here. However, our approach starts to crucially differ also conceptually instead of just methodologically from the work in [8] in the way that infinite unions and integration are treated in our systems and in the way that the uniformities of associated (extracted) bounds are guaranteed.

1.4. Applications, case studies, and extensions

Besides these logical considerations, the applications of the present metatheorems to probability theory are arguably the most important consequence of the present work. We hence use the last two sections of this paper to substantiate the applicability of the logical metatheorems introduced here.

At first, we outline in detail how the quantitative analysis of Egorov's theorem as presented in the seminal case study for proof mining in measure theory from [2] formalizes in our systems. Thereby, as mentioned before, we also provide the first logical explanation of the uniformities of the bounds extracted therein.[6]

Furthermore, the analysis of the results from [2] provided later highlights that these remain true for probability contents. This therefore illustrates that the notions and proofs produced in the work [2], by following the finitary perspective of proof mining, allow for a lift of the underlying convergence result to the theory of contents, a qualitative result that complements the quantitative results produced in [2] in the form of a corresponding rate. In that way, this points to an apparent empirical phenomenon which we want to highlight here: Finitary quantitative variants of notions and results from the theory of probability measures, as suggested by proof mining, seem to provide analogous versions suitable for the theory of probability contents. This view is largely corroborated by the further case studies developed since this work (as discussed below). Further, and more generally, these results thereby also highlight the naturalness of the theory of contents as an underlying medium for developing a logical account of probability theory in the sense of proof mining and fuel our confidence that already this system for probability contents will provide a suitable base for proof mining developments in the context of probability theory in the future. In that way, the fundamental relevance of (probability) contents, relating to the statement by Bochner quoted above, is essentially rediscovered by the proof-theoretic approach presented in this paper.

As another application of our novel logical approach, we establish a general so-called proof-theoretic transfer principle that allows for a lift of computational information on the relation between modes of convergence of sequences of real numbers to sequences of random variables. These results thereby provide a formal footing for this type of strategy, which is rather abundant in probability theory and in particular features in some recent case studies on proof mining and probability theory by the first author [43].

Besides these examples of application discussed here, the applicability of the systems presented in this paper is further substantiated by the fact that they explain the previously mentioned application [1] preceding this work as well as the very recent works in laws of large numbers [42, 43], asymptotic behavior of stochastic processes [45, 47], and stochastic optimization [44, 46, 53] (along with forthcoming

[6]This is in particular to be compared with the work by Dueñez and Iovino [8] where a model-theoretic account of the uniformities of a quantitative variant of the dominated convergence theorem, partially akin to the results of [2], is given. These considerations of [8], however, avoid Egorov's theorem and with that the central object of study in [2] by which their quantitative version of the dominated convergence theorem is established therein.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

work by the authors together with Thomas Powell) as instances of the present methodology. In particular, in the context of these recent works, the extractive proof-theoretic perspective of this work was crucial for obtaining the respective results. The growing number of case studies thereby in particular empirically illustrates that the present abstract approach indeed has a rather broad applicability, as hoped for initially.

Beyond that, the present work lends itself both to further theoretical investigations on the extension of proof mining methods to further notions from probability and measure theory like Bochner integrals and martingales, among many others, as well as to substantiate and carry out many further (and potentially much more sophisticated) applications of proof mining in this area beyond those already mentioned. In particular, we want to mention that most ideas developed here could be extended, mutatis mutandis, to general finite contents and measures.

2. Preliminaries

The basic system that we rely on is the system $\mathcal{A}^\omega = \mathsf{WE\text{-}PA}^\omega + \mathsf{QF\text{-}AC} + \mathsf{DC}$ for classical analysis in all finite types as commonly used in proof mining, formalized via (a weakly extensional variant of) Peano arithmetic in all finite types together with a few choice principles (see, e.g., [25] where this notation for the system was, presumably, first introduced). As all systems introduced here will be extensions of this system $\mathcal{A}^\omega$, we in this section sketch the essential features relevant for this paper. For any further details, we refer to the works [26, 60].

Here, we follow the definition of weakly extensional Peano arithmetic in all finite types $\mathsf{WE\text{-}PA}^\omega$ as, for example, given in [26] (see also [60]) and, in that way, we do not recall all the defining features $\mathsf{WE\text{-}PA}^\omega$ here and only focus on the four main aspects which are relevant in detail for this paper. In general, we denote function types using the bracket notation used in [26], that is, $\rho(\tau)$ is the type of functions that map objects of type τ to objects of type ρ, and we use T to denote the set of all finite types as usual, that is,

$$0 \in T, \quad \rho, \tau \in T \to \rho(\tau) \in T.$$

As usual, we denote pure types by natural numbers by setting $n+1 := 0(n)$. The four central properties of $\mathsf{WE\text{-}PA}^\omega$ that we need here are that, for one, the only primitive relation is equality at type 0 (denoted by $=_0$) and higher-type equality is only defined as an abbreviation via recursion with

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

For another, $\mathsf{WE\text{-}PA}^\omega$ crucially does not contain the full extensionality principles

$$\forall x^{\tau(\rho)}, y^\rho, y'^\rho \big(y =_\rho y' \to xy =_\tau xy'\big). \tag{$\mathsf{E}_{\rho,\tau}$}$$

Instead, it only contains the quantifier-free extensionality rule

$$\frac{A_0 \to s =_\rho t}{A_0 \to r[s/x^\rho] =_\tau r[t/x^\rho]} \tag{QF–ER}$$

where A_0 is a quantifier-free formula, s and t are terms of type ρ and r is a term of type τ. This lack of full extensionality is essential for establishing results on program extraction from classical proofs as the full extensionality axiom is not admissible in the context of the Dialectica interpretation, the main tool used later extract bounds from (classical) proofs. However, the extensionality rule is admissible in that context and so, from an applied perspective, serves the much-needed purpose of reintroducing an admissible fraction of extensionality back into the system to be practically as flexible as possible when dealing with proofs from the literature. We refer to [26] for a detailed discussion on this.

Further, $\mathsf{WE\text{-}PA}^\omega$ contains constants $\underline{R}_\rho$ for simultaneous primitive recursion in the sense of Gödel [13] and Hilbert [18] as governed by the axioms

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

$$\begin{cases} (R_i)_{\underline{\rho}} 0 \underline{y} \underline{z} =_{\rho_i} y_i \\ (R_i)_{\underline{\rho}} (Sx) \underline{y} \underline{z} =_{\rho_i} z_i (\underline{R}_{\underline{\rho}} x \underline{y} \underline{z}) x \end{cases} \quad \text{for } i = 1, \ldots, k \qquad (\underline{R})$$

where $\underline{\rho} = \rho_1, \ldots, \rho_k$ is a tuple of types, $\underline{y} = y_1, \ldots, y_k$ with y_i of type ρ_i and $\underline{z} = z_1, \ldots, z_k$ with z_i of type $\rho_i(0)\underline{\rho}^t$ where we write $\underline{\rho}^t := (\rho_k) \ldots (\rho_1)$.

Lastly, due to the inclusion of the combinators of Schönfinkel [54] in the language of WE-PA$^\omega$, the system allows the definition of λ-abstraction in the sense that for any term t of type τ and any variable x of type ρ, we can construct a term $\lambda x.t$ of type $\tau(\rho)$ such that the free variables of $\lambda x.t$ are exactly those of t without x and so that

$$\text{WE-PA}^\omega \vdash (\lambda x.t)(s) =_\tau t[s/x]$$

for any term s of type ρ.

Next to WE-PA$^\omega$, we define, as usual, the principle of quantifier-free choice QF-AC, that is,[7]

$$\forall \underline{x} \exists \underline{y} A_0(\underline{x}, \underline{y}) \to \exists \underline{Y} \forall \underline{x} A_0(\underline{x}, \underline{Y}\underline{x}) \qquad (\text{QF–AC})$$

with A_0 quantifier-free and where the types of the variable tuples $\underline{x}$, $\underline{y}$ are arbitrary, as well as the principle of dependent choice DC defined as the collection of DC$^{\underline{\rho}}$ for all tuples of types $\underline{\rho}$ with

$$\forall x^0, \underline{y}^{\underline{\rho}} \exists \underline{z}^{\underline{\rho}} A(x, \underline{y}, \underline{z}) \to \exists \underline{f}^{\underline{\rho}(0)} \forall x^0 A(x, \underline{f}(x), \underline{f}(S(x))) \qquad (\text{DC}^{\underline{\rho}})$$

where $\underline{f}^{\underline{\rho}(0)}$ stands for $f_1^{\rho_1(0)}, \ldots, f_k^{\rho_k(0)}$ and A may now be arbitrary.

Over $\mathcal{A}^\omega$, we will have to rely on some chosen representation of the real numbers as a Polish space and for that we follow definitions and conventions given in [26]. In particular, rational numbers are represented using pairs of natural numbers and, in that context, we fix the same pairing function j as in [25]:

$$j(n^0, m^0) := \begin{cases} \min u \leq_0(n+m)^2 + 3n + m[2u =_0 (n+m)^2 + 3n + m] & \text{if existent,} \\ 0^0 & \text{otherwise.} \end{cases}$$

The usual arithmetical operations $+_{\mathbb{Q}}, \cdot_{\mathbb{Q}}, |\cdot|_{\mathbb{Q}}$, etc., are then primitive recursively definable through terms that operate on such codes and the usual relations $=_{\mathbb{Q}}$, $<_{\mathbb{Q}}$, etc., are definable via quantifier-free formulas.

For real numbers we then rely on a representation via fast converging Cauchy sequences of rational numbers with a fixed Cauchy modulus 2^{-n} (see [26] for details), that is, via objects of type 1, and we consider $\mathbb{N}$ and $\mathbb{Q}$ as being embedded in that representation via the constant sequences. Also here, 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., now operating on type 1 objects, are representable by formulas of the underlying language. Naturally, these relations are not decidable anymore but are given by Π^0_1- and Σ^0_1-formulas, respectively.

In the context of this representation of reals, we will later rely on an operator $\widehat{\cdot}$ which allows for an implicit quantification over all such fast-converging Cauchy sequences of rationals. Following [26], we define this operator via

$$\widehat{x}n := \begin{cases} xn & \text{if } \forall k <_0 n \big(|xk -_{\mathbb{Q}} x(k+1)|_{\mathbb{Q}} <_{\mathbb{Q}} 2^{-k-1}\big), \\ xk & \text{for } k <_0 n \text{ least with } |xk -_{\mathbb{Q}} x(k+1)|_{\mathbb{Q}} \geq_{\mathbb{Q}} 2^{-k-1} \text{ otherwise,} \end{cases}$$

turning x of type 1 into a fast-converging Cauchy sequence $\widehat{x}$, and we refer to [26] for any further discussions of its properties.

[7] Here, and in the following, we use the notation $\underline{Y}\underline{x}$ to abbreviate $Y_1\underline{x}, \ldots, Y_k\underline{x}$ for $\underline{Y} = Y_1, \ldots, Y_k$.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

In the context of the bound extraction theorems later on, we will rely on a canonical selection of a Cauchy sequence representing a given real number. Naturally, such an association will be noneffective. However, it will suffice that the operation behaves well enough w.r.t. the notion of majorization. Following [25], this can be achieved for non-negative numbers via the function $(\cdot)_\circ$ defined by

$$(r)_\circ(n) := j(2k_0, 2^{n+1} - 1),$$

where

$$k_0 := \max k \left[\frac{k}{2^{n+1}} \leq r \right].$$

Later, we will need an extension of this function $(\cdot)_\circ$ to all real numbers such that we retain these nice properties regarding majorizability and so, for $r < 0$, we consider $(r)_\circ$ to be defined by[8]

$$(r)_\circ(n) = j(2\bar{k}_0 \dot{-} 1, 2^{n+1} - 1)$$

where

$$\bar{k}_0 := \max k \left[\frac{k}{2^{n+1}} \leq |r| \right].$$

Then $(r)_\circ(n) = -_{\mathbb{Q}}(|r|)_\circ(n)$ and we get the following lemma containing exactly the properties that we later need for this notion to be useful in the context of majorizability (extending Lemma 2.10 from [25]):

Lemma 2.1 (essentially [25, Lemma 2.10], see also [51, Lemma 2.1]). *Let $r \in \mathbb{R}$. Then:*

1. *$(r)_\circ$ is a representation of r in the sense of the above (see again, e.g., [26]).*
2. *For $s \in [0, \infty)$, if $|r| \leq s$, then $(r)_\circ \leq_1 (s)_\circ$, that is, $(r)_\circ(n) \leq (s)_\circ(n)$ for all $n \in \mathbb{N}$.*
3. *$(r)_\circ$ is nondecreasing (as a type 1 function).*

Lastly, we write r_α for the unique real represented by $\widehat{\alpha}$ for a given sequence $\alpha \in \mathbb{N}^{\mathbb{N}}$ and we sometimes write $[\alpha](n)$ for the rational number represented by the n-th element of that sequence for better readability.

In terms of notation, we want to note that, to enhance readability, we will omit the subscripts of the arithmetical operations for $\mathbb{R}$ everywhere. Further, we will omit the operation $\cdot_{\mathbb{R}}$ often altogether. Similarly, we will also omit types of variables whenever convenient. Also, we will almost always omit types or related subscripts in proofs. Lastly, we throughout denote the powerset of a set X by 2^X.

3. Systems for algebras of sets

In this section, we develop the underlying systems on which we will later bootstrap our treatment of probability contents and probability measures as well as of all the other respective extensions discussed before. As such, we begin with a treatment of algebras of sets as the most basic underlying algebraic notion that is essential for the theory of contents. For references on these basic definitions and their properties, if nothing else is mentioned otherwise, we mainly refer to [6].

Definition 3.1 (Algebra of sets). Let Ω be a set and $S \subseteq 2^\Omega$. Then S is called an algebra of sets (or simply an algebra) if $\emptyset \in S$ and for any $A, B \in S$, it holds that $A^c := \Omega \setminus A \in S$ and $A \cup B \in S$.

The approach that we take towards a formal system for algebras of sets is to use abstract types to represent both the underlying ground set Ω as well as the algebra $S \subseteq 2^\Omega$. One then has to restore the structure of S as a collection of subsets over Ω with certain operations on them by including additional constants that reintroduce these operations in this abstract setting.

[8] Here, $\dot{-}$ is defined via $n \dot{-} m := \max\{n - m, 0\}$ for $n, m \in \mathbb{N}$.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

Concretely, to form a system for the treatment of algebras, we extend the previously discussed set of types T by two new abstract types Ω and S, forming the extended set of types $T^{\Omega,S}$ defined by

$$0, \Omega, S \in T^{\Omega,S}, \quad \rho, \tau \in T^{\Omega,S} \to \rho(\tau) \in T^{\Omega,S},$$

and, over the resulting language, we then utilize this augmented set of types to introduce the following new constants to induce the usual structure on the set represented by S in relation to Ω as mentioned before:

- eq of type $0(\Omega)(\Omega)$;
- $\in$ of type $0(S)(\Omega)$;
- $\cup$ of type $S(S)(S)$;
- $(\cdot)^c$ of type $S(S)$;
- $\emptyset$ of type S;
- c_Ω of type Ω.

The constant eq serves as an abstract account of the equality relation between objects of type Ω while the constant $\in$ serves as an abstract account of the element relation between elements as objects of type Ω and sets as objects of type S. The constants $\cup$ and $(\cdot)^c$ reintroduce the respective operations of union and complement for the abstract type S and $\emptyset$ provides a constant representing the empty set. The constant c_Ω in particular is intended to witness that the underlying set Ω is nonempty. We often simply write A^c instead of $(A)^c$ for A of type S. Further, we abbreviate $\in xA =_0 0$ by $x \in A$ and, similarly, we write $x \notin A$ for $\in xA \neq_0 0$. Lastly, we define $\Omega := \emptyset^c$ as a notation for the top element of S.[9] Also regarding notation, we introduce intersections as an abbreviation by defining

$$A \cap B := (A^c \cup B^c)^c$$

for terms A^S, B^S.

We write $x =_\Omega y$ as an abbreviation of $\mathrm{eq}xy =_0 0$ for objects x^Ω, y^Ω. Using $\in$, we introduce equality on S via the following abbreviation: for A^S and B^S, we define

$$A =_S B :\equiv \forall x^\Omega (x \in A \leftrightarrow x \in B).$$

Note that $=_S$ clearly is, provably, an equivalence relation. Furthermore, we introduce the abbreviation

$$A \subseteq_S B :\equiv \forall x^\Omega (x \in A \to x \in B)$$

for A, B of type S and it is straightforward to show that $\subseteq_S$ forms a partial order with respect to equality defined by $=_S$.

For axioms, we first specify that eq represents an equivalence relation:

$$\forall x^\Omega, y^\Omega (\mathrm{eq}xy \leq_0 1), \tag{eq}_1$$

$$\forall x^\Omega, y^\Omega, z^\Omega (x =_\Omega x \wedge (x =_\Omega y \to y =_\Omega x) \wedge (x =_\Omega y \wedge y =_\Omega z \to x =_\Omega z)). \tag{eq}_2$$

Further, we axiomatize that $\in$, as a relation, is bounded by 1 on all inputs and behaves as an element relation regarding the operations of union and complement as well as with respect to the empty set:

$$\forall x^\Omega \forall A^S (\in xA \leq_0 1), \tag{$\in$}_1$$

$$\forall x^\Omega (x \notin \emptyset), \tag{$\in$}_2$$

$$\forall x^\Omega \forall A^S, B^S (x \in A \cup B \leftrightarrow x \in A \vee x \in B), \tag{$\in$}_3$$

$$\forall x^\Omega \forall A^S (x \in A^c \leftrightarrow x \notin A). \tag{$\in$}_4$$

[9] This is not to be confused with the type Ω but the context will make it clear which of these two readings is intended.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

Based on the fact that inclusions of elements x^Ω in elements A^S as facilitated by $\in$ are quantifier-free assertions, the above axioms are (generalized) Π_1-sentences and so they are in particular immediately admissible in the context of bound extraction theorems based on the Dialectica interpretation (as will be discussed later in more detail).

Definition 3.2. We write $\mathcal{F}^\omega$ for the system resulting from $\mathcal{A}^\omega$ over the augmented language including the types Ω, S (where all the respective constants and axioms now are allowed to also refer to these new types, if applicable) by extending this system with the constants $\mathrm{eq}, \in, \cup, (\cdot)^c, \emptyset, c_\Omega$ and the axioms $(\mathrm{eq})_1 - (\mathrm{eq})_2$ as well as $(\in)_1 - (\in)_4$.

We now begin by showing some basic properties of the above operations on algebras provable in this system $\mathcal{F}^\omega$ which, for one, amount to deriving the essential algebraic properties of S as a subalgebra of the full Boolean algebra of the power set of Ω. Further, for another, all algebraic operations on S behave in a provably extensional way.

Proposition 3.3. *The operations $\cup$ and $(\cdot)^c$ are provably extensional in $\mathcal{F}^\omega$, that is, $\mathcal{F}^\omega$ proves:*

1. $\forall A^S, A'^S, B^S, B'^S (A =_S A' \wedge B =_S B' \to A \cup B =_S A' \cup B')$,
2. $\forall A^S, A'^S (A =_S A' \to A^c =_S A'^c)$.

Further, all the axioms of Boolean algebras, instantiated using $\cap, \cup, (\cdot)^c$, and $=_S$, can be derived in $\mathcal{F}^\omega$. Lastly, over $\mathcal{F}^\omega$, $A \subseteq_S B$ is equivalent to both $A =_S B \cap A$ and $B =_S A \cup B$ for terms A^S, B^S.

Proof. We only show items (i) and (ii) as these illustrate the style of proof that one typically follows in $\mathcal{F}^\omega$ to reason about the algebraic structure of S. The identities of Boolean algebras and the equivalent formulations of the order in terms of meet and join are then easily derived from the axioms $(\in)_2, \dots, (\in)_4$.

1. Fix A, A', B, B' and assume $A = A'$ as well as $B = B'$. We need to show that
$$\forall x (x \in A \cup B \leftrightarrow x \in A' \cup B').$$
Let x be arbitrary. Then $x \in A \cup B$ is equivalent to $x \in A \vee x \in B$ by $(\in)_3$. By assumption of $A = A'$ and $B = B'$, we have that $x \in A \vee x \in B$ is equivalent to $x \in A' \vee x \in B'$ and so to $x \in A' \cup B'$ by $(\in)_3$, which yields the claim.
2. Fix A, A' and assume $A = A'$. We need to show that
$$\forall x (x \in A^c \leftrightarrow x \in A'^c).$$
Let x be arbitrary. Then $x \in A^c$ is equivalent to $x \notin A$ by $(\in)_4$. By assumption of $A = A'$, we have that $x \notin A$ is equivalent to $x \notin A'$ and thus to $x \in A'^c$ again by $(\in)_4$. □

Remark 3.4. Also the constants eq and $\in$ are immediately provably extensional in $\mathcal{F}^\omega$, as can be shown using the quantifier-free extensionality rule.

Using the recursor constants of the underlying language of $\mathcal{F}^\omega$ in combination with the union operation $\cup$ immediately allows one to also talk about arbitrary finite unions. Concretely, given a sequence of events $A^{S(0)}$ and two natural numbers $n^0 \leq_0 m^0$, we use the abbreviation

$$\bigcup_{i=n}^{m} A(i) := R_S(m - n, A(n), \lambda B, x.(B \cup A(n + x + 1)))$$

where R_S is a (single) type S recursor constant. For $m <_0 n$, we simply set $\bigcup_{i=n}^m A(i) := \emptyset$. We then dually write

$$\bigcap_{i=n}^{m} A(i) := \left(\bigcup_{i=n}^{m} (A(i))^c \right)^c.$$

It is easy to show by induction that the previous extensionality result for $\cup$ extends to these finite unions.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

4. Systems for contents on algebras of sets

We now augment the previous system $\mathcal{F}^\omega$ for the treatment of algebras so that we arrive at a system suitable for treating proofs from the theory of probability contents in the sense of the following definition:[10]

Definition 4.1 (Contents). Let Ω be a set and $S \subseteq 2^\Omega$ be an algebra. A content on S is a mapping $\mu : S \to [0, \infty]$ such that $\mu(\emptyset) = 0$ and $\mu(A \cup B) = \mu(A) + \mu(B)$ for $A, B \in S$ with $A \cap B = \emptyset$.

We say that μ is a probability content if $\mu(\Omega) = 1$.

We mainly denote probability contents by the symbol $\mathbb{P}$. Again, we mainly refer to [6] as a standard reference for the theory of contents.

The concrete approach that we now take for a formal system for probability contents on algebras is to introduce an additional constant

- $\mathbb{P}$ of type $1(S)$

to the language of the system $\mathcal{F}^\omega$. The first defining properties of $\mathbb{P}$ as a probability content are then easily formalized in the underlying language as

$$\forall A^S (0 \leq_{\mathbb{R}} \mathbb{P}(A) \leq_{\mathbb{R}} 1), \tag{$\mathbb{P}$}_1$$

$$\mathbb{P}(\emptyset) =_{\mathbb{R}} 0. \tag{$\mathbb{P}$}_2$$

These statements $(\mathbb{P})_1$ and $(\mathbb{P})_2$ are again purely universal statements and therefore immediately admissible in the context of metatheorems based on the (monotone) functional interpretation.

The last property of $\mathbb{P}$, that is, additivity, if formalized naively via

$$\forall A^S, B^S (A \cap B =_S \emptyset \to \mathbb{P}(A \cup B) =_{\mathbb{R}} \mathbb{P}(A) + \mathbb{P}(B)),$$

is not purely universal based on the internal definition of $=_S$ and is instead equivalent over $\mathcal{F}^\omega$ (extended with the constant $\mathbb{P}$) to the following generalized Π_3-sentence:

$$\forall A^S, B^S \exists x^\Omega (x \notin A \cap B \to \mathbb{P}(A \cup B) =_{\mathbb{R}} \mathbb{P}(A) + \mathbb{P}(B)).$$

Similar to how the class of so-called type Δ sentences is treated in, for example, [16], it is clear that this statement would be admissible in the context of bound extraction theorems based on the monotone Dialectica interpretation if the x could be conceived of as being bounded in a suitable sense relative to A and B. Now, as a matter of fact, a crucial perspective for our formal approach to deriving bound extraction theorems for these systems for algebras and probability contents will be that the whole space Ω can be naturally regarded as uniformly bounded. In the context of a corresponding suitable extension of the notion of majorizability to Ω which reflects this perspective via assuming that there is a uniform majorant for all x^Ω, the above axiom actually has a trivial monotone functional interpretation and is thus admissible in the context of the approach to proof mining metatheorems via such a variant of the Dialectica interpretation. This will be discussed in full formal detail later on so that here, for now, we are content with just considering quantification over Ω as "bounded" and so as "proof-theoretically harmless."

However, if we were to admit the above sentence as the sole axiom, we would be tasked with deriving all the other properties of $\mathbb{P}$ from this axiom, including monotonicity and thus extensionality of the content as a function, which would require many subtle manipulations of various equalities using the quantifier-free extensionality rule. We therefore instead opt for the following axiomatization which eases

[10] While contents are also studied over much sparser structures than algebras of sets, we here only consider the case of a content defined on such an algebra.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

the formal development of these properties in the resulting system: For one, instead of additivity, we include the following generalized additivity law which holds for probability contents on algebras:

$$\forall A^S, B^S(\mathbb{P}(A \cup B) =_{\mathbb{R}} \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)). \tag{$\mathbb{P}$}_3$$

This statement is purely universal and thus immediately admissible in the context of our approach to bound extraction theorems as before. The other property of $\mathbb{P}$ that we then axiomatically add is that of the monotonicity of $\mathbb{P}$, that is,

$$\forall A^S, B^S(A \subseteq_S B \to \mathbb{P}(A) \leq_{\mathbb{R}} \mathbb{P}(B)).$$

Similar to the above, this statement is equivalent to the following (generalized) Π_3-statement

$$\forall A^S, B^S \exists x^{\Omega}(\mathbb{P}(A) >_{\mathbb{R}} \mathbb{P}(B) \to x \in A \wedge x \notin B) \tag{$\mathbb{P}$}_4$$

which in the context of the previously sketched extended notion of majorizability will later have a trivial monotone functional interpretation and thus be admissible in the context of our approach to proof mining metatheorems.

Adding these statements as axioms to the underlying system for algebras of sets, we derive the following system for probability contents on such algebras:

Definition 4.2. We write $\mathcal{F}^{\omega}[\mathbb{P}]$ for the system resulting from $\mathcal{F}^{\omega}$ by adding the above constant $\mathbb{P}$ together with the axioms $(\mathbb{P})_1 - (\mathbb{P})_4$.

We now begin with some immediate properties of $\mathbb{P}$ provable in the above system.

Proposition 4.3. *The following properties of* $\mathbb{P}$ *are provable in* $\mathcal{F}^{\omega}[\mathbb{P}]$*:*

1. $\mathbb{P}$ *is extensional w.r.t.* $=_S$ *and* $=_{\mathbb{R}}$*, that is,*

$$\forall A^S, B^S(A =_S B \to \mathbb{P}(A) =_{\mathbb{R}} \mathbb{P}(B)).$$

2. $\mathbb{P}$ *is definite on* $\emptyset$*, that is,*

$$\forall A^S(\mathbb{P}(A) >_{\mathbb{R}} 0 \to A \neq_S \emptyset).$$

3. $\mathbb{P}$ *is additive, that is,*

$$\forall A^S, B^S(A \cap B =_S \emptyset \to \mathbb{P}(A \cup B) =_{\mathbb{R}} \mathbb{P}(A) + \mathbb{P}(B)).$$

4. $\mathbb{P}$ *respects the relative complements of subsets, that is,*

$$\forall A^S, B^S(B \subseteq_S A \to \mathbb{P}(A \cap B^c) =_{\mathbb{R}} \mathbb{P}(A) - \mathbb{P}(B)).$$

In particular, we also have

$$\forall A^S(\mathbb{P}(A^c) =_{\mathbb{R}} 1 - \mathbb{P}(A)).$$

5. $\mathbb{P}$ *satisfies Boole's inequality, that is,*

$$\forall A^{S(0)}, n^0\left(\mathbb{P}\left(\bigcup_{i=0}^{n} A(i)\right) \leq_{\mathbb{R}} \sum_{i=0}^{n} \mathbb{P}(A(i))\right).$$

Proof.

1. Assume $\mathbb{P}(A) > \mathbb{P}(B)$. By axiom $(\mathbb{P})_4$, there exists an x such that $x \in A$ and $x \notin B$, that is, $A \neq B$. Similarly we derive $A \neq B$ from $\mathbb{P}(A) < \mathbb{P}(B)$. Combined, we get that $A = B$ implies $\mathbb{P}(A) = \mathbb{P}(B)$.
2. Assume $\mathbb{P}(A) > 0 = \mathbb{P}(\emptyset)$. Then by axiom $(\mathbb{P})_4$, we get an $x \in A$, that is, $A \neq \emptyset$.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

3. Let A, B be arbitrary with $A \cap B = \emptyset$. We have $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ by axiom $(\mathbb{P})_3$. As $\mathbb{P}$ is extensional, we get $\mathbb{P}(A \cap B) = \mathbb{P}(\emptyset) = 0$ so that the above implies $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ as desired.
4. Let $E := A \cap B$ and $F := A \cap B^c$. Then $E \cap F = \emptyset$ (by the properties of algebras of sets). Thus, by additivity, $\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F)$. We have that $E \cup F = A$ (again by the properties of algebras of sets). Thus, by extensionality of $\mathbb{P}$, we have $\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c)$. Now, $B \subseteq A$ implies $A \cap B = B$ by Proposition 3.3 and so the result follows from the extensionality of $\mathbb{P}$.
5. This follows via a simple induction from the axiom $(\mathbb{P})_3$. □

Contents on algebras enjoy certain continuity properties similar to continuity from above and below for measures but without the existence of limiting sets, that is, infinite unions, etc. (see, e.g., [6]) and we now discuss how the system $\mathcal{F}^\omega[\mathbb{P}]$ recognizes Cauchy-variants of these properties.

For that, we introduce 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 into a sequence of disjoint sets $A\uparrow$ with the same (partial) union(s) and if A was already a disjoint family, then it is left unchanged by the operation.

We now begin with a Cauchy-type form of σ-additivity of $\mathbb{P}$ as a content. For this, note that for a given $A^{S(0)}$, the sequence of partial sums

$$\sum_{i=0}^{n} \mathbb{P}((A\uparrow)(i)) = \mathbb{P}\left(\bigcup_{i=0}^{n} (A\uparrow)(i)\right) = \mathbb{P}\left(\bigcup_{i=0}^{n} A(i)\right)$$

is a monotone and bounded sequence of real numbers and thus is Cauchy. The following result that (already a weak fragment of) $\mathsf{WE\text{-}PA}^\omega$ suffices to prove the Cauchy formulation of the convergence of monotone and bounded sequences is well known:

Lemma 4.4 (folklore, see essentially [26]). *The system $\mathsf{WE\text{-}PA}^\omega$ (and actually already a weak fragment thereof) proves that*

$$\forall a^{1(0)}\Big(\forall n^0(0 \leq_{\mathbb{R}} a(n) \leq_{\mathbb{R}} 1 \wedge a(n) \leq_{\mathbb{R}} a(n+1)) \\ \to \forall k^0 \exists N^0 \forall n^0, m^0 \geq_0 N\Big(|a(n) - a(m)| <_{\mathbb{R}} 2^{-k}\Big)\Big).$$

So, instantiating the above result with $a(n) = \sum_{i=0}^{n} \mathbb{P}((A\uparrow)(i))$, we can derive that $\mathcal{F}^\omega[\mathbb{P}]$ (and actually already a weak fragment thereof) can prove the Cauchy property of sequences of contents of increasing disjoint unions:

Proposition 4.5. *The system $\mathcal{F}^\omega[\mathbb{P}]$ (and actually already a weak fragment thereof) proves*

$$\forall A^{S(0)} \forall k^0 \exists N^0 \forall n^0, m^0 \geq_0 N\left(\left|\sum_{i=0}^{n} \mathbb{P}((A\uparrow)(i)) - \sum_{i=0}^{m} \mathbb{P}((A\uparrow)(i))\right| <_{\mathbb{R}} 2^{-k}\right).$$

From this Proposition 4.5, we can then immediately derive the following continuity theorems for contents.

Proposition 4.6. *The system $\mathcal{F}^\omega[\mathbb{P}]$ (and actually already a weak fragment thereof) proves:*

1. *$\mathbb{P}$ is continuous from below, that is,*

$$\forall A^{S(0)}\Big(\forall n^0(A(n) \subseteq_S A(n+1)) \to \forall k^0 \exists N^0 \forall n^0, m^0 \geq_0 N\Big(|\mathbb{P}(A(n)) - \mathbb{P}(A(m))| <_{\mathbb{R}} 2^{-k}\Big)\Big).$$

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

2. $\mathbb{P}$ *is continuous from above, that is,*

$$\forall A^{S(0)}\Big(\forall n^0(A(n+1) \subseteq_S A(n)) \to \forall k^0 \exists N^0 \forall n^0, m^0 \geq_0 N\Big(|\mathbb{P}(A(n)) - \mathbb{P}(A(m))| <_{\mathbb{R}} 2^{-k}\Big)\Big).$$

Proof.

1. Note that $A(n) \subseteq A(n+1)$ for all n implies that

$$(A\uparrow)(n+1) = A(n+1) \cap A(n)^c$$

for any n. Thus by Proposition 4.3, (4), we have

$$\mathbb{P}((A\uparrow)(n+1)) = \mathbb{P}(A(n+1)) - \mathbb{P}(A(n))$$

for all n. Thus we have

$$\sum_{i=0}^{n} \mathbb{P}((A\uparrow)(i)) = \mathbb{P}(A(0)) + \sum_{i=0}^{n-1} (\mathbb{P}(A(i+1)) - \mathbb{P}(A(i))) = \mathbb{P}(A(n))$$

for any n. The result now follows from Proposition 4.5.

2. Observe that $A(n+1) \subseteq A(n)$ for any n implies that $A(n)^c \subseteq A(n+1)^c$ for all n. Thus, by (1), we have

$$\forall k^0 \exists N^0 \forall n^0, m^0 \geq_0 N\Big(|\mathbb{P}(A(n)^c) - \mathbb{P}(A(m)^c)| <_{\mathbb{R}} 2^{-k}\Big).$$

By Proposition 4.3, (4), we get $\mathbb{P}(A(n)^c) = 1 - \mathbb{P}(A(n))$ and from this the result follows. □

5. Systems for σ-algebras and probability measures

If we now further require the closure of the underlying algebra of sets under countable unions, we arrive at the notion of a σ-algebra which forms the algebraic basis for probability measures.

Definition 5.1 (σ-algebra). Let Ω be a set and $S \subseteq 2^\Omega$ be an algebra. Then S is called a σ-algebra if for any $(A_n) \subseteq S$, it also holds that $\bigcup_{n=0}^{\infty} A_n \in S$.

The requirement that a content on a σ-algebra is also well-behaved w.r.t. these countable unions then leads to the notion of a measure on such an algebra.

Definition 5.2 (Measure). Let Ω be a set and $S \subseteq 2^\Omega$ be a σ-algebra. A measure on S is a content $\mu : S \to [0, \infty]$ that is also σ-additive, that is,

$$\mu\left(\bigcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} \mu(A_n)$$

for any $(A_n) \subseteq S$ with $A_i \cap A_j = \emptyset$ for $i \neq j$.

The map μ is called a probability measure if $\mu(\Omega) = 1$. In that case, (Ω, S, μ) is called a probability space.

In this section, we will now discuss how the previous system for algebras $\mathcal{F}^\omega$ and its extension $\mathcal{F}^\omega[\mathbb{P}]$ for treating probability contents can be augmented by a certain intensional treatment of countably infinite unions to provide an apt and tame formal basis for these notions.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

5.1. Treating infinite unions tamely

Concretely, to treat countably infinite unions over algebras of sets tamely, we now extend the previous system $\mathcal{F}^\omega$ with the following further constant

$\circ$ $\bigcup$ of type $S(S(0))$,

providing a term of type S for the resulting union of the sequence of sets coded by the input of type $S(0)$. So, in the context of suitable axioms specifying that $\bigcup A$ for a given $A^{S(0)}$ represents the union of all $A(n)$, we can formally induce that the algebra S is closed under these countable unions. The immediate axioms specifying the property that $\bigcup A$ is the corresponding union are

$$\forall A^{S(0)} \forall n^0 \Big(A(n) \subseteq_S \bigcup A \Big) \tag{$\bigcup_1$}$$

as well as

$$\forall A^{S(0)} \forall B^S \Big(\forall n^0 (A(n) \subseteq_S B) \to \bigcup A \subseteq_S B \Big),$$

specifying that $\bigcup$ is the join of the elements $A(n)$ in S (seen as a lattice). The first statement $(\bigcup)_1$ is immediately admissible in the context of the Dialectica interpretation as it is purely universal. The latter statement is naturally not admissible as is in the context of the usual approach to proof mining metatheorems as it contains a negative universal quantifier of type 0 that cannot be majorized (which, after all, is also why a uniform variant of arithmetical comprehension is necessary to fully define countable unions of sets, see, e.g., [35] for further discussions).

Since we want to avoid this strong form of comprehension as to not in general distort the strength of the extracted bounds to be able to a priori guarantee the extractability of proof-theoretically tame bounds from proofs, we opt for the next best thing we can do and instead specify the union only intensionally by adding the following rule-variant of the above converse implication

$$\frac{F_{qf} \to \forall n^0 (A(n) \subseteq_S B)}{F_{qf} \to \bigcup A \subseteq_S B} \tag{$\bigcup_2$}$$

where A is a term of type $S(0)$, B is a term of type S and F_{qf} is a quantifier-free formula. So: If $A(n)$ is provably bounded above by B w.r.t. $\subseteq_S$ under some quantifier-free assumptions, then also $\bigcup A$ is provably bounded above by B w.r.t. $\subseteq_S$ under the same assumptions.

Definition 5.3. We write $\mathcal{F}^\omega[\bigcup]$ for the system resulting from $\mathcal{F}^\omega$ by extending it with the constant $\bigcup$ together with the axiom $(\bigcup)_1$ and the rule $(\bigcup)_2$. Similarly, we write $\mathcal{F}^\omega[\bigcup, \mathbb{P}]$ for the system that arises from $\mathcal{F}^\omega[\mathbb{P}]$ by adding the same constants, axioms, and rules.

Using this intensional approach to countable unions, we can also immediately provide an intensional treatment of countable intersections. For this, we first define A^c for an $A^{S(0)}$ by setting $A^c(n) := A(n)^c$ for any n^0. Using this notation, we then define the countable intersection of a collection of sets represented by an $A^{S(0)}$ via

$$\bigcap A := \Big(\bigcup A^c \Big)^c.$$

We then get that analogs of the axiom $(\bigcup)_1$ and the rule $(\bigcup)_2$, formulated appropriately for the intersection, are provable in our system $\mathcal{F}^\omega[\bigcup]$:

Lemma 5.4. *The following statement is provable in* $\mathcal{F}^\omega[\bigcup]$*:*

$$\forall A^{S(0)} \forall n^0 \Big(\bigcap A \subseteq_S A(n) \Big).$$

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

Further, in $\mathcal{F}^\omega[\bigcup]$ the following rule is derivable:

$$\frac{F_{qf} \to \forall n^0(B \subseteq_S A(n))}{F_{qf} \to B \subseteq_S \bigcap A}.$$

Proof. For the provability of the first statement, let $x \in \bigcap A$. By definition we have $x \notin \bigcup A^c$. Then by axiom $(\bigcup)_1$, we get $x \notin A(n)^c$, that is, $x \in A(n)$ for any n.

Now, for the rule, suppose that we provably have $F_{qf} \to \forall n(B \subseteq A(n))$. Then we also provably have $F_{qf} \to \forall n(A(n)^c \subseteq B^c)$ and using the rule $(\bigcup)_2$, we get $F_{qf} \to \bigcup A^c \subseteq B^c$. Thus also $F_{qf} \to B \subseteq (\bigcup A^c)^c = \bigcap A$ as desired. □

5.2. *Handling probability measures*

As we have seen in Propositions 4.5 and 4.6, the system $\mathcal{F}^\omega[\mathbb{P}]$ already provides Cauchy-variants of the convergence of monotone sequences of events as well as of sums of disjoint events. In the theory of measures on σ-algebras, the resulting limits of course correspond to the measure of respective infinite unions or intersections. Thus, the natural question of whether and how this can be formally represented in the system $\mathcal{F}^\omega[\bigcup, \mathbb{P}]$ immediately arises. And while for a disjoint family represented by $A^{S(0)}$, the limit

$$0 \leq \mathbb{P}\left(\bigcup A\right) - \mathbb{P}\left(\bigcup_{i=0}^{n} A(i)\right) \to 0 \text{ for } n \to \infty$$

holds true, there is in general no computable rate of convergence for this expression in the sense that even a function φ of type 1 such that

$$\forall k^0 \exists n \leq_0 \varphi(k)\left(\mathbb{P}\left(\bigcup A\right) - \mathbb{P}\left(\bigcup_{i=0}^{n} A(i)\right) \leq_{\mathbb{R}} 2^{-k}\right) \tag{$\circ$}$$

is in general not computable (see Remark 5.5 for an example). Nevertheless, we want to point out that while therefore the convergence of the sequence $\sum_{i=0}^{n} \mathbb{P}(A(i))$ towards $\mathbb{P}(\bigcup A)$ cannot be provable in any system that allows for the extraction of computable and uniform bounds, the system $\mathcal{F}^\omega[\bigcup, \mathbb{P}]$ still provides an intensional version of that convergence in the following sense:

1. By Proposition 4.5, the sequence of partial sums $\sum_{i=0}^{n} \mathbb{P}(A(i))$ is provably Cauchy.
2. Using the additivity and monotonicity of $\mathbb{P}$, that is, axioms $(\mathbb{P})_3$ and $(\mathbb{P})_4$, we get by $\bigcup_{i=0}^{n} A(i) \subseteq_S \bigcup A$ that

$$\sum_{i=0}^{n} \mathbb{P}(A(i)) =_{\mathbb{R}} \mathbb{P}\left(\bigcup_{i=0}^{n} A(i)\right) \leq_{\mathbb{R}} \mathbb{P}\left(\bigcup A\right)$$

   holds provably.
3. For any object B^S such that we provably have $\forall n^0(A(n) \subseteq_S B)$, we get $\bigcup A \subseteq_S B$ using the rule $(\bigcup)_2$ and so we get provably

$$\mathbb{P}\left(\bigcup A\right) \leq_{\mathbb{R}} \mathbb{P}(B)$$

   in that case by monotonicity of $\mathbb{P}$.

So the value $\mathbb{P}(\bigcup A)$ is at least intensionally specified to be the limit of the partial sums $\sum_{i=0}^{n} \mathbb{P}(A(i))$ as $\mathbb{P}(\bigcup A)$ is bounded below by this nondecreasing sequence of partial sums and intensionally bounded above by the probability of any set which provably sits above the given partial unions $\bigcup_{i=0}^{n} A(i)$.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

The case that we want to make is now twofold: For one, as mentioned in the introduction, the theory of contents already exhausts large parts of the theory of measures in the sense that often already the properties of contents on algebras suffice to carry out proofs for properties of measures on σ-algebras (as will also be the case in the applications discussed later). For another, we want to argue that even in situations where one cannot do just with finite unions and content, such an intensional specification of countable unions and their measures might suffice for formalizing a given proof and all the while guaranteeing the extractability of tame bounds a priori. If that is not the case, then the result under consideration might be considered to be inherently "untame" and a full treatment of the comprehension principle needed to define the respective unions could be necessary. We therefore regard $\mathcal{F}^{\omega}[\bigcup, \mathbb{P}]$ as a suitable tame base system for treating probability measures on σ-algebras.

Remark 5.5. For an example where φ in $(\circ)$ is not computable, we take $(r_i) \subseteq (0,1)$ to be a sequence of computable real numbers such that

$$a_n = \sum_{i=0}^{n} r_i \to a \leq 1$$

without a computable rate of convergence.[11] We now define $\Omega = \mathbb{N} \cup \{\infty\}$ as well as $S = 2^{\Omega}$. On the discrete sample space Ω, we then define a probability mass function p via $p(i) = r_i$ for $i \in \mathbb{N}$ as well as $p(\infty) = 1 - a$. This p then as usual induces a probability measure $\mathbb{P}$ on the σ-algebra S defined for $A \subseteq \Omega$ via

$$\mathbb{P}(A) = \sum_{a \in A} p(a).$$

Clearly $(\Omega, S, \mathbb{P})$ is a probability space where

$$\mathbb{P}(\mathbb{N}) - \sum_{i=1}^{n+1} \mathbb{P}(\{i\}) = a - \sum_{i=0}^{n} r_i = a - a_n$$

cannot have a computable rate of convergence to 0.

6. Intensional intervals, inverse mappings and measurable functions

In this section, we now extend the machinery of the previous logical systems so that we are able to deal with functions $f : \Omega \to \mathbb{R}$ that are measurable in a certain suitable sense relative to algebras. As such, the treatment given here will be instrumental for our approach to integrable functions in Section 7, for the applications discussed in Section 9, and for the proof-theoretic transfer principles for implications between modes of convergence in Section 10. For this, we now first recall the essential definitions and basic results.

Definition 6.1 (Borel σ-algebra). Let X be a topological space. The Borel σ-algebra $\mathcal{B}(X)$ on X is the smallest σ-algebra on X that contains all open subsets of X.

We refer to [17] as a standard reference for Borel σ-algebras in particular and measure theory in general (in particular regarding the well-definedness of the above definition for which one needs to see that the intersection of any family of (σ-)algebras is again a (σ-)algebra).

Crucial for us will be the notion of a generating set of a (σ-)algebra.

[11] The existence of such a sequence is due to [57]. For a direct construction, we proceed similarly to [61]: pick an enumeration $f : \mathbb{N} \to \mathbb{N}$ of the special Halting Problem without repetitions. Then defining $r_i = 2^{-f(i)-1}$ yields a suitable sequence such that a_n defined as above naturally converges to an element $a \in [0,1]$ as it is monotone and bounded above but the rate of convergence cannot be computable as this would allow one to decide the special Halting Problem.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

Definition 6.2 (Generators of a (σ-)algebra). Let Ω be a set and $S \subseteq 2^\Omega$ be a (σ-)algebra. A generating set for S is a set $S_0 \subseteq 2^\Omega$ such that S is the smallest (σ-)algebra containing S_0.

In that terminology, the Borel σ-algebra $\mathcal{B}(X)$ is the σ-algebra generated by the open subsets of the underlying topological space.

For the special case of the real numbers as a topological space with the usual topology induced by the metric distance, we in particular get the following canonical generators besides the open subsets of $\mathbb{R}$.

Lemma 6.3 (folklore, see, e.g., [17]). *The Borel σ-algebra $\mathcal{B}(\mathbb{R})$ on $\mathbb{R}$ is generated by the collection of all closed intervals $\{[a, b] \mid a, b \in \mathbb{R}\}$.*

One then arrives at the notion of a measurable function which we for simplicity only formulate for real-valued functions here.

Definition 6.4 (Borel-measurable function). Let (Ω, S, μ) be a content space. A function $f : \Omega \to \mathbb{R}$ is called Borel-measurable if

$$f^{-1}(B) := \{x \in \Omega \mid f(x) \in B\} \in S$$

for all $B \in \mathcal{B}(\mathbb{R})$.

As we will crucially use later on, Borel-measurability is simply characterized by a similar condition on a generating set.

Proposition 6.5 (folklore, see, e.g., [17]). *Let (Ω, S, μ) be a measure space. A function $f : \Omega \to \mathbb{R}$ is Borel-measurable if, and only if, $f^{-1}([a, b]) \in S$ for all $a, b \in \mathbb{R}$.*

If the underlying algebra S is not a σ-algebra and if μ is only a content, then the requirement that the preimages of the above collection of intervals are included is still statable but might result in something weaker than full Borel-measurability. We call a function $f : \Omega \to \mathbb{R}$ where the preimages of all intervals $[a, b]$ for $a, b \in \mathbb{R}$ are included in a corresponding algebra $S \subseteq 2^\Omega$ to be *weakly Borel-measurable*. It will in particular be this notion of weakly Borel-measurable functions that we will rely on later in the context of our approach towards Lebesgue integrals for probability contents. It should be noted that by requiring the inclusion $f^{-1}([a, b]) \in S$ for two real numbers $a, b \in \mathbb{R}$ and an algebra S, we in particular also obtain that $f^{-1}([a, b)) = f^{-1}([a, b]) \cap \left(f^{-1}([b, b])\right)^c \in S$ as S is an algebra.[12]

To formally deal with the notion of (weak) Borel-measurability, we thus need access to the collection of the closed intervals $[a, b]$ for $a, b \in \mathbb{R}$ generating the Borel-algebra. For this, we will introduce an intensional approach to real intervals in the next subsection to provide formal means of operating on these generators. These intensional variants of real intervals can then be processed by a general type of inverse map using which we will be able to state the measurability of a function formally.

6.1. Intensional Intervals

Concretely, we now provide a quantifier-free (and thus in a way intensional) account of the closed intervals $[a, b]$ (and thus also of the half-open intervals $[a, b)$ as discussed above) by introducing a further constant to the language:

- $[\cdot, \cdot]$ of type $0(1)(1)(1)$.

Given two inputs a^1, b^1, this function shall return a characteristic function for an intensional representation of the corresponding interval. For this, we add the following axioms:

$$\forall a^1, b^1, r^1([a, b](r) \leq_0 1), \qquad ([\cdot,\cdot])_1$$

$$\forall a^1, b^1, r^1(r \in [a, b] \to a \leq_{\mathbb{R}} r \leq_{\mathbb{R}} b), \qquad ([\cdot,\cdot])_2$$

[12] Note that in the context of algebras, this is a particular benefit from working with closed intervals in the above notion of weak Borel-measurability. If, for example, one instead were to work with half-open intervals $[a, b)$, then defining the closed intervals requires the use of a countably infinite intersection via $[a, b] = \bigcap_{k \in \mathbb{N}} [a, b + 2^{-k})$ which can only be sustained in σ-algebras.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

$$\forall a^1, b^1, r^1(a <_{\mathbb{R}} r <_{\mathbb{R}} b \to r \in [a,b]), \tag{$([\cdot,\cdot])_3$}$$

$$\forall a^1, b^1(a, b \in [a,b]). \tag{$([\cdot,\cdot])_4$}$$

Here, we wrote $[a,b]$ for $[\cdot,\cdot]ab$ as well as $r \in [a,b]$ for $[a,b](r) =_0 0$. Note that this is an intensional representation of the set as we have $a, b \in [a,b]$ but we cannot conclude from $r = a$ or $r = b$ that $r \in [a,b]$.

We also introduce the following notation used later: if we are given a system $\mathcal{C}^\omega$, we write $\mathcal{C}^\omega[\mathrm{Int}]$ to denote the extension of $\mathcal{C}^\omega$ by the above constant $[\cdot,\cdot]$ and the axioms $([\cdot,\cdot])_1 - ([\cdot,\cdot])_4$ for treating the closed intervals.

In similarity to the discussion above, these closed intervals then also provide a quantifier-free access to the half-open intervals in the following way: We define $[\cdot,\cdot)$ of type $0(1)(1)(1)$ via

$$[\cdot,\cdot) := \lambda a^1, b^1, r^1.\, \max\{[a,b](r), 1 - [b,b](r)\}.$$

Also for $[\cdot,\cdot)$, we write $[a,b)$ for $[\cdot,\cdot)ab$ as well as $r \in [a,b)$ for $[a,b)(r) =_0 0$.

In the context of that definition, we in particular obtain relatively immediately that $[\cdot,\cdot)$ defined as such satisfies properties that intensionally specify the half-open intervals similar to how we have specified closed intervals above with the axioms $([\cdot,\cdot])_1 - ([\cdot,\cdot])_4$, that is, we for one have $a \in [a,b)$ but from $r = a$, we cannot infer $r \in [a,b)$ and we have $b \notin [a,b)$ but from $r = b$, we cannot infer $r \notin [a,b)$. This is collected in the following lemma:

Lemma 6.6. *The system* $\mathcal{A}^\omega[\mathrm{Int}]$ *proves the following properties of* $[\cdot,\cdot)$*:*

1. $\forall a^1, b^1, r^1([a,b)(r) \leq_0 1)$,
2. $\forall a^1, b^1, r^1(r \in [a,b) \to a \leq_{\mathbb{R}} r <_{\mathbb{R}} b)$,
3. $\forall a^1, b^1, r^1(a <_{\mathbb{R}} r <_{\mathbb{R}} b \to r \in [a,b))$,
4. $\forall a^1, b^1(a <_{\mathbb{R}} b \to a \in [a,b) \wedge b \notin [a,b))$.

We omit the proof as it is rather immediate.

6.2. The inverse map

We now provide a treatment of the inverse map for a given function f of type $1(\Omega)$. For this, we actually introduce a uniform operator into the language via a constant

- $(\cdot)^{-1}$ of type $0(\Omega)(0(1))(1(\Omega))$

that provides an inverse map for any given function $f^{1(\Omega)}$ in the sense that, writing f^{-1} for $(\cdot)^{-1}f$, the functional f^{-1} receives a subset of the reals coded via a characteristic function $A^{0(1)}$ and maps this into a characteristic function $f^{-1}A$ of type $0(\Omega)$ coding a subset of the underlying space Ω.

This type of map is then governed by the following two axioms:

$$\forall f^{1(\Omega)} \forall A^{0(1)} \forall x^{\Omega} \Big(f^{-1}Ax \leq_0 1 \Big), \tag{Inv_1}$$

$$\forall f^{1(\Omega)} \forall A^{0(1)} \forall x^{\Omega} \Big(x \in f^{-1}A \leftrightarrow f(x) \in A \Big). \tag{Inv_2}$$

Here, we wrote $f(x) \in A$ for $Af(x) =_0 0$ and $x \in f^{-1}A$ for $f^{-1}Ax =_0 0$.

Also for this type of extension, we introduce the following generic notation: given a system $\mathcal{C}^\omega$, we write $\mathcal{C}^\omega[\mathrm{Inv}]$ to denote the extensions of $\mathcal{C}^\omega$ by the constant $(\cdot)^{-1}$ and the above axioms for treating the inverse map.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

6.3. Measurability of functions

In the context of the intensional representations for closed intervals, generating the Borel σ-algebra on the reals, as well as using the general inverse map introduced before, we are now in the position of formulating the (weak) Borel-measurability of a function $f^{1(\Omega)}$ formally in the underlying language by

$$\forall a^1, b^1 \exists A^S \forall x^{\Omega}\Big(x \in f^{-1}([a,b]) \leftrightarrow x \in A\Big).$$

The inner matrix (based on the fact that stating element relations via $\in$ is quantifier-free and as we use the abbreviation $x \in f^{-1}([a,b])$ for $f^{-1}[a,b]x =_0 0$ as introduced in Section 6.2) is quantifier-free and thus the above sentence is a generalized Π_3-statement. Similar to the monotonicity statement of the content $\mathbb{P}$, this statement would be admissible in the context of the monotone Dialectica interpretation if the quantification over elements of type S could be conceived of as being bounded in a suitable sense. Similar to how we argued with the monotonicity of $\mathbb{P}$, a crucial perspective for our formal approach to bound extraction theorems later on will be that, besides the whole space Ω, we will also be able to regard the space S as uniformly bounded, formally encapsulated via a corresponding suitable extension of the notion of majorizability to S used later. In that context, such a sentence has a trivial monotone functional interpretation and we will use this later to formulate admissible axioms stating that certain classes of functions are indeed measurable.

7. Treating integration over probability contents

We now want to extend the previous system $\mathcal{F}^{\omega}[\mathbb{P}]$ for probability contents on algebras so that we can treat a certain class of integrable functions $f : \Omega \to \mathbb{R}$, in particular so that we obtain a suitable base for random variables and their moments as used in various applications, especially those illustrated in Section 9.

For the usual approach to the integral over contents, which mimics that of the Lebesgue integral, we mainly follow the exposition given in [6] (where the corresponding notion is introduced under the name of the "D-integral") which we detail here to some degree to provide the necessary mathematical basics for the axiomatizations chosen later. Concretely, let Ω be a set and S an algebra on it and let μ be a finite content on S (i.e., $\mu(\Omega) < +\infty$). One then first arrives at a notion of simple function that is completely analogous to how it is usually defined in the context of measure theory, that is, a *simple function* is a function $f : \Omega \to \mathbb{R}$ of the form

$$f(x) = \sum_{i=0}^{n} b_i \chi_{A_i}$$

for given sets $A_i \in S$ and values $b_i \in \mathbb{R}$.[13] For such a function f, the integral over μ is simply defined as

$$\int f \, \mathrm{d}\mu = \sum_{i=0}^{n} b_i \mu(A_i),$$

also in similarity to usual Lebesgue integrals over measures. A general function $f : \Omega \to \mathbb{R}$ is now declared *integrable* if there is a sequence of simple functions f_n such that (1) the f_n converge to f in a suitable sense[14] and (2) the sequence satisfies

[13] In some cases, as, for example, also in [6], one requires the sets A_i to be mutually disjoint and to cover the whole space Ω but we do not include these requirements here for simplicity.

[14] The corresponding notion of convergence is dubbed *hazy convergence* in [6] and relies on the use of a corresponding outer content (which is similar to an outer measure) constructed from μ. Here, we will however not rely on any precise details regarding this notion.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

$$\lim_{n,m\to\infty} \int |f_n - f_m| \, \mathrm{d}\mu = 0.$$

Such a sequence is called a *determining sequence* for f and using such a determining sequence, one then defines the integral of f via

$$\int f \, \mathrm{d}\mu = \lim_{n\to\infty} \int f_n \, \mathrm{d}\mu.$$

Crucially, this limit is well-defined as the following general result shows:

Lemma 7.1 (Lemma 4.4.12 in [6]). *Let f be an integrable function and let (f_n) be a determining sequence for f. Then $f_n - f$ is integrable and*

$$\lim_{n\to\infty} \int |f_n - f| \, \mathrm{d}\mu = 0.$$

This notion of a (D-)integral defined for contents then shares many of the familiar properties of Lebesgue integrals defined for measures. One particularly useful property is that every integrable function f is measurable in the following extended sense, called T_2-measurable in [6]: for any $\varepsilon > 0$, there is a partition $A_0, \ldots, A_n \in S$ of Ω such that $\mu(A_0) < \varepsilon$ and $|f(x) - f(y)| < \varepsilon$ for any $x, y \in A_i$ and any i. In particular, for any function f that is measurable in this sense, one gets that f is integrable if, and only if, $|f|$ is integrable (see, e.g., Corollary 4.4.19 in [6]).

Now, a major part of the theory of integrals over contents (like, e.g., a nice correspondence between the so-called D- and S-integrals, the latter being similar in spirit to a Riemann-Stieltjes integral, and the fact that the notions of T_2-measurability and integrability coincide as shown by Theorem 4.5.7 in [6], among many others) depends on the assumption that the integrated functions are bounded. In that way, we will similarly require that all integrated functions are bounded.

In fact, using the proof-theoretic perspective of the approach taken here, we find that this assumption of the boundedness of functions is also suggested as a necessity in our formal approach by the notion of majorizability employed later. Concretely, as discussed before, to develop a proof-theoretically tame theory of algebras and contents, we have regarded Ω and S as uniformly bounded in the sense that we later regard all elements of these spaces as uniformly majorized by the content of the full space $\mu(\Omega)$, that is, by 1 in the context of a probability content, which we denote in writing by $1 \gtrsim_\Omega x$ and $1 \gtrsim_S A$ for $x \in \Omega$ and $A \in S$. While this will be discussed comprehensively and in full formal detail later, we here already look at what this definition entails for majorizable functions f of type $1(\Omega)$: a function f^* of type $0(0)(0)$ is a majorant for f, written $f^* \gtrsim_{1(\Omega)} f$, if

$$f^*(m)(k) \geq f^*(n)(j) \geq f(x)(i) \text{ whenever } m \geq n \gtrsim_\Omega x \text{ and } k \geq j \geq i.$$

Therefore, as $1 \gtrsim_\Omega x$ for any x, we in particular derive that

$$|f(x)| \leq [|f(x)|](0) + 1 \leq f(x)(0) + 1 \leq f^*(1)(0) + 1$$

for any x, using the monotonicity of the coding of rational numbers (see, e.g., the discussion on p.430 in [26]). Thus, the real number represented by $f(x)$ is uniformly bounded by $f^*(1)(0) + 1$ for any x and so any majorizable function of type $1(\Omega)$ represents a bounded function $\Omega \to \mathbb{R}$. In other words, boundedness of integrable functions is suggested as a necessary assumption by the chosen proof-theoretic methodology.

Lastly, we will actually require that the integrated functions are not only T_2-measurable in the sense discussed above but that they even are weakly Borel-measurable in the sense discussed before (i.e., that the preimages of closed intervals are included in the underlying algebra). Clearly, any bounded and weakly Borel-measurable function is also T_2-measurable in the previous sense and thus integrable as discussed before. While this class is slightly restricted compared to that of all integrable functions, we find

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

that it has two central advantages: For one, it allows for a particularly smooth and proof-theoretically tame approach to the integral that is amenable to bound extraction theorems. For another, in virtually all previous ad hoc proof mining applications to measure and probability theory involving integrals, already the stronger (or, in the context of σ-algebras, equivalent) property of being Borel-measurable is assumed so that a treatment of this class still allows for our metatheorems established later to provide a proof-theoretic explanation of the respective extractions. Moreover, this is in particular true for the applications discussed later in Section 9. In that way, the approach to the integral via this assumption should be understood to be merely indicative regarding the possibilities of the present formal approach, where various other measurability and integrability notions could similarly be accommodated.

Now, the formal approach to the integral over a probability content is to abstractly encode a suitable subspace of bounded and weakly Borel-measurable functions closed under linear combinations, multiplications with characteristic functions, and absolute values[15] intensionally via the use of a characteristic function and then to introduce the integral for all functions from this space as well as the relevant closure properties using further constants and axioms. For this, we now initially introduce two further constants

- I of type $0(1(\Omega))$,
- $\|\cdot\|_\infty$ of type $1(1(\Omega))$,

into the language of $\mathcal{F}^\omega[\mathbb{P}, \mathrm{Int}, \mathrm{Inv}]$. The first of these is the previously mentioned characteristic function providing an intensional account of a space closed under linear combinations, multiplications with characteristic functions and absolute values as well as containing only bounded and weakly Borel-measurable functions and the latter is used to formally introduce the supremum norm on these functionals. As initial axioms, we now therefore stipulate the following:

$$\forall f^{1(\Omega)}(If \leq_0 1), \tag{I)$_1$}$$

$$\forall A^S((\lambda x.(x \in A)) \in I), \tag{I)$_2$}$$

$$\forall f^{1(\Omega)}, a^1, b^1 \exists A^S \forall x^\Omega \Big(f \in I \to \Big(x \in f^{-1}([a,b]) \leftrightarrow x \in A\Big)\Big), \tag{I)$_3$}$$

$$\forall f^{1(\Omega)}, x^\Omega (f \in I \to (|f(x)| \leq_{\mathbb{R}} \|f\|_\infty)), \tag{I)$_4$}$$

$$\forall f^{1(\Omega)}, k^0 \exists x^\Omega \Big(f \in I \to \Big(\|f\|_\infty - 2^{-k} \leq_{\mathbb{R}} |f(x)|\Big)\Big), \tag{I)$_5$}$$

$$\forall f^{1(\Omega)}, g^{1(\Omega)}, \alpha^1, \beta^1 (f, g \in I \to (\lambda x.(\alpha f(x) + \beta g(x))) \in I), \tag{I)$_6$}$$

$$\forall f^{1(\Omega)}(f \in I \to \lambda x.|f(x)| \in I), \tag{I)$_7$}$$

$$\forall f^{1(\Omega)}, A^S (f \in I \to \lambda x.(f(x)(x \in A)) \in I). \tag{I)$_8$}$$

The axioms $(I)_3$ and $(I)_5$ will again be admissible later because of the extended notion of majorizability on Ω and S. Note also that axioms $(I)_4$ and $(I)_5$ together specify $\|f\|_\infty$ as the least upper bound on $|f(x)|$.[16]

In the following, we will write χ_A for the function $\lambda x.(x \in A^c)$.[17] Also, in the following we just briefly write $\alpha f + \beta g$ for $\lambda x.(\alpha f(x) + \beta g(x))$ as well as $|f|$ for $\lambda x.(|f(x)|)$ and $f\chi_A$ for $\lambda x.(f(x)\chi_A(x))$.

As the operations max and min can be defined on functions using the absolute value via

$$\max\{f, g\} = (f + g + |f - g|)/2 \text{ and } \min\{f, g\} = -\max\{-f, -g\},$$

we immediately get that axiom $(I)_7$ implies the closure of I under these operations and thus we have effectively axiomatized that I in particular is a Riesz space of bounded functions with the respective

[15] In the context of a probability measure over a σ-algebra, such a space could for example be the space of all bounded and Borel-measurable functions.

[16] In fact, these two axioms can be seen as an instantiation of the general approach to tame suprema over bounded sets developed in [51].

[17] As customary in the context of integration, we want characteristic functions to take the value 1 if the element is included in the set. As we previously have used 0 for this, we used A^c in the above definition.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

operations and thus our approach is similar to how abstract integration spaces are approached in the context of Daniell integrals [7] where a similar collection of functions is presumed.

To deal with the integral, we add a further constant

- $\int \cdot \,\mathrm{d}\mathbb{P}$ of type $1(1(\Omega))$.

The first two axioms for the integral are now that it behaves as expected on characteristic functions and that the integral is a linear function:

$$\forall A^S \left(\int \chi_A \,\mathrm{d}\mathbb{P} =_{\mathbb{R}} \mathbb{P}(A) \right), \tag{$\int_1$}$$

$$\forall f^{1(\Omega)}, g^{1(\Omega)}, \alpha^1, \beta^1 \left(f, g \in I \to \int (\alpha f + \beta g) \,\mathrm{d}\mathbb{P} =_{\mathbb{R}} \alpha \int f \,\mathrm{d}\mathbb{P} + \beta \int g \,\mathrm{d}\mathbb{P} \right). \tag{$\int_2$}$$

Using these two axioms, we immediately get that the integral behaves as expected on simple functions.

The major other statement that we need to axiomatize is that any function in I is actually integrable in the sense that its integral is well-defined and arises as the limit of a sequence of integrals of simple functions. In the context of our standing assumption that our functions are bounded and weakly Borel-measurable, we can in particular derive the following general result that guarantees this property, and inspires the subsequent axiom:

Lemma 7.2 (essentially folklore). *Let Ω be a set, S an algebra on Ω and μ a probability content on S. Let f be a weakly Borel-measurable function and assume that $|f|$ is bounded by $b \in \mathbb{N}^*$, that is, $|f(x)| \leq b$ for all $x \in \Omega$. For a given k, define*

$$I_{k,i} = \left[-b + \frac{i}{2^k}, -b + \frac{i+1}{2^k} \right) \text{ for } i = 0, 1, \ldots, b2^{k+1} - 2 \text{ and } I_{k,b2^{k+1}-1} = \left[\frac{b2^k - 1}{2^k}, b \right].$$

Then:

$$\forall k \in \mathbb{N} \left(\int \left| f - \sum_{i=0}^{b2^{k+1}-1} \left(-b + \frac{i}{2^k} \right) \chi_{f^{-1}(I_{k,i})} \right| \mathrm{d}\mathbb{P} \leq 2^{-k} \right).$$

Proof. Let $k \in \mathbb{N}$. As all $I_{k,i}$ are disjoint and cover $[-b, b]$ and since $|f|$ is bounded by b, their preimages under f are disjoint and cover Ω. Thus

$$1 = \mathbb{P}(\Omega) = \mathbb{P}\left(\bigcup_{i=0}^{b2^{k+1}-1} f^{-1}(I_{k,i}) \right) = \sum_{i=0}^{b2^{k+1}-1} \mathbb{P}(f^{-1}(I_{k,i}))$$

and for any $k \in \mathbb{N}$:

$$f(x) = \sum_{i=0}^{b2^{k+1}-1} f(x) \chi_{f^{-1}(I_{k,i})}(x).$$

Further, for $x \in f^{-1}(I_{k,i})$, it clearly holds that

$$\left| f(x) - \left(-b + \frac{i}{2^k} \right) \right| \leq \frac{1}{2^k}.$$

As k was arbitrary, the function f is integrable by Theorem 4.5.7 in [6] (use, e.g., the equivalence between (viii) and (v)). Thus, using the monotonicity and linearity of the integral on contents (see, e.g., Theorem 4.4.13 in [6]), we have

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

$$\begin{aligned}\int \left| f - \sum_{i=0}^{b2^{k+1}-1} \left(-b + \frac{i}{2^k}\right) \chi_{f^{-1}(I_{k,i})} \right| \mathrm{d}\mathbb{P} &= \int \left| \sum_{i=0}^{b2^{k+1}-1} \left(f + b - \frac{i}{2^k}\right) \chi_{f^{-1}(I_{k,i})} \right| \mathrm{d}\mathbb{P} \\ &\leq \sum_{i=0}^{b2^{k+1}-1} \int \left| f - \left(-b + \frac{i}{2^k}\right) \right| \chi_{f^{-1}(I_{k,i})} \, \mathrm{d}\mathbb{P} \\ &\leq \sum_{i=0}^{b2^{k+1}-1} \int \frac{1}{2^k} \chi_{f^{-1}(I_{k,i})} \, \mathrm{d}\mathbb{P} \\ &\leq \sum_{i=0}^{b2^{k+1}-1} \frac{1}{2^k} \mathbb{P}(f^{-1}(I_{k,i})) \\ &\leq \frac{1}{2^k} \sum_{i=0}^{b2^{k+1}-1} \mathbb{P}(f^{-1}(I_{k,i})) = \frac{1}{2^k}. \end{aligned}$$ □

To axiomatize the integrability of a function $f \in I$, it thus suffices to state the conclusion of the above lemma and although we could formalize this directly by employing the general inverse mapping and intensional intervals, we can actually avoid this machinery here at the mild expense of quantifying over the sequence of sets used in the simple functions instead of explicitly specifying them. Concretely, we consider the following third axiom[18]

$$\forall f^{1(\Omega)} \forall k^0 \exists A^{S(0)} \left(f \in I \to \int \left| f - \sum_{i=0}^{2^{k+1} b_f - 1} \left(-b_f + \frac{i}{2^k}\right) \chi_{A(i)} \right| \mathrm{d}\mathbb{P} \leq_{\mathbb{R}} 2^{-k} \right), \qquad (\int)_3$$

where we wrote $b_f := \|f\|_\infty(0) + 1$ and have used that, since $\|f\|_\infty \geq_{\mathbb{R}} |f(x)|$ for all x, it holds that the natural number b_f similarly bounds $|f|$. Again, by the later considerations on majorizability whereby also A of type $S(0)$ can be regarded as uniformly bounded, this axiom will later be admissible in the context of our approach to proof mining metatheorems via the monotone functional interpretation.

Lastly, to also devise a practical system, it will be convenient to also axiomatically include (instead of discussing how it might be provable in the system) that the integral of a positive function $f \in I$ is positive. Naively, this statement can be written as

$$\forall f^{1(\Omega)} \left(f \in I \wedge \forall x^{\Omega} (f(x) \geq_{\mathbb{R}} 0) \to \int f \, \mathrm{d}\mathbb{P} \geq_{\mathbb{R}} 0 \right)$$

which is not a priori admissible in the context of our approach to bound extraction theorems due to the hidden negative universal type 0 quantifier in $\geq_{\mathbb{R}}$. However, if we rewrite the above statement in the prenexed form

$$\forall f^{1(\Omega)} \forall k^0 \exists x^{\Omega} \exists j^0 \left(f \in I \wedge \int f \, \mathrm{d}\mathbb{P} <_{\mathbb{R}} -2^{-k} \to \left(f(x) \leq_{\mathbb{R}} -2^{-j} \right) \right),$$

we can witness the quantifier over j simply by $j = k$ which can be immediately seen using only very basic properties of the integral (see, e.g., Theorem 4.4.13 in [6]) so that we arrive at the axiom

$$\forall f^{1(\Omega)} \forall k^0 \exists x^{\Omega} \left(f \in I \wedge \int f \, \mathrm{d}\mathbb{P} <_{\mathbb{R}} -2^{-k} \to \left(f(x) \leq_{\mathbb{R}} -2^{-k} \right) \right) \qquad (\int)_4$$

[18]Note that in the context of this axiom, we can introduce the sum expression by the recursor constants of the underlying system $\mathcal{A}^\omega$.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

which, in the context of our perspective that we regard quantification over Ω as bounded, will later be admissible in the context of the monotone functional interpretation.

We want to note that the axioms $(\int)_2 - (\int)_4$ roughly correspond to the four central properties of an abstract "I-integral" in the context of Daniell's approach to integration [7] and so, in some sense, our approach to the integral here can be regarded as a sort of effectivized implementation of the Daniell integral.

With all of these constants and axioms, we then arrive at the following system for integrals:

Definition 7.3. We write $\mathcal{F}^\omega[\mathbb{P}, \text{Integral}]$ for the system resulting from $\mathcal{F}^\omega[\mathbb{P}, \text{Int}, \text{Inv}]$ by adding the above constants $I, \|\cdot\|_\infty, \int \cdot \, \mathrm{d}\mathbb{P}$ together with the axioms $(I)_1 - (I)_8$ as well as $(\int)_1 - (\int)_4$. Further, we write $\mathcal{F}^\omega[\bigcup, \mathbb{P}, \text{Integral}]$ for the system $\mathcal{F}^\omega[\mathbb{P}, \text{Integral}]$ extended with the constant $\bigcup$ and the axiom $(\bigcup)_1$ as well as the rule $(\bigcup)_2$.

We end this section with some immediate properties of the integral over contents that are provable in our system. A more intricate use of the integrals will then be made in Section 9 where they (together with the boundedness and weak Borel-measurability) feature crucially in the formal explanation of a previous proof-mining application (and in particular highlight the usability of the above axiomatic approach to the integral with regard to the proof mining practice). As is common in proof mining, however, this approach to the integral enjoys a large degree of flexibility in the sense that it can, of course, be immediately augmented by further constants and axioms specifying other properties of the integral that might be crucial in a certain application, as also already mentioned before.

Lemma 7.4. *The system* $\mathcal{F}^\omega[\mathbb{P}, \text{Integral}]$ *proves:*

1. $\int \cdot \, \mathrm{d}\mathbb{P}$ *is monotone w.r.t. pointwise inequality, that is,*

$$\forall f^{1(\Omega)}, g^{1(\Omega)} \Big(f, g \in I \wedge \forall x^\Omega (f(x) \leq_{\mathbb{R}} g(x)) \to \int f \, \mathrm{d}\mathbb{P} \leq_{\mathbb{R}} \int g \, \mathrm{d}\mathbb{P} \Big).$$

2. $\int \cdot \, \mathrm{d}\mathbb{P}$ *is extensional w.r.t. pointwise equality, that is,*

$$\forall f^{1(\Omega)}, g^{1(\Omega)} \Big(f, g \in I \wedge \forall x^\Omega (f(x) =_{\mathbb{R}} g(x)) \to \Big| \int f \, \mathrm{d}\mathbb{P} - \int g \, \mathrm{d}\mathbb{P} \Big| =_{\mathbb{R}} 0 \Big).$$

3. $\int \cdot \, \mathrm{d}\mathbb{P}$ *is monotone w.r.t. inequality almost everywhere, that is,*

$$\forall f^{1(\Omega)}, g^{1(\Omega)} \Big(f, g \in I \wedge \exists A^S \Big(\mathbb{P}(A) = 0 \wedge \forall x^\Omega (x \in A^c \to f(x) \leq_{\mathbb{R}} g(x)) \Big) \to \int f \, \mathrm{d}\mathbb{P} \leq_{\mathbb{R}} \int g \, \mathrm{d}\mathbb{P} \Big).$$

4. $\int \cdot \, \mathrm{d}\mathbb{P}$ *is extensional w.r.t. equality almost everywhere, that is,*

$$\forall f^{1(\Omega)}, g^{1(\Omega)} \Big(f, g \in I \wedge \exists A^S \Big(\mathbb{P}(A) = 0 \wedge \forall x^\Omega (x \in A^c \to f(x) =_{\mathbb{R}} g(x)) \Big) \to \Big| \int f \, \mathrm{d}\mathbb{P} - \int g \, \mathrm{d}\mathbb{P} \Big| =_{\mathbb{R}} 0 \Big).$$

5. $\int \cdot \, \mathrm{d}\mathbb{P}$ *behaves well with absolute values, that is,*

$$\forall f^{1(\Omega)} \Big(f \in I \to \Big| \int f \, \mathrm{d}\mathbb{P} \Big| \leq_{\mathbb{R}} \int |f| \, \mathrm{d}\mathbb{P} \Big).$$

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

Proof.

1. If $\forall x(f(x) \leq g(x))$, note that we have $(g - f)(x) \geq 0$ for all x. By axiom $(\int)_4$, we have thus that $\int (g - f)\,\mathrm{d}\mathbb{P} \geq 0$ and therefore, we get $\int f\,\mathrm{d}\mathbb{P} \leq \int g\,\mathrm{d}\mathbb{P}$ by axiom $(\int)_2$.
2. This immediately follows from item (1).
3. By axiom $(I)_8$, we have $f\chi_{A^c}, g\chi_{A^c} \in I$. As in particular $f\chi_{A^c}(x) \leq g\chi_{A^c}(x)$ holds for any x, we get
$$\int f\chi_{A^c}\,\mathrm{d}\mathbb{P} \leq \int g\chi_{A^c}\,\mathrm{d}\mathbb{P}$$
by item (1). Similarly, as the axioms for the supremum norm imply that $(f - g)\chi_A(x) \leq \|f - g\|_\infty \chi_A(x)$ for all x, item (1) together with axiom $(\int)_1$ implies
$$\int (f - g)\chi_A\,\mathrm{d}\mathbb{P} \leq \int \|f - g\|_\infty \chi_A\,\mathrm{d}\mathbb{P} = \|f - g\|_\infty \mathbb{P}(A) = 0$$
which yields
$$\int f\chi_A\,\mathrm{d}\mathbb{P} \leq \int g\chi_A\,\mathrm{d}\mathbb{P}.$$
As $f(x) = f\chi_A(x) + f\chi_{A^c}(x)$ holds for all x (and similarly for g), we thus in particular get the claim using axiom $(\int)_2$.
4. This immediately follows from item (3).
5. Note that we have provably that $-|f(x)| \leq f(x) \leq |f(x)|$ for any x so that by item (1) and axiom $(\int)_2$, we have
$$-\int |f|\,\mathrm{d}\mathbb{P} \leq \int f\,\mathrm{d}\mathbb{P} \leq \int |f|\,\mathrm{d}\mathbb{P},$$
that is, that $\left|\int f\,\mathrm{d}\mathbb{P}\right| \leq \int |f|\,\mathrm{d}\mathbb{P}$. □

Note that by item (2) of the above lemma together with the axioms on the supremum norm $\|\cdot\|_\infty$, we in particular have that $\int \cdot\,\mathrm{d}\mathbb{P}$ is extensional w.r.t. $\|\cdot\|_\infty$ in the sense that
$$\forall f^{1(\Omega)}, g^{1(\Omega)}\left(\|f - g\|_\infty =_{\mathbb{R}} 0 \to \left|\int f\,\mathrm{d}\mathbb{P} - \int g\,\mathrm{d}\mathbb{P}\right| =_{\mathbb{R}} 0\right).$$

8. A bound extraction theorem

We now establish our main results, the bound extraction theorems, for the systems introduced previously. For that, as hinted at in the introduction, we follow the approach of the first metatheorems using abstract types presented in [25] (see also [12, 26]).[19] As the outline of our approach is rather standard in that way, we will sometimes only sketch the arguments instead of giving full, detailed proofs, only spelling out those parts that are sensitive to the new ideas introduced in this paper. Throughout, to ease notation, we write $\mathcal{C}^\omega$ for the system $\mathcal{F}^\omega$ or one of its extensions as discussed previously.

As in the works mentioned above, the main tool for the metatheorems presented here is Gödel's Dialectica interpretation [13] which is combined with a negative translation by Kuroda [36]. We recall the definitions of those central proof interpretations here:

Definition 8.1 [13, 60]. The Dialectica interpretation $A^D = \exists \underline{x} \forall \underline{y} A_D(\underline{x}, \underline{y})$ of a formula A in the language of $\mathcal{C}^\omega$ (and its extensions) is defined via the following recursion on the structure of the formula:

[19] As mentioned in the introduction already, this approach has been adapted to provide a treatment amenable to proof mining methods of a wide array of different mathematical notions in the past. We again refer to the references in the introduction for these results.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

1. $A^D := A_D := A$ for A being a prime formula.

If $A^D = \exists \underline{x} \forall \underline{y} A_D(\underline{x}, \underline{y})$ and $B^D = \exists \underline{u} \forall \underline{v} B_D(\underline{u}, \underline{v})$, we set

2. $(A \wedge B)^D := \exists \underline{x}, \underline{u} \forall \underline{y}, \underline{v} (A \wedge B)_D$
   where $(A \wedge B)_D(\underline{x}, \underline{u}, \underline{y}, \underline{v}) := A_D(\underline{x}, \underline{y}) \wedge B_D(\underline{u}, \underline{v})$,
3. $(A \vee B)^D := \exists z^0, \underline{x}, \underline{u} \forall \underline{y}, \underline{v} (A \vee B)_D$
   where $(A \vee B)_D(z^0, \underline{x}, \underline{u}, \underline{y}, \underline{v}) := (z = 0 \to A_D(\underline{x}, \underline{y})) \wedge (z \neq 0 \to B_D(\underline{u}, \underline{v}))$,
4. $(A \to B)^D := \exists \underline{U}, \underline{Y} \forall \underline{x}, \underline{v} (A \to B)_D$
   where $(A \to B)_D(\underline{U}, \underline{Y}, \underline{x}, \underline{v}) := A_D(\underline{x}, \underline{Y}\underline{x}\underline{v}) \to B_D(\underline{U}\underline{x}, \underline{v})$,
5. $(\exists z^\tau A(z))^D := \exists z, \underline{x} \forall \underline{y} (\exists z^\tau A(z))_D$
   where $(\exists z^\tau A(z))_D(z, \underline{x}, \underline{y}) := A_D(\underline{x}, \underline{y}, z)$,
6. $(\forall z^\tau A(z))^D := \exists \underline{X} \forall z, \underline{y} (\forall z^\tau A(z))_D$
   where $(\forall z^\tau A(z))_D(\underline{X}, z, \underline{y}) := A_D(\underline{X}z, \underline{y}, z)$.

Definition 8.2 [36]. The negative translation of A is defined by $A' := \neg\neg A^*$ where A^* is defined by the following recursion on the structure of A:

1. $A^* := A$ for prime A;
2. $(A \circ B)^* := A^* \circ B^*$ for $\circ \in \{\wedge, \vee, \to\}$;
3. $(\exists x^\tau A)^* := \exists x^\tau A^*$;
4. $(\forall x^\tau A)^* := \forall x^\tau \neg\neg A^*$.

For the combination of these two interpretations, the following soundness result is one of the two central technical tools in the context of the proof of the proof mining metatheorems. In that context, we define $\mathcal{C}^{\omega -}$ as the system $\mathcal{C}^\omega$ *without* the schemas QF-AC and DC.

Lemma 8.3 (essentially [25]). *Let $\mathcal{P}$ be a set of universal sentences and let $A(\underline{a})$ be an arbitrary formula (with only the variables $\underline{a}$ free) in the language of $\mathcal{F}^\omega$. Then the rule*

$$\begin{cases} \mathcal{F}^\omega + \mathcal{P} \vdash A(\underline{a}) \Rightarrow \\ \mathcal{F}^{\omega -} + (\mathrm{BR}) + \mathcal{P} \vdash \forall \underline{a}, \underline{y} (A')_D(\underline{t}\underline{a}, \underline{y}, \underline{a}) \end{cases}$$

holds where $\underline{t}$ is a tuple of closed terms of $\mathcal{F}^{\omega -} + (\mathrm{BR})$ which can be extracted from the respective proof and (BR) *is the schema of simultaneous bar-recursion of Spector [58], here extended to all types from $T^{\Omega, S}$ (similar to, e.g., [26]).*

This result extends to any suitable extension of the language of $\mathcal{F}^\omega$ (e.g., by any kind of new types and constants) together with any number of additional universal axioms in that language.

We omit the proof as it is almost exactly the same as the proof given for the analogous soundness result in [25] (although this result from [25] is of course not formulated for the system $\mathcal{F}^\omega$).

Besides the soundness of the Dialectica interpretation (together with the negative translation), the other one of the two central tools utilized in the metatheorems is that of majorizability. Originally introduced by Howard [19], the notion of majorizable functionals was later extended by Bezem [5] to that of strongly majorizable functionals to provide a model for finite type arithmetic extended by the schema of bar recursion discussed before. In that way, this model of strongly majorizable functionals provides the crucial basis for proof mining metatheorems of systems allowing for dependent choice. In the context of the abstract types, we further need to consider an extension of this notion of strongly majorizable functionals to these new types. The first such extensions to abstract types have been devised in [12, 25]. However, in this setting (and essentially in all other settings for metatheorems proved afterwards), this extension was motivated and based on the metric structure assumed for the respective classes of spaces that were treated. We thus find ourselves here at a "fork in the road," where we have to extend the notion of majorizability sensibly to our types Ω and S, both representing spaces which do not carry any metric structure. The key insight, already mentioned and motivated throughout the

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

previous sections many times (e.g., in the context of the admissibility of the axioms containing existential quantifiers over variables of types Ω, S or $S(0)$, etc.) is to

1. majorize objects A^S by natural numbers bounding the measure of A,
2. majorize objects x^Ω by natural numbers bounding the measure of the full set Ω in S.

In the case of a probability measure, any object of type Ω or S is therefore uniformly majorized by 1 but with the phrasing of (1) and (2), we wanted to highlight the general idea of this approach as it might be feasible also for more general finite contents.

In any way, similar to [12, 25], the majorants for objects with types from $T^{\Omega,S}$ are objects with types from T according to the following extended projection:

Definition 8.4 (essentially [12]). Define $\widehat{\tau} \in T$, given $\tau \in T^{\Omega,S}$, by recursion on the structure via

$$\widehat{0} := 0,\ \widehat{\Omega} := 0,\ \widehat{S} := 0,\ \widehat{\tau(\xi)} := \widehat{\tau}(\widehat{\xi}).$$

The majorizability relation $\gtrsim$ is then defined in tandem with the structure of all strongly majorizable functionals.

Definition 8.5 (essentially [12, 25]). Let Ω be a nonempty 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 $\gtrsim_\rho$ are defined by

$$\begin{cases} \mathcal{M}_0 := \mathbb{N}, n \gtrsim_0 m := n \geq m \wedge n, m \in \mathbb{N}, \\ \mathcal{M}_\Omega := \Omega, n \gtrsim_\Omega x := n \geq \mathbb{P}(\Omega) \wedge n \in \mathcal{M}_0, x \in \mathcal{M}_\Omega, \\ \mathcal{M}_S := S, n \gtrsim_S A := n \geq \mathbb{P}(A) \wedge n \in \mathcal{M}_0, A \in \mathcal{M}_S, \\ f \gtrsim_{\tau(\xi)} x := f \in \mathcal{M}_{\widehat{\tau}}^{\mathcal{M}_{\widehat{\xi}}} \wedge x \in \mathcal{M}_\tau^{\mathcal{M}_\xi} \\ \qquad\qquad \wedge \forall g \in \mathcal{M}_{\widehat{\xi}}, y \in \mathcal{M}_\xi (g \gtrsim_\xi y \to fg \gtrsim_\tau xy) \\ \qquad\qquad \wedge \forall g, y \in \mathcal{M}_{\widehat{\xi}} (g \gtrsim_{\widehat{\xi}} y \to fg \gtrsim_{\widehat{\tau}} fy), \\ \mathcal{M}_{\tau(\xi)} := \left\{ x \in \mathcal{M}_\tau^{\mathcal{M}_\xi} \mid \exists f \in \mathcal{M}_{\widehat{\tau}}^{\mathcal{M}_{\widehat{\xi}}} : f \gtrsim_{\tau(\xi)} x \right\}. \end{cases}$$

So, as already discussed previously, though only informally, as $\mathbb{P}$ is a probability content, we have $1 \gtrsim_S A$ for any $A \in S$ (as $\mathbb{P}(A) \leq \mathbb{P}(\Omega) = 1$) as well as $1 \gtrsim_\Omega x$ for any $x \in \Omega$ (but in the way the model is defined above, the definition immediately makes sense in the context of general finite contents).

Before we move on further, we now just quickly note that majorizability behaves nicely w.r.t. functions with multiple arguments as represented by their curried variants.

Lemma 8.6 ([12, 25], see also Kohlenbach [26, Lemma 17.80]). *Let $\xi = \tau(\xi_k) \dots (\xi_1)$. For $x^* : \mathcal{M}_{\widehat{\xi_1}} \to (\mathcal{M}_{\widehat{\xi_2}} \to \dots \to \mathcal{M}_{\widehat{\tau}}) \dots)$ and $x : \mathcal{M}_{\xi_1} \to (\mathcal{M}_{\xi_2} \to \dots \to \mathcal{M}_\tau) \dots)$, we have $x^* \gtrsim_\xi x$ iff*

1. $\forall y_1^*, y_1, \dots, y_k^*, y_k \left(\wedge_{i=1}^k (y_i^* \gtrsim_{\xi_i} y_i) \to x^* y_1^* \dots y_k^* \gtrsim_\tau x y_1 \dots y_k \right)$ *and*
2. $\forall y_1^*, y_1, \dots, y_k^*, y_k \left(\wedge_{i=1}^k (y_i^* \gtrsim_{\widehat{\xi_i}} y_i) \to x^* y_1^* \dots y_k^* \gtrsim_{\widehat{\tau}} x^* y_1 \dots y_k \right)$.

The other main structure featuring in the metatheorems is the structure of all set-theoretic functionals $\mathcal{S}^{\omega,\Omega,S}$, defined via $\mathcal{S}_0 := \mathbb{N}$, $\mathcal{S}_\Omega := \Omega$, $\mathcal{S}_S := S$ and

$$\mathcal{S}_{\tau(\xi)} := \mathcal{S}_\tau^{\mathcal{S}_\xi}.$$

Both structures $\mathcal{S}^{\omega,\Omega,S}$ and $\mathcal{M}^{\omega,\Omega,S}$ later turn into models of our systems if equipped with corresponding interpretations for the respective additional constants, with $\mathcal{S}^{\omega,\Omega,S}$ serving as the structure for the intended standard models.

The proof of the bound extraction theorems now follows the following general high-level outline of most other such metatheorems: using functional interpretation and negative translation, one extracts realizers from (essentially) $\forall\exists$-theorems. These realizers have types from $T^{\Omega,S}$. We then use

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

majorizability to construct bounds for these realizers, depending only on majorants of the parameters, which are validated in a model based on $\mathcal{M}^{\omega,\Omega,S}$. In a final step, we can then recover to the truth in a model based on the usual full set-theoretic structure $\mathcal{S}^{\omega,\Omega,S}$ if the types occurring in the axioms and the theorem are "low enough," which we will call admissible. Concretely, following [12, 25], we introduce the following specific classes of types: We call a type ξ *of degree* n if $\xi \in T$ and it has degree $\leq n$ in the usual sense (see, e.g., [26]). Further we call ξ *small* if it is of the form $\xi = \xi_0(0)\dots(0)$ for $\xi_0 \in \{0, \Omega, S\}$ (including $0, \Omega, S$) and call it *admissible* if it is of the form $\xi = \xi_0(\tau_k)\dots(\tau_1)$ where each τ_i is small and $\xi_0 \in \{0, \Omega, S\}$ (also including $0, \Omega, S$).

Further, also in analogy to [12, 25], we define certain subclasses of formulas satisfying certain type restrictions: A formula A is called a $\forall$-formula if $A = \forall \underline{a}^{\underline{\xi}} A_{qf}(\underline{a})$ with A_{qf} quantifier-free and all types ξ_i in $\underline{\xi} = (\xi_1, \dots, \xi_k)$ are admissible. A formula A is called an $\exists$-formula if $A = \exists \underline{a}^{\underline{\xi}} A_{qf}(\underline{a})$ with similar A_{qf} and $\underline{\xi}$.

The class Δ already mentioned previously, which was originally introduced in [20, 21] (and then lifted to abstract types in [16]) to signify a collection of commonly occurring formulas with trivial monotone functional interpretations, is now similarly introduced in the context of the systems studied in this paper: a formula of type Δ is any formula of the form

$$\forall \underline{a}^{\underline{\delta}} \exists \underline{b} \leq_{\underline{\sigma}} \underline{r}\underline{a} \forall \underline{c}^{\underline{\gamma}} A_{qf}(\underline{a}, \underline{b}, \underline{c})$$

where A_{qf} is quantifier-free, the types in $\underline{\delta}$, $\underline{\sigma}$, and $\underline{\gamma}$ are admissible, $\underline{r}$ is a tuple of closed terms of appropriate type, $\leq$ is defined by recursion on the type via

1. $x \leq_0 y := x \leq_0 y$,
2. $x \leq_\Omega y := \mathbb{P}(\Omega) \leq_{\mathbb{R}} \mathbb{P}(\Omega)$,
3. $A \leq_S B := \mathbb{P}(A) \leq_{\mathbb{R}} \mathbb{P}(B)$,
4. $x \leq_{\tau(\xi)} y := \forall z^{\xi}(xz \leq_\tau yz)$,

and $\underline{x} \leq_{\underline{\sigma}} \underline{y}$ is an abbreviation for $x_1 \leq_{\sigma_1} y_1 \wedge \dots \wedge x_k \leq_{\sigma_k} y_k$ where $\underline{x}$, $\underline{y}$, and $\underline{\sigma}$ are k-tuples of terms and types, respectively, such that x_i and y_i are of type σ_i.

Given a set $\beth$ of formulas of type Δ, we write $\overline{\beth}$ for the set of all Skolem normal forms

$$\exists \underline{B} \leq_{\underline{\sigma}(\underline{\delta})} \underline{r} \forall \underline{a}^{\underline{\delta}} \forall \underline{c}^{\underline{\gamma}} A_{qf}(\underline{a}, \underline{B}\underline{a}, \underline{c})$$

for any $\forall \underline{a}^{\underline{\delta}} \exists \underline{b} \leq_{\underline{\sigma}} \underline{r}\underline{a} \forall \underline{c}^{\underline{\gamma}} A_{qf}(\underline{a}, \underline{b}, \underline{c})$ in $\beth$.

Remark 8.7. We want to note briefly that all axioms that were previously discussed as admissible based on our extended notion of majorizability can actually be seen as statements of type Δ in the context of the above definition. At first, the axiom $(\mathbb{P})_4$ can be equivalently rewritten as

$$\forall A^S, B^S \exists x^\Omega \leq_\Omega c_\Omega(\mathbb{P}(A) >_{\mathbb{R}} \mathbb{P}(B) \to (x \in A \wedge x \notin B))$$

and is thus immediately of type Δ.[20] Second, also the axiom $(I)_3$ can be rewritten with the additional boundedness information via

$$\forall f^{1(\Omega)}, a^1, b^1 \exists A^S \leq_S \Omega \forall x^\Omega \Big(f \in I \to \Big(x \in f^{-1}([a,b]) \leftrightarrow x \in A \Big) \Big)$$

and thus is of type Δ. Similarly, also the axiom $(I)_5$ can be rewritten as an axiom of type Δ as

$$\forall f^{1(\Omega)}, k^0 \exists x^\Omega \leq_\Omega c_\Omega \Big(f \in I \to \Big(\|f\|_\infty - 2^{-k} \leq_{\mathbb{R}} |f(x)| \Big) \Big).$$

[20] Note the importance of the constant c_Ω witnessing the nonemptiness of Ω for writing $(\mathbb{P})_4$ in that way.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

Lastly, also the integrability axioms $(\int)_4$ and $(\int)_3$ can be rewritten as

$$\forall f^{1(\Omega)}, k^0 \exists x^{\Omega} \leq_{\Omega} c_{\Omega}\left(f \in I \wedge \int f \, \mathrm{d}\mathbb{P} <_{\mathbb{R}} -2^{-k} \to \left(f(x) \leq_{\mathbb{R}} -2^{-k}\right)\right)$$

and

$$\forall f^{1(\Omega)}, k^0 \exists A^{S(0)} \leq_{S(0)} \lambda n^0.\Omega\left(f \in I \to \int \left| f - \sum_{i=0}^{2^{k+1}b_f - 1} \left(-b_f + \frac{i}{2^k}\right) \chi_{A(i)} \right| \mathrm{d}\mathbb{P} \leq_{\mathbb{R}} 2^{-k}\right),$$

respectively, with $b_f := \|f\|_{\infty}(0) + 1$ as before, which turns them into axioms of type Δ.

Crucially, axioms of type Δ are trivialized under the monotone functional interpretation[21] and we treat any axiom of type Δ in $\mathcal{C}^{\omega}$ (or any suitable extension) "in this spirit." We here only write "in this spirit" as we actually do not use a monotone variant of the Dialectica interpretation but treat the functional interpretation part and the majorization part of the combined interpretation separately. In that way, we follow the approach given in [16] (see also the recent [51]) and treat axioms of type Δ by employing a construction that converts a theory with axioms of such a type into a theory using only additional purely universal axioms formulated using the Skolem functions of these axioms. This new theory is then used in combination with the functional interpretation to extract the respective terms and then the proof proceeds as outlined before.

Concretely, we now proceed as follows: Let $\beth$ be a set of axioms of type Δ and write $\widehat{\mathcal{C}}^{\omega}$ for $\mathcal{C}^{\omega}$ without any of its axioms of type Δ. Then, we form a new theory $\overline{\mathcal{C}}^{\omega}_{\beth}$ from $\widehat{\mathcal{C}}^{\omega}$ by adding the Skolem functionals $\underline{B}$ of any axiom of type Δ of $\mathcal{C}^{\omega} + \beth$, say of the form

$$\forall \underline{a}^{\underline{\delta}} \exists \underline{b} \leq_{\underline{\sigma}} \underline{r}\underline{a} \forall \underline{c}^{\underline{\gamma}} A_{qf}(\underline{a}, \underline{b}, \underline{c}),$$

as new constants to the language and simultaneously adding the corresponding "instantiated Skolem normal form," that is,

$$\underline{B} \leq_{\underline{\sigma}(\underline{\delta})} \underline{r} \wedge \forall \underline{a}^{\underline{\delta}} \forall \underline{c}^{\underline{\gamma}} A_{qf}(\underline{a}, \underline{B}\underline{a}, \underline{c}),$$

as a new axiom. Therefore, the system $\overline{\mathcal{C}}^{\omega}_{\beth}$ only extends $\mathcal{F}^{\omega}$ by new types, constants, and universal axioms. Therefore, as mentioned before, Lemma 8.3 also applies to this system where the conclusion is then proved in $\overline{\mathcal{C}}^{\omega^-}_{\beth}$ + (BR), that is, $\overline{\mathcal{C}}^{\omega}_{\beth}$ with the principles QF-AC and DC removed and where the scheme of simultaneous bar-recursion is added.

In the case where the extension $\mathcal{C}^{\omega}$ contains the rule $(\bigcup)_2$, we for simplicity assume that in the process of forming the extended theory, this rule is also removed in the sense that $\overline{\mathcal{C}}^{\omega}_{\beth}$ does not contain the rule and for any provable premise

$$\mathcal{C}^{\omega} + \beth \vdash F_{qf} \to \forall n^0 (A(n) \subseteq_S B),$$

we add the corresponding conclusion

$$F_{qf} \to \bigcup A \subseteq_S B$$

as an axiom of $\overline{\mathcal{C}}^{\omega}_{\beth}$.

We now move on to the central result of the majorization part of the chosen approach to the bound extraction theorems, stating that every closed term in the underlying language of the system in question is majorizable. As such, the result is similar to Lemma 9.11 in [12].

[21] While the interpretation was introduced under this name in [23], the idea of combining the Dialectica interpretation and majorization is already due to [21].

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

Lemma 8.8. *Let $\mathcal{C}^\omega$ be (one of the previously discussed extensions of) the system $\mathcal{F}^\omega[\mathbb{P}]$ and let $\beth$ be a set of additional axioms of type Δ. Let Ω be a nonempty set and let $S \subseteq 2^\Omega$ be an algebra (or, in the context of the constant $\bigcup$, a σ-algebra). Let $\mathbb{P}$ be a probability content on S. Then $\mathcal{M}^{\omega,\Omega,S}$ is a model of $\overline{\mathcal{C}}_\beth^{\omega-} + (\mathrm{BR})$, provided $\mathcal{S}^{\omega,\Omega,S} \models \beth$ (with $\mathcal{M}^{\omega,\Omega,S}$ and $\mathcal{S}^{\omega,\Omega,S}$ defined via suitable interpretations of the additional constants in $\mathcal{C}^\omega$). Moreover, for any closed term t of $\overline{\mathcal{C}}_\beth^{\omega-} + (\mathrm{BR})$, one can construct a closed term t^* of $\mathcal{A}^\omega + (\mathrm{BR})$ such that*

$$\mathcal{M}^{\omega,\Omega,S} \models (t^* \gtrsim t).$$

Proof. The structure of the proof is standard and similar to proofs of related results from the literature (see, e.g., [26]). As such, we only discuss the interpretations of the new constants added to $\mathcal{A}^\omega$ to form the respective theories as well as their majorizations. For the constants already contained in $\mathcal{A}^\omega$, we may choose suitable interpretations as in [26] and for majorizing a composition of terms, we may similarly proceed as outlined therein. For that, we now first focus on $\mathcal{F}^\omega[\mathbb{P}]$ and assume that there are no further axioms of type Δ beyond those already contained in $\mathcal{F}^\omega[\mathbb{P}]$. For any $\mathcal{C}^\omega$, we deal with any set $\beth$ of additional axioms of type Δ and the respectively induced constants later on by moving to the theory $\overline{\mathcal{C}}_\beth^\omega$.

Now, for the new constants added to $\mathcal{A}^\omega$ to form $\mathcal{F}^\omega[\mathbb{P}]$, we consider the following interpretations (writing $\mathcal{M}$ for $\mathcal{M}^{\omega,\Omega,S}$):

- $[\mathrm{eq}]_\mathcal{M}$:= the characteristic function of the equality relation in Ω;
- $[\in]_\mathcal{M}$:= the characteristic function of the element relation in S;
- $[\cup]_\mathcal{M}$:= union in S;
- $[(\cdot)^c]_\mathcal{M}$:= complement in S;
- $[\emptyset]_\mathcal{M}$:= the empty set in S;
- $[\mathbb{P}]_\mathcal{M} := \lambda A^S.(\mathbb{P}(A))_\circ$ where $\mathbb{P}$ is the content fixed in the context of this lemma.

This is only well-defined in $\mathcal{M}^{\omega,\Omega,S}$ if we can construct majorants of these objects. This we can do as follows:

- $\lambda x^0, y^0.1 \gtrsim \mathrm{eq}$;
- $\lambda x^0, y^0.1 \gtrsim \in$;
- $\lambda x^0, y^0.1 \gtrsim \cup$;
- $\lambda x^0.1 \gtrsim (\cdot)^c$;
- $0 \gtrsim \emptyset$;
- $\lambda x^0.(x)_\circ \gtrsim \mathbb{P}$.

Note that in the last item, the operation $(x)_\circ$ is definable in $\mathcal{A}^\omega$ via a closed term as x is of type 0.

For justifying that those terms really are majorants of the respective constants, we argue as follows: The first four items immediately follow from the fact that $\mathbb{P}(A) \leq \mathbb{P}(\Omega) = 1$ (i.e., that $\mathbb{P}$ is a probability content) and that $\mathbb{P}(\emptyset) =_\mathbb{R} 0$. The last item follows immediately from Lemma 2.1 as clearly, if $x \geq_\mathbb{R} \mathbb{P}(A)$, then $(x)_\circ \gtrsim (\mathbb{P}(A))_\circ$.

In the case where $\mathcal{C}^\omega$ contains the respective additional constant $\bigcup$, a corresponding interpretation is naturally defined by

- $[\bigcup]_\mathcal{M}$:= countably infinite union in S,

which is well-defined since we in this context assume that S is a σ-algebra. We can achieve majorization as before by exploiting that $\mathbb{P}$ is a finite content with

- $\lambda f^{0(0)}.1 \gtrsim \bigcup$.

Lastly, if the system $\mathcal{C}^\omega$ contains the respective constants and axioms for treating integrals, we choose corresponding interpretations of the additional constants as follows:

- $\left[[\cdot,\cdot]\right]_\mathcal{M} := \lambda a^1, b^1, x^1.\begin{cases} 0 & \text{if } r_x \in [r_a, r_b]; \\ 1 & \text{otherwise;} \end{cases}$

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

- $[(\cdot)^{-1}]_{\mathcal{M}} := \lambda f^{1(\Omega)}, A^{0(1)}, x^{\Omega}. \begin{cases} 0 & \text{if } x \in f^{-1}(\{r_a \mid a^1 : A(a) =_0 0\}); \\ 1 & \text{otherwise;} \end{cases}$
- $[I]_{\mathcal{M}} :=$ the characteristic function of a set of bounded and weakly Borel-measurable functions $f^{1(\Omega)}$ which is closed under linear combinations, multiplication with characteristic functions, and absolute values;
- $[\|\cdot\|_\infty]_{\mathcal{M}} := \lambda f^{1(\Omega)}. \begin{cases} (\sup_{x\in\Omega} |f(x)|)_\circ & \text{if } f \text{ is bounded;} \\ 0 & \text{otherwise;} \end{cases}$
- $[\int \cdot \, \mathrm{d}\mathbb{P}]_{\mathcal{M}} := \lambda f^{1(\Omega)}. \begin{cases} (\int f \, \mathrm{d}\mathbb{P})_\circ & \text{if } f \text{ is bounded and weakly Borel-measurable;} \\ 0 & \text{otherwise;} \end{cases}$

  where the latter $\int f \, \mathrm{d}\mathbb{P}$ represents the usual integral defined over the content.

Note that here, we now rely on the extended operator $(\cdot)_\circ$ operating on all real numbers as the integral of a general integrable function may be negative. As for majorization, we rely on the following constructions:

- $\lambda a^1, b^1, r^1.1 \gtrsim [\cdot,\cdot]$;
- $\lambda f^{1(0)}, A^{0(1)}, x^0.1 \gtrsim (\cdot)^{-1}$;
- $\lambda f^{1(0)}.1 \gtrsim I$;
- $\lambda f^{1(0)}.(f(1)(0)+1)_\circ \gtrsim \|\cdot\|_\infty$;
- $\lambda f^{1(0)}.(f(1)(0)+1)_\circ \gtrsim \int \cdot \, \mathrm{d}\mathbb{P}$.

The first three items are immediate as we deal with characteristic functions. For the fourth, note that if $f^{*1(0)} \gtrsim f^{1(\Omega)}$, then

$$\forall n^0, x^{\Omega}(n \gtrsim x \to f^*(n) \gtrsim f(x))$$

and as $1 \gtrsim x$ for any x^{Ω} as $1 = \mathbb{P}(\Omega)$, we have $f^*(1) \gtrsim f(x)$ for any x^{Ω}. Therefore, we have

$$f^*(1)(0)+1 \geq_{\mathbb{R}} f(x)(0)+1 \geq_{\mathbb{R}} [|f(x)|](0)+1 \geq_{\mathbb{R}} |f(x)|$$

for any x^{Ω}, by the monotonicity of the coding of rational numbers (again, see, e.g., the discussion on p.430 in [26]). This implies $f^*(1)(0)+1 \geq_{\mathbb{R}} \|f\|_\infty$ so that the result follows from Lemma 2.1.

Lastly, note that for any bounded and weakly Borel-measurable function f, we have that $|\int f \, \mathrm{d}\mathbb{P}| \leq_{\mathbb{R}} \int |f| \, \mathrm{d}\mathbb{P}$ so that

$$\left|\int f \, \mathrm{d}\mathbb{P}\right| \leq_{\mathbb{R}} \int |f| \, \mathrm{d}\mathbb{P} \leq_{\mathbb{R}} \||f|\|_\infty =_{\mathbb{R}} \|f\|_\infty \leq_{\mathbb{R}} f^*(1)(0)+1$$

as before. The majorizability result then follows again from Lemma 2.1.

That $\mathcal{M}^{\omega,\Omega,S}$ with these chosen interpretations is a model of $\mathcal{C}^{\omega-}$ + (BR) can be shown similarly to analogous results (see, e.g., [26]). The intended interpretations of the constants of $\mathcal{C}^{\omega}$ in $\mathcal{S}^{\omega,\Omega,S}$, turning $\mathcal{S}^{\omega,\Omega,S}$ into a model of these systems, are defined in analogy to the corresponding model $\mathcal{M}^{\omega,\Omega,S}$ defined above.

For treating the other additional axioms in $\mathcal{C}^{\omega} + \beth$ of type Δ beyond the axioms already contained in $\mathcal{C}^{\omega}$, we rely on the following argument (akin to [16], Lemma 5.11) showing that $\mathcal{S}^{\omega,\Omega,S} \models \beth$ implies $\mathcal{M}^{\omega,\Omega,S} \models \widetilde{\beth}$. For this, the proof given in [16] for Lemma 5.11 carries over which we sketch here: While $\mathcal{M}^{\omega,\Omega,S}$ in general is not a model of the axiom of choice [22], one can show (similar to [22]) that $\mathcal{M}^{\omega,\Omega,S} \models \mathsf{b\text{-}AC}_{\Omega,S}$ where

$$\mathsf{b\text{-}AC}_{\Omega,S} := \bigcup_{\delta,\rho\in T^{\Omega,S}} \mathsf{b\text{-}AC}^{\delta,\rho}$$

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

with

$$\mathsf{b\text{-}AC}^{\delta,\rho} := \forall Z^{\rho(\delta)}\Big(\forall x^{\delta}\exists y \leq_{\rho} ZxA(x,y,Z) \to \exists Y \leq_{\rho(\delta)} Z\forall x^{\delta} A(x,Yx,Z)\Big).$$

Now, for small types ρ, we have $M_\rho = S_\rho$ while for admissible types ρ, we have $M_\rho \subseteq S_\rho$ (for which it is important that admissible types take arguments of small types). For this, the proof given in [12] carries over. Further, we need that it is provable in $\mathcal{C}^{\omega-}$ that

$$\forall x', x, y\big(x' \gtrsim_{\rho} x \wedge x \geq_{\rho} y \to x' \gtrsim_{\rho} y\big) \tag{$\dagger$}$$

holds for all types ρ which can be shown similar to, for example, [26].

Suppose now that

$$\mathcal{S}^{\omega,\Omega,S} \models \forall \underline{a}^{\underline{\delta}}\exists \underline{b} \leq_{\underline{\sigma}} \underline{ra}\forall \underline{c}^{\underline{\gamma}} A_{qf}(\underline{a},\underline{b},\underline{c}).$$

Then also $\mathcal{M}^{\omega,\Omega,S}$ is a model of this sentence: First the types of the variables which are universally quantified are admissible, so over $\mathcal{M}^{\omega,\Omega,S}$ the domain of the universal quantifiers is reduced. For the witnesses for $\underline{b}$, which exist in $\mathcal{S}^{\omega,\Omega,S}$, note first that these could potentially live in $\mathcal{M}^{\omega,\Omega,S}$ as the types of the variables in $\underline{b}$ are admissible, that is, they take arguments of small types and map into small types. It thus only remains to be seen whether such a witness is majorizable for majorizable inputs. However, by the above argument, the terms in $\underline{r}$ are all majorizable and if $\underline{a}$ comes from $\mathcal{M}^{\omega,\Omega,S}$, then $\underline{ra}$ is majorizable. That we have $\underline{b} \leq_{\underline{\sigma}} \underline{ra}$ in $\mathcal{M}^{\omega,\Omega,S}$ now implies that $\underline{b}$ is majorizable by ($\dagger$) (and consequently the corresponding interpretations exist in $\mathcal{M}^{\omega,\Omega,S}$ too). Lastly, it is rather immediate to see that $\mathcal{M}^{\omega,\Omega,S} \models \exists$ implies $\mathcal{M}^{\omega,\Omega,S} \models \widetilde{\exists}$ using $\mathsf{b\text{-}AC}_{\Omega,S}$.

From $\mathcal{M}^{\omega,\Omega,S} \models \widetilde{\exists}$, we immediately get that the corresponding Skolem functions have interpretations in $\mathcal{M}^{\omega,\Omega,S}$, that the corresponding structures defined by some canonical interpretations of those additional constants are indeed models of those variants of the systems where the corresponding Skolem functionals of these axioms are added and where the axioms themselves are replaced by their instantiated Skolem normal forms (i.e., $\overline{\mathcal{C}}_{\exists}^{\omega-}$ and its extensions) and, lastly, that the above majorizability result extends to these systems.

Note that, technically, these arguments were already needed in the above considerations to see that $\mathcal{M}^{\omega,\Omega,S}$ really is a model of $\mathcal{C}^{\omega-}$ + (BR). However, we did not discuss this there explicitly as for those specific axioms of type Δ belonging to $\mathcal{C}^{\omega-}$ + (BR), the types of the variables occurring in them are not only small but actually all among $\{0, 1, \Omega, S, S(0)\}$ so that it was immediately clear that the models coincide at that level (essentially just by definition) and we thus omitted such a general discussion there. □

Combined with the Dialectica interpretation, the main result we then arrive at is the following bound extraction result for classical proofs:

Theorem 8.9. *Let $\mathcal{C}^{\omega}$ be (one of the previously discussed extensions of) the system $\mathcal{F}^{\omega}[\mathbb{P}]$ and let $\exists$ be a set of formulas of type Δ. Let τ be admissible, δ be of degree 1 and s be a closed term of $\mathcal{C}^{\omega}$ of type $\sigma(\delta)$ for admissible σ and let $B_\forall(x,y,z,u)/C_\exists(x,y,z,v)$ be $\forall$-/$\exists$-formulas of $\mathcal{C}^{\omega}$ with only $x,y,z,u/x,y,z,v$ free. If*

$$\mathcal{C}^{\omega} + \exists \vdash \forall x^{\delta}\forall y \leq_{\sigma} s(x)\forall z^{\tau}\Big(\forall u^{0} B_\forall(x,y,z,u) \to \exists v^{0} C_\exists(x,y,z,v)\Big),$$

then one can extract a partial functional $\Phi : \mathcal{S}_\delta \times \mathcal{S}_{\widehat{\tau}} \rightharpoonup \mathbb{N}$ which is total and (bar-recursively) computable on $\mathcal{M}_\delta \times \mathcal{M}_{\widehat{\tau}}$ and such that for all $x \in \mathcal{S}_\delta$, $z \in \mathcal{S}_\tau$, $z^* \in \mathcal{S}_{\widehat{\tau}}$, if $z^* \gtrsim z$, then

$$\mathcal{S}^{\omega,\Omega,S} \models \forall y \leq_{\sigma} s(x)(\forall u \leq_0 \Phi(x,z^*) B_\forall(x,y,z,u) \to \exists v \leq_0 \Phi(x,z^*) C_\exists(x,y,z,v))$$

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

holds whenever $\mathcal{S}^{\omega,\Omega,S} \models \Box$ *for* $\mathcal{S}^{\omega,\Omega,S}$ *defined via any nonempty set* Ω *and any algebra* $S \subseteq 2^\Omega$ *(or, in the context of the constant* $\bigcup$*, any* σ*-algebra) together with any probability content* $\mathbb{P}$ *on* S *(and with suitable interpretations of the additional constants). Further:*

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, y, z, u, v *and a finite conjunction instead of a single premise* $\forall u^0 B_\forall(x, y, z, u)$*.*
3. *If the claim is proved without* DC*, then* τ *may be arbitrary and* Φ *will be a total functional on* $\mathcal{S}_\delta \times \mathcal{S}_{\widehat{\tau}}$ *which is primitive recursive in the sense of Gödel [13] and Hilbert [18]. In that case, also plain majorization can be used instead of strong majorization (see, e.g., [26]).*

Proof. First, assume that $\mathcal{S}^{\omega,\Omega,S} \models \Box$ and (for simplicity) that

$$\mathcal{C}^\omega + \Box \vdash \forall z^\tau \Big(\forall u^0 B_\forall(z, u) \to \exists v^0 C_\exists(z, v) \Big).$$

Clearly, the same statement is then also provable in $\overline{\mathcal{C}}^\omega_\Box$. By assumption, $B_\forall(z, u) = \forall \underline{a} B_{qf}(z, u, \underline{a})$ and $C_\exists(z, v) = \exists \underline{b} C_{qf}(z, v, \underline{b})$ for quantifier-free B_{qf} and C_{qf}. Thus, by prenexiation, we get

$$\overline{\mathcal{C}}^\omega_\Box \vdash \forall z^\tau \exists u, v, \underline{a}, \underline{b} (B_{qf}(z, u, \underline{a}) \to C_{qf}(z, v, \underline{b})).$$

Using Lemma 8.3 (which is applicable as $\overline{\mathcal{C}}^\omega_\Box$ is an extension of $\mathcal{F}^\omega$ only by new constants and purely universal axioms) and disregarding the realizers for $\underline{a}, \underline{b}$, we get closed terms t_u, t_v of $\overline{\mathcal{C}}^{\omega-}_\Box + (\mathrm{BR})$ such that

$$\overline{\mathcal{C}}^{\omega-}_\Box + (\mathrm{BR}) \vdash \forall z^\tau (B_\forall(z, t_u(z)) \to C_\exists(z, t_v(z))).$$

By Lemma 8.8 there are closed terms t^*_u, t^*_v of $\mathcal{A}^\omega + (\mathrm{BR})$ such that

$$\mathcal{M}^{\omega,\Omega,S} \models t^*_u \gtrsim t_u \wedge t^*_v \gtrsim t_v \wedge \forall z^\tau (B_\forall(z, t_u(z)) \to C_\exists(z, t_v(z)))$$

for all nonempty sets Ω, any algebra (or σ-algebra) $S \subseteq 2^\Omega$ and any probability content $\mathbb{P}$ on S and where the constants are interpreted as in Lemma 8.8. Define

$$\Phi(z^*) := \max\{t^*_u(z^*), t^*_v(z^*)\}.$$

Then

$$\mathcal{M}^{\omega,\Omega,S} \models \forall u \leq_0 \Phi(z^*) B_\forall(z, u) \to \exists v \leq_0 \Phi(z^*) C_\exists(z, v)$$

holds for all $z \in \mathcal{M}_\tau$ and $z^* \in \mathcal{M}_{\widehat{\tau}}$ with $z^* \gtrsim z$. The conclusion that $\mathcal{S}^{\omega,\Omega,S}$ satisfies the same sentence can be achieved as in the proof of Theorem 17.52 in [26] which we sketch here: Note that in the conclusion, we restrict ourselves to those z which have majorants z^*. As the type of z is admissible, it takes arguments of small type for which $\mathcal{M}^{\omega,\Omega,S}$ and $\mathcal{S}^{\omega,\Omega,S}$ coincide (recall the proof of Lemma 8.8). Therefore, any such z, z^* from $\mathcal{S}^{\omega,\Omega,S}$ also live in $\mathcal{M}^{\omega,\Omega,S}$ so that $\Phi(z^*)$ is well-defined for z, z^* belonging to $\mathcal{S}^{\omega,\Omega,S}$ with $z^* \gtrsim z$. In $B_\forall$, all types are admissible so that truth in $\mathcal{S}^{\omega,\Omega,S}$ implies truth in $\mathcal{M}^{\omega,\Omega,S}$ and similarly for $C_\exists$ where thus truth in $\mathcal{M}^{\omega,\Omega,S}$ implies truth in $\mathcal{S}^{\omega,\Omega,S}$. Lastly, as in Lemma 17.84 in [26], we can show that as Φ is of type $0(\widehat{\tau})$, the interpretations of Φ in $\mathcal{S}^{\omega,\Omega,S}$ and $\mathcal{M}^{\omega,\Omega,S}$ coincide on majorizable elements. All in all, the arguments above imply that

$$\mathcal{S}^{\omega,\Omega,S} \models \forall u \leq_0 \Phi(z^*) B_\forall(z, u) \to \exists v \leq_0 \Phi(z^*) C_\exists(z, v)$$

holds for all $z \in S_\tau$ and $z^* \in S_{\widehat{\tau}}$ with $z^* \gtrsim z$.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

The additional $\forall x^{\delta}\forall y \leq_{\sigma} s(x)$ can be treated as, for example, discussed in [25] and we thus omit any details. Similarly, item (1) can be shown as in the proof of Theorem 17.52 from [26] (see page 428 therein). Further, (2) is immediate and (3) follows from the fact that without DC, bar recursion becomes superfluous, and the model $\mathcal{M}^{\omega,\Omega,S}$ can be avoided. □

9. Applications of the metatheorems

In this section, we are now concerned with the applications of the above metatheorems. Concretely, we want to indicate how the systems introduced prior can be used together with their metatheorems to recognize previous (ad hoc) applications in the spirit of the proof mining program as proper instances of proof-theoretic bound extraction theorems. For this, we here focus on the seminal work [2] by Avigad, Dean, and Rute. There, we find that the quantitative results obtained in [2] for Egorov's theorem, as well as the dominated convergence theorem, although in general being of bar-recursive complexity (which is due to the use of a certain principle of countable choice as we will later see formally), are nevertheless highly uniform, being in particular independent of the space, the measurable sets, and the measure. As already mentioned in the introduction, the authors of [2] presumed that this independence can be explained using some instantiation of the notion of majorizability. In that way, the metatheorems proved before and the discussion in this section show that this intuition was correct and that the uniformities are a necessary consequence of the novel form of majorizability introduced in this paper.

However, as already discussed in the introduction, we want to mention that the focus on the work [2] shall be understood to be merely indicative of the usefulness of the systems and metatheorems discussed before. In particular, essentially all other quantitative works on probability theory in the spirit of proof mining that have so far been considered in the literature can be similarly recognized as instances of the metatheorems proved here. Similarly, there too, the peculiar uniformities observed in practice are a priori guaranteed by the approach towards the metatheorems chosen here.

Now, the work [2] is concretely concerned with interrelations between different modes of convergence for sequences of random variables. The most prominent of these modes, also based on its similarity to a usual notion of pointwise convergence of functions, is that of almost sure convergence.

Definition 9.1 (Almost sure convergence). Let $(\Omega, S, \mathbb{P})$ be a probability space and (X_n) be a sequence of random variables $X_n : \Omega \to \mathbb{R}$ (i.e., X_n is measurable w.r.t. S and the Borel σ-algebra on $\mathbb{R}$). Then (X_n) is said to converge almost surely to a random variable $X : \Omega \to \mathbb{R}$ if

$$\mathbb{P}(\{x \in \Omega \mid X_n(x) \to X(x)\}) = 1.$$

This notion of almost sure convergence does not lend itself easily to a quantitative account of that convergence. Thus, in many cases where probability theorists are concerned with quantitative results (see, e.g., [39, 55]), they opt for a different, but equivalent, formulation of almost sure convergence known as almost uniform convergence.

Definition 9.2 (Almost uniform convergence). Let $(\Omega, S, \mathbb{P})$ be a probability space and (X_n) be a sequence of random variables $X_n : \Omega \to \mathbb{R}$. Then (X_n) is said to converge almost uniformly[22] to a random variable $X : \Omega \to \mathbb{R}$ if for all $\varepsilon, \delta > 0$ there exists an $N \in \mathbb{N}$ such that

$$\mathbb{P}(\{x \in \Omega \mid \forall n \geq N(|X_n(x) - X(x)| \leq \varepsilon)\}) > 1 - \delta.$$

The seminal result that these two notions of convergence are indeed equivalent is known as Egorov's theorem [9]. Note that since $(\Omega, S, \mathbb{P})$ is a probability space and the X_n's are random variables (and therefore measurable), the sets we take the probability of in the above definitions are measurable sets.

[22] Almost uniform convergence is usually formulated by requiring that for every $\varepsilon > 0$, there exists a measurable set $F_\varepsilon \in S$ with $\mathbb{P}(F_\varepsilon) \leq \varepsilon$ such that (X_n) converges uniformly on F_ε^c to X. The (equivalent) formulation given here is, however, more fruitful from an applied proof-theoretic perspective as it immediately allows for a very natural Cauchy-type variant.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

This result was analyzed quantitatively in [2] and then was used in turn also to provide a quantitative dominated convergence theorem for the Lebesgue integral. With the results from this section, we will be able to recognize this analysis as an instance of the preceding metatheorems for probability contents on algebras.

We now move towards formalizing the results from [2] and for that purpose introduce a sequence of random variables into the system $\mathcal{F}^\omega[\mathbb{P}, \text{Integral}]$ by adding an additional constant $X^{1(\Omega)(0)}$ together with the axiom

$$\forall n^0 (X(n) \in I)$$

to stipulate that the sequence in question belongs to our subspace of bounded and weakly Borel-measurable functions. It is clear that Theorem 8.9 extends from $\mathcal{F}^\omega[\mathbb{P}, \text{Integral}]$ to its extension by this constant X as any $X(n)$ is trivially majorizable as it is bounded and the whole constant X is thus majorized by a maximum construction. In that system, using axiom $(I)_3$ that asserts the weak Borel-measurability of any $X(n)$, it is then in particular a consequence of Π_1^Ω-AC (and actually of $\mathsf{b}\text{-}\mathsf{AC}_{\Omega,S}$ by regarding quantification over the type S as bounded) that there exists a functional $P^{S(0)(0)(0)}$ such that

$$\forall a^0, b^0, c^0, x^\Omega (x \in P(a,b,c) \leftrightarrow x \in (|X(c) - X(b)|)^{-1}([0, 2^{-a}])).$$

We add such a P directly into the language of the system via a new constant of the appropriate type together with the above defining property as an axiom and denote the resulting system by $\mathcal{F}^\omega[\mathbb{P}, \text{Integral}, X]$.

Using this functional P, we can then formally introduce the alternative way of formulating almost uniform convergence using finite unions as (implicitly) introduced in [2], which allows for both a natural metastable variant as well as to extend any discussion regarding this notion naturally to the context of contents:

Definition 9.3. We say that X converges almost uniformly with respect to finite unions if

$$\forall k^0, a^0 \exists b^0 \forall c^0 \left(\mathbb{P}\left(\bigcup_{i=b}^{c} \bigcup_{j=b}^{c} P(a,i,j)^c \right) \leq_{\mathbb{R}} 2^{-k} \right).$$

It is rather immediately clear that this notion of almost uniform convergence w.r.t. finite unions is equivalent over probability spaces to the usual notion of almost uniform convergence. Further, we want to emphasize that this mode was not explicitly introduced in [2] but rather *implicitly* as already mentioned above as it is naturally suggested through the quantitative rendering of almost uniform convergence used in [2] in their main quantitative result given in Theorem 3.1. Concretely, one immediately finds that a solution of the monotone functional interpretation of the negative translation of almost uniform convergence w.r.t. finite unions is exactly a function $M(k,a)$ providing a 2^{-k}-uniform bound for the 2^{-a}-metastable convergence of the sequence coded by X as introduced in [2].

Similarly, we can also give a formal representation of the notion of almost uniform metastable pointwise convergence as introduced in [2] in the context of the system $\mathcal{F}^\omega[\mathbb{P}, \text{Integral}, X]$:

Definition 9.4. We say that X converges almost uniform metastable pointwisely if

$$\forall k^0, a^0, F^1 \exists b^0 \left(\mathbb{P}\left(\bigcap_{m=0}^{b} \bigcup_{i=m}^{F(m)} \bigcup_{j=m}^{F(m)} P(a,i,j)^c \right) \leq_{\mathbb{R}} 2^{-k} \right).$$

Contrary to Definition 9.3, this mode of convergence was explicitly introduced in [2] as a "more quantitatively friendly" version of the notion of almost sure convergence. In particular, as shown by Proposition 4.1 in [2], the notion of almost uniform metastable pointwise convergence coincides over probability spaces with that of almost sure convergence. Also, as essentially already observed in [2], a

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

solution to the monotone functional interpretation (of the negative translation of) the statement of almost uniform metastable pointwise convergence is exactly a function $M'(k,a)$ providing a 2^{-k}-uniform bound on the 2^{-a}-metastable pointwise convergence of the sequence coded by X as introduced in [2].

In [2], the authors now provide a quantitative version of Egorov's theorem by constructing a solution of the monotone functional interpretation of (the negative translation of) almost uniform convergence w.r.t. finite unions from a solution of the monotone functional interpretation of (the negative translation of) the statement of almost uniform metastable pointwise convergence. In that way, they in particular provide a uniform quantitative rendering of the corresponding implication. We justify this application by showing in the following that already the system $\mathcal{F}^{\omega}[\mathbb{P}, \text{Integral}, X]$ proves this implication between the above two modes of convergence. Besides thereby explaining the success and the uniformities of the quantitative version of Egorov's theorem from [2], the provability in $\mathcal{F}^{\omega}[\mathbb{P}, \text{Integral}, X]$ established here shows in particular that the corresponding results from [2] are true for probability contents and not just probability measures as illustrated in [2]. That is, the authors inadvertently provided an "Egorov-like theorem" for probability contents. Concretely, this seems to be in particular due to the above renderings of the notions of almost sure and almost uniform convergence introduced via a finitary perspective informed by proof mining in [2], which provide exactly those alternative phrasings of these notions that are much more nicely compatible with the notion of a content due to a computationally effective formulation using finite unions. In that way, the results from this section tie into the comments made in the introduction that the notions and proofs produced through the finitary perspective of proof mining seem to be suitable so that they allow for a simultaneous lift of the results to the theory of contents.

We now first note that one direction of that equivalence can be easily witnessed in the system $\mathcal{F}^{\omega}[\mathbb{P}, \text{Integral}, X]$ discussed previously.

Theorem 9.5. *The system $\mathcal{F}^{\omega}[\mathbb{P}, \text{Integral}, X]$ proves that if X converges almost uniformly with respect to finite unions, then X converges almost uniformly metastable pointwisely.*

Proof. We reason in $\mathcal{F}^{\omega}[\mathbb{P}, \text{Integral}, X]$. Let k, a, and F be given. Using that X converges almost uniformly w.r.t. finite unions, there exists a b such that

$$\mathbb{P}\left(\bigcup_{i=b}^{c}\bigcup_{j=b}^{c} P(a,i,j)^c\right) \leq 2^{-k}$$

for all c. Now, we in particular have

$$\bigcap_{m=0}^{b}\bigcup_{i=m}^{F(m)}\bigcup_{j=m}^{F(m)} P(a,i,j)^c \subseteq \bigcup_{i=b}^{F(b)}\bigcup_{j=b}^{F(b)} P(a,i,j)^c$$

and therefore

$$\mathbb{P}\left(\bigcap_{m=0}^{b}\bigcup_{i=m}^{F(m)}\bigcup_{j=m}^{F(m)} P(a,i,j)^c\right) \leq \mathbb{P}\left(\bigcup_{i=b}^{F(b)}\bigcup_{j=b}^{F(b)} P(a,i,j)^c\right) \leq 2^{-k}$$

follows from the monotonicity of $\mathbb{P}$. This yields that X converges almost uniformly metastable pointwisely. □

As mentioned before, a quantitative version of the converse of the above theorem is one of the main results of [2] and to obtain this, the authors of [2] mainly utilized a quantitative version of the following property about sequences of events.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

Theorem 9.6 (Theorem 2.2 of [2]). *For every sequence of events (A_n), any functional $M : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ and any $\lambda > \lambda' > 0$, there exists an $n \in \mathbb{N}$ such that*

$$\mathbb{P}\left(\bigcap_{m=0}^{M(F)} \bigcup_{j=m}^{F(m)} A_j\right) < \lambda' \text{ for all } F : \mathbb{N} \to \mathbb{N}$$

implies $\mathbb{P}(A_n) < \lambda$.

We now discuss in the following how (the proof of) this result can be formalized in our system for probability contents on algebras $\mathcal{F}^\omega[\mathbb{P}]$, justifying the existence and the uniformity of the quantitative result given in [2] by means of our metatheorems. Concretely, we show:

Theorem 9.7. *The system $\mathcal{F}^\omega[\mathbb{P}]$ proves:*

$$\forall A^{S(0)}, M^{0(1)}, u^0, v^0 >_0 u \exists n^0 \left(\forall F^1 \left(\mathbb{P}\left(\bigcap_{m=0}^{M(F)} \bigcup_{j=m}^{F(m)} A(j)\right) \leq_{\mathbb{R}} 2^{-v}\right) \to \mathbb{P}(A(n)) <_{\mathbb{R}} 2^{-u}\right).$$

In particular, as this theorem of $\mathcal{F}^\omega[\mathbb{P}]$ has the correct logical form, our main Theorem 8.9 on the extraction of uniform computable bounds applies and we thus find that the existence of a computable bound on the existential quantifier on n can be guaranteed to exist a priori and even further, based on our notion of majorizability, it can be guaranteed that this bound is independent of the content space and the sequence of events which matches exactly the properties of the bound explicitly calculated in [2].

To now demonstrate Theorem 9.7, we in particular rely on the following lemma:

Lemma 9.8. *The system $\mathcal{F}^\omega[\mathbb{P}]$ proves:*

$$\forall A^{S(0)}, k^0 \exists N^0 \forall n^0 \left(\mathbb{P}\left(\bigcup_{i=0}^{n} A(i) \cap \left(\bigcup_{i=0}^{N} A(i)\right)^c\right) <_{\mathbb{R}} 2^{-k}\right).$$

Proof. We reason in $\mathcal{F}^\omega[\mathbb{P}]$. Let $A^{S(0)}$ and k^0 be given. At first, note that Proposition 4.5 implies that

$$\exists N \forall n \left(n \geq N \to \left|\sum_{i=0}^{n} \mathbb{P}((A\uparrow)(i)) - \sum_{i=0}^{N} \mathbb{P}((A\uparrow)(i))\right| < 2^{-k}\right). \tag{$*$}$$

Take such an N and let n be arbitrary. If $n < N$, then

$$\bigcup_{i=0}^{n} A(i) \cap \left(\bigcup_{i=0}^{N} A(i)\right)^c = \emptyset$$

and so by extensionality of $\mathbb{P}$, we get

$$\mathbb{P}\left(\bigcup_{i=0}^{n} A(i) \cap \left(\bigcup_{i=0}^{N} A(i)\right)^c\right) = 0$$

and are done. So suppose $n \geq N$. Then by $(*)$, we get

$$\left|\sum_{i=0}^{n} \mathbb{P}((A\uparrow)(i)) - \sum_{i=0}^{N} \mathbb{P}((A\uparrow)(i))\right| < 2^{-k}.$$

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

Since all the $(A\uparrow)(i)$ are disjoint (by definition of $A\uparrow$) and since we have $\bigcup_{i=0}^{j}(A\uparrow)(i) = \bigcup_{i=0}^{j} A(i)$ for any j, we immediately derive

$$\sum_{i=0}^{j} \mathbb{P}((A\uparrow)(i)) = \mathbb{P}\left(\bigcup_{i=0}^{j} A(i)\right)$$

for any j by finite additivity and extensionality of $\mathbb{P}$. Thus, we in particular have

$$\left|\mathbb{P}\left(\bigcup_{i=0}^{n} A(i)\right) - \mathbb{P}\left(\bigcup_{i=0}^{N} A(i)\right)\right| < 2^{-k}$$

and since $n \geq N$ implies $\bigcup_{i=0}^{N} A(i) \subseteq \bigcup_{i=0}^{n} A(i)$, we obtain

$$\mathbb{P}\left(\bigcup_{i=0}^{n} A(i) \cap \left(\bigcup_{i=0}^{N} A(i)\right)^{c}\right) = \mathbb{P}\left(\bigcup_{i=0}^{n} A(i)\right) - \mathbb{P}\left(\bigcup_{i=0}^{N} A(i)\right) < 2^{-k}$$

by Proposition 4.3. □

With that lemma, we are now in the position for a formal proof of the main combinatorial theorem from [2]:

Proof of Theorem 9.7. Let $A^{S(0)}$, $M^{0(1)}$, u^0 and v^0 with $v > u$ be given and suppose

$$\forall F^1\left(\mathbb{P}\left(\bigcap_{m=0}^{M(F)} \bigcup_{j=m}^{F(m)} A(j)\right) \leq 2^{-v}\right).$$

So, by the previous Lemma 9.8 applied to the sequence of events $f_m^{S(0)}$ defined by $f_m(k) = A(k+m)$, we have

$$\forall m \exists N \forall n\left(\mathbb{P}\left(\bigcup_{i=m}^{n+m} A(i) \cap \left(\bigcup_{i=m}^{N+m} A(i)\right)^{c}\right) < \frac{2^{-u} - 2^{-v}}{2^{m+1}}\right)$$

and so in particular

$$\forall m \exists N \geq m \forall n \geq m\left(\mathbb{P}\left(\bigcup_{i=m}^{n} A(i) \cap \left(\bigcup_{i=m}^{N} A(i)\right)^{c}\right) < \frac{2^{-u} - 2^{-v}}{2^{m+1}}\right).$$

Thus, using AC (actually, by switching from $<$ to $\leq$ in the above formulation and manipulating the bound slightly, Π_1^0-AC suffices), there exists a functional F^1 such that for all m and $n \geq m$:

$$\mathbb{P}\left(\bigcup_{i=m}^{n} A(i) \cap \left(\bigcup_{i=m}^{F(m)} A(i)\right)^{c}\right) < \frac{2^{-u} - 2^{-v}}{2^{m+1}}.$$

It is now easy to see that for this functional F, we have

$$\begin{aligned} A(M(F)) &\subseteq \bigcap_{m=0}^{M(F)} \bigcup_{i=m}^{M(F)} A(i) \\ &\subseteq \left(\bigcap_{m=0}^{M(F)} \bigcup_{j=m}^{F(m)} A(j)\right) \cup \bigcup_{m=0}^{M(F)}\left(\bigcup_{i=m}^{M(F)} A(i) \cap \left(\bigcup_{j=m}^{F(m)} A(j)\right)^{c}\right) \end{aligned}$$

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

and so, by the sub-additivity and monotonicity of $\mathbb{P}$, we derive

$$\mathbb{P}(A(M(F))) < 2^{-v} + \sum_{m=0}^{M(F)} \frac{2^{-u} - 2^{-v}}{2^{m+1}} < 2^{-u}$$

and so taking $n := M(F)$ for this functional F yields the claim. □

Using this combinatory lemma, we can now also prove the converse of Theorem 9.5 and thereby exhibit how a quantitative solution for Theorem 9.7 as obtained in [2] immediately can be used in conjunction with the proof-theoretic metatheorem established in Theorem 8.9 to derive a quantitative version of Egorov's theorem in the sense of the above notions incorporating finite unions as presented in Theorem 3.1 of [2] and in particular guarantees the observed uniformities of the rates a priori.

Theorem 9.9. *The system* $\mathcal{F}^{\omega}[\mathbb{P}, \text{Integral}, X]$ *proves that if X converges almost uniformly metastable pointwisely, then X converges almost uniformly with respect to finite unions.*

Proof. Suppose that X does not converge almost uniformly with respect to finite unions, that is, that we have k and a such that

$$\forall m \exists g \left(\mathbb{P}\left(\bigcup_{i=m}^{g} \bigcup_{j=m}^{g} P(a,i,j)^c \right) > 2^{-k} \right).$$

Using QF-AC (after suitably prenexing the hidden quantifiers), we get a functional G such that

$$\forall m \left(\mathbb{P}\left(\bigcup_{i=m}^{G(m)} \bigcup_{j=m}^{G(m)} P(a,i,j)^c \right) > 2^{-k} \right).$$

For a contradiction, suppose now that X converges almost uniformly metastable pointwisely. By instantiating the corresponding notion with $k+1$ and a, we get

$$\forall F \exists b \left(\mathbb{P}\left(\bigcap_{m=0}^{b} \bigcup_{i=m}^{F(m)} \bigcup_{j=m}^{F(m)} P(a,i,j)^c \right) \leq 2^{-(k+1)} \right).$$

Thus, by QF-AC (again after suitably prenexing), we get a functional M such that

$$\forall F \left(\mathbb{P}\left(\bigcap_{m=0}^{M(F)} \bigcup_{i=m}^{F(m)} \bigcup_{j=m}^{F(m)} P(a,i,j)^c \right) \leq 2^{-(k+1)} \right).$$

Defining $M'(F) = M(\lambda n.\tilde{G}(F(n)))$ where $\tilde{G}(n) = \max_{m \leq n} G(m)$, we get

$$\forall F \left(\mathbb{P}\left(\bigcap_{m=0}^{M'(F)} \bigcup_{i=m}^{\tilde{G}(F(m))} \bigcup_{j=m}^{\tilde{G}(F(m))} P(a,i,j)^c \right) \leq 2^{-(k+1)} \right).$$

We now define a sequence of events A via

$$A(n) := \bigcup_{i=n}^{G(n)} \bigcup_{j=n}^{G(n)} P(a,i,j)^c.$$

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

Then we have

$$\bigcup_{n=m}^{F(m)} A(n) = \bigcup_{n=m}^{F(m)} \bigcup_{i=n}^{G(n)} \bigcup_{j=n}^{G(n)} P(a,i,j)^c \subseteq \bigcup_{i=m}^{\tilde{G}(F(m))} \bigcup_{j=m}^{\tilde{G}(F(m))} P(a,i,j)^c$$

for all m and F and therefore this implies

$$\forall F\left(\bigcap_{m=0}^{M'(F)} \bigcup_{n=m}^{F(m)} A(n) \subseteq \bigcap_{m=0}^{M'(F)} \bigcup_{i=m}^{\tilde{G}(F(m))} \bigcup_{j=m}^{\tilde{G}(F(m))} P(a,i,j)^c\right).$$

By the monotonicity of $\mathbb{P}$, we get

$$\forall F\left(\mathbb{P}\left(\bigcap_{m=0}^{M'(F)} \bigcup_{n=m}^{F(m)} A(n)\right) \leq 2^{-(k+1)}\right),$$

which yields, by Theorem 9.7, that there exists an m such that

$$\mathbb{P}(A(m)) = \mathbb{P}\left(\bigcup_{i=m}^{G(m)} \bigcup_{j=m}^{G(m)} P(a,i,j)^c\right) < 2^{-k}$$

which is a contradiction. □

As a last formal elucidation of some of the results from [2], we turn to Theorem 3.2 therein, where the authors provide a quantitative version for a special case of the dominated convergence theorem, strengthening preceding results from Tao [59]. Concretely, they assume for this special case that the random variables are positive and bounded (w.l.o.g.) by 1 (which immediately yields the "uniform" majorizability of the sequence X and thus guarantees the full independence of the rate from X via the preceding metatheorems). The following result now establishes that the corresponding infinitary convergence result can be proved in our system for integrals over probability contents $\mathcal{F}^{\omega}[\mathbb{P}, \text{Integral}, X]$ and therefore makes it possible to recognize the quantitative results extracted in [2] as an application of the metatheorems established in this paper.

Theorem 9.10. *The system* $\mathcal{F}^{\omega}[\mathbb{P}, \text{Integral}, X]$ *proves: if* X *converges almost uniformly metastable pointwisely and satisfies* $\forall n^0, x^{\Omega}(0 \leq_{\mathbb{R}} X(n)(x) \leq_{\mathbb{R}} 1)$*, it holds that*

$$\forall k^0 \exists n^0 \forall i^0, j^0\left(i, j \geq_0 n \to \left|\int X(i)\,\mathrm{d}\mathbb{P} - \int X(j)\,\mathrm{d}\mathbb{P}\right| \leq_{\mathbb{R}} 2^{-k}\right).$$

Proof. Let k be given. Since, by Theorem 9.9, X converges almost uniformly w.r.t. finite unions, there exists an n such that

$$\forall m\left(\mathbb{P}\left(\bigcup_{a=n}^{m} \bigcup_{b=n}^{m} P(k+1,a,b)^c\right) \leq 2^{-(k+2)}\right).$$

Take $i, j \geq n$ and define $m = \max\{i, j\}$ as well as

$$A := \bigcup_{a=n}^{m} \bigcup_{b=n}^{m} P(k+1,a,b)^c.$$

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

Similar to the proof of item (4) from Lemma 7.4, we have

$$\left|\int X(i)\,\mathrm{d}\mathbb{P} - \int X(j)\,\mathrm{d}\mathbb{P}\right| \leq \int |X(i) - X(j)|\,\mathrm{d}\mathbb{P}$$
$$= \int |X(i) - X(j)|\chi_A\,\mathrm{d}\mathbb{P} + \int |X(i) - X(j)|\chi_{A^c}\,\mathrm{d}\mathbb{P}.$$

As we have $|X(i)(x) - X(j)(x)| \leq 2$ for all x, we get

$$\int |X(i) - X(j)|\chi_A\,\mathrm{d}\mathbb{P} \leq \int 2\chi_A\,\mathrm{d}\mathbb{P} = 2\mathbb{P}(A) \leq 2^{-(k+1)}.$$

On the other hand, $x \in A^c$ yields

$$x \in \bigcap_{a=n}^{m}\bigcap_{b=n}^{m} P(k+1,a,b)$$

which in particular gives, by definition of m, that $x \in P(k+1,i,j)$. From the definition of P, this in particular implies that

$$|X(i)(x) - X(j)(x)| \leq 2^{-(k+1)}$$

and so we have $|X(i)(x) - X(j)(x)|\chi_{A^c}(x) \leq 2^{-(k+1)}$ which yields

$$\int |X(i) - X(j)|\chi_{A^c}\,\mathrm{d}\mathbb{P} \leq 2^{-(k+1)}$$

and we are done. □

10. Proof-theoretic transfer principles

In this last section, we present how our systems and metatheorems allow for the proof of a general type of result, which we call a proof-theoretic transfer principle, that allows one to transfer quantitative information on implications between modes of convergence of real numbers to corresponding quantitative information on implications between analogous modes of convergence for bounded random variables. In particular, as this type of reasoning is very common in the literature on the convergence of various iterations of random variables (see, e.g., [32] among many others), this transfer principle allows for a logical explanation of the strategy of providing a proof-theoretic analysis of such results in practice by mainly analyzing the underlying result on real numbers and then lifting this result together with some (often) simple modifications to random variables.

Concretely, to allow for a discussion of general modes of convergence for real numbers and random variables, we consider the following abstract formal setup: We throughout this section fix two Π_3-formulas

$$P(x^{1(0)}, \underline{p}^{\underline{\sigma}}) = \forall a^0 \exists b^0 \forall c^0 P_0(a,b,c,x,\underline{p})$$

and

$$Q(x^{1(0)}, \underline{p}^{\underline{\sigma}}) = \forall u^0 \exists v^0 \forall w^0 Q_0(u,v,w,x,\underline{p})$$

where P_0 and Q_0 are quantifier-free formulas which only have the indicated variables free. We understand P and Q as abstract representations of modes of convergence, with parameters $\underline{p}$, for a sequence of real numbers represented by x.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

For an example, we may take

$$P_0(a,b,c,x) := \forall i^0, j^0(b \leq_0 i, j \wedge i, j \leq_0 c \to |x(i) - x(j)| \in [0, 2^{-a}]), \tag{$\sim$}$$

using the previous intensional intervals (where the above statement can thus be regarded as a quantifier-free statement). In that case, P represents the usual Cauchy property for x.

To allow for a discussion of these modes applied to random variables, we extend the system $\mathcal{F}^\omega[\mathbb{P}]$ with four further constants

$$X^{1(\Omega)(0)},\ P^{S\underline{\sigma}^t(0)(0)(0)},\ Q^{S\underline{\sigma}^t(0)(0)(0)},\ \tau^{0(0)},$$

together with the axioms

$$\begin{gathered}\forall \underline{p^\sigma}, a^0, b^0, c^0, z^\Omega(z \in P(a,b,c,\underline{p}) \leftrightarrow P_0(a,b,c,\lambda n.X(n)(z),\underline{p})),\\ \forall \underline{p^\sigma}, a^0, b^0, c^0, z^\Omega(z \in Q(a,b,c,\underline{p}) \leftrightarrow Q_0(a,b,c,\lambda n.X(n)(z),\underline{p})),\\ \forall n^0, z^\Omega(\tau(n) \geq_{\mathbb{R}} |X(n)(z)|),\end{gathered}$$

specifying that the properties P_0 and Q_0 (inducing the predicates P and Q) induce measurable sets pointwisely relative to the sequence of random variables[23] specified by X and that these random variables are all bounded via a suitable monotone sequence of bounds (i.e., that X as a constant is majorized by τ). It is clear that Theorem 8.9 extends to this system, which we denote by $\mathcal{U}^\omega$, as all constants are majorizable and since the new axioms are purely universal.

In this extended language, we can then provide a formula that represents the property P if suitably lifted to the sequence of random variables represented by X:

Definition 10.1. We say that X satisfies P almost uniformly, and write $P(X)$ a.u., if

$$\forall \underline{p^\sigma}, k^0, a^0 \exists b^0 \forall c^0 \left(\mathbb{P}\left(P(a,b,c,\underline{p})^c\right) \leq_{\mathbb{R}} 2^{-k}\right).$$

Similarly, we define $Q(X)$ a.u.

If we consider the previous example for P_0 given in ($\sim$), then by formulating $P(X)$ a.u. in this case we recover the notion of almost uniform convergence with respect to finite unions as given in Definition 9.3 (i.e., the variant of almost uniform convergence implicitly considered in [2]).

We now turn to our main result that provides a relationship between statements of the form

$$\forall \underline{p^\sigma}, x^{1(0)}(P(x,\underline{p}) \to Q(x,\underline{p}))$$

and statements of the form

$$P(X) \text{ a.u.} \to Q(X) \text{ a.u.}$$

and which thereby not only establishes an upgrade-type theorem from relations between modes of convergence for sequences of reals to sequences of random variables but also allows for a transfer of the computational information obtainable for the implication in the premise to the implication in the conclusion.

Theorem 10.2. *Provably in $\mathcal{U}^\omega$, given functionals V, A, C such that*

$$\forall x, \underbrace{\underline{p}, x^*, B, u, w}_{\omega}\Big(\forall n(x^*(n) \geq |x(n)|) \wedge P_0(A\omega, B(A\omega), C\omega, x, \underline{p}) \to Q_0(u, V\underline{p}x^*Bu, w, x, \underline{p})\Big),$$

[23] If we assume that X is weakly Borel-measurable in the context of the above example ($\sim$), the corresponding point sets P, Q above are indeed measurable.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

we can construct V', A', C' such that

$$\forall \underbrace{\underline{p}, B, k, u, w}_{\alpha}\Big(\mathbb{P}(P(A'\alpha, Bk(A'\alpha), C'\alpha, \underline{p})^c) \leq 2^{-k} \to \mathbb{P}(Q(u, V'\underline{p}Bku, w, \underline{p})^c) \leq 2^{-k}\Big).$$

Proof. Given such V, A, C, and $\alpha = (\underline{p}, B, k, u, w)$, we define

$$\begin{aligned} A'\alpha &:= A\underline{p}\tau(Bk)uw, \\ C'\alpha &:= C\underline{p}\tau(Bk)uw, \\ V'\underline{p}Bku &:= V\underline{p}\tau(Bk)u. \end{aligned}$$

Let z be arbitrary with $z \in Q(u, V'\underline{p}Bku, w, \underline{p})^c$. By the axioms of $\mathcal{U}^\omega$ and the definition of V', we have

$$\begin{aligned} z \in Q(u, V'\underline{p}Bku, w, \underline{p})^c &\leftrightarrow z \in Q(u, V\underline{p}\tau(Bk)u, w, \underline{p})^c \\ &\leftrightarrow \neg Q_0(u, V\underline{p}\tau(Bk)u, w, \lambda n.X(n)(z), \underline{p}) \end{aligned}$$

and the latter implies

$$\neg P_0(A\underline{p}\tau(Bk)uw, Bk(A\underline{p}\tau(Bk)uw), C\underline{p}\tau(Bk)uw, \lambda n.X(n)(z), \underline{p})$$

using the assumptions on V, A, C and that $\tau(n) \geq |X(n)(z)|$ for any z. This in turn is by definition of A', V', C' equivalent to

$$\neg P_0(A'\alpha, Bk(A'\alpha), C'\alpha, \lambda n.X(n)(z), \underline{p})$$

and thus to

$$z \in P(A'\alpha, Bk(A'\alpha), C'\alpha, \underline{p})^c.$$

Thus, we have

$$Q(u, V'\underline{p}Bku, w, \underline{p})^c \subseteq P(A'\alpha, Bk(A'\alpha), C'\alpha, \underline{p})^c$$

as z above was arbitrary and therefore, we get

$$\mathbb{P}(Q(u, V'\underline{p}Bku, w, \underline{p})^c) \leq \mathbb{P}(P(A'\alpha, Bk(A'\alpha), C'\alpha, \underline{p})^c)$$

by the monotonicity of $\mathbb{P}$. This yields the claim. □

This result, while at first glance rather technical and abstract, has a very concrete use recently observed in applications by the first author [43] and to illustrate this, we will now shortly discuss the extent of the above result and its use in mathematics:

1. Observe that the conclusion of Theorem 10.2 is just a witnessed version of the Dialectica interpretation of

$$P(X) \text{ a.u.} \to Q(X) \text{ a.u.} \tag{+}$$

and therefore, under AC, this witnessed Dialectica interpretation in particular implies (+). In that way, whenever the premise of Theorem 10.2 is established in $\mathcal{U}^\omega$, one immediately obtains the truth of (+) and so the above result allows for a lift from a (quantitative) result on real numbers to a true result for random variables. Furthermore, another main benefit of the conclusion of Theorem 10.2 is that it allows for the extraction of quantitative information in the sense that the functional V' provides

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

a transformation of a rate for the premise $P(X)$ a.u. into a rate for the conclusion $Q(X)$ a.u. Even further, as V' can be constructed from V, this transformation of rates can be directly inferred from the transformation of rates V of the presumed result for real numbers.

2. The premise of Theorem 10.2 is "essentially" the Dialectica interpretation of the statement

$$\forall \underline{p}^{\sigma}, x^{1(0)}(P(x,\underline{p}) \to Q(x,\underline{p})) \tag{++}$$

in the sense that the functionals V, A, C represent realizers for this interpretation, with the additional assumption that these realizers are suitably uniform, depending only on upper bounds of the sequence $x^{1(0)}$. Although one can construct examples where such realizers do not possess this kind of uniformity, in practice, for many theorems of the form (++) that have a semi-constructive proof, such uniform realizers can be given. In particular, this is true for the forthcoming work by the first author on Kronecker's Lemma [43] and an upcoming work by Oliva and Arthan on quantitative stochastic optimization [48]. In both these cases, one uses reasoning about sequences of real numbers to obtain the analogous result about sequences of random variables and a computational interpretation can be given to this line of reasoning. Theorem 10.2 then in particular provides an abstract generalization of this procedure and explains how this reasoning is substantiated by logical results.

Lastly, in the following remark we discuss a counterexample illustrating the necessity of the majorizability of the sequence of random variables in Theorem 10.2:

Remark 10.3. For the above transfer principle to hold, the assumption of the boundedness of the sequence of random variables is necessary as the following example shows: Take $\Omega := \mathbb{N}$ and let S be the collection of all finite and co-finite subsets of $\mathbb{N}$, that is,

$$S := \{A \subseteq \mathbb{N} \mid A \text{ is finite or } A^c \text{ is finite}\}.$$

Furthermore, define the content $\mathbb{P}$ by $\mathbb{P}(A) = 0$ if A is finite and $\mathbb{P}(A) = 1$ if A^c is finite, for all $A \in S$. Now, we consider the two properties

$$P(x) \equiv P_0(x) \equiv 0 = 0 \text{ and } Q(x) \equiv \exists n \forall m Q_0(n,m,x) \equiv \exists n \forall m (n \geq_{\mathbb{Q}} [x_0](m))$$

for a sequence $x = (x_n)$ of type $1(0)$. Clearly, both P and Q are Π^0_3-formulas and are trivially true for all sequences x. Therefore also $P(x) \to Q(x)$ is trivially true. Further, we can easily give V, A, C that satisfy the assumptions of Theorem 10.2. Now, consider

$$X_n : \mathbb{N} \to \mathbb{N}, \quad k \mapsto (k)_{\mathbb{Q}}$$

for each n. Then the set $Q(n,m)$ corresponding to Q_0 is just

$$Q(n,m) = \{k \in \mathbb{N} \mid Q_0(n,m,\lambda l.X_l(k)\} = \{k \in \mathbb{N} \mid n \geq_{\mathbb{Q}} [X_0(k)](m)\} = \{k \in \mathbb{N} \mid n \geq k\}$$

which belongs to S as it is finite. P_0 is just represented by the full set $\mathbb{N}$. Therefore, X satisfies P almost uniformly and does not satisfy Q almost uniformly as any $Q(n,m)$ has measure 0.

Acknowledgements. Both authors want to thank Thomas Powell, Ulrich Kohlenbach, Henry Towsner, and José Iovino for insightful discussions on the topics of this paper. We also want to thank the anonymous referee for the very careful reading of the manuscript and the resulting helpful suggestions which improved the paper in various places, in particular its presentation.

Competing interests. The authors have no competing interest to declare.

Financial support. The first author was partially supported by the EPSRC Centre for Doctoral Training in Digital Entertainment (EP/L016540/1). The second author was supported by the 'Deutsche Forschungsgemeinschaft' Project DFG KO 1737/6-2.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

References

[1] R. Arthan and P. Oliva, 'On the Borel-Cantelli Lemmas, the Erdős-Rényi Theorem, and the Kochen-Stone Theorem', *J. Log. Anal.* **13**(6) (2021), 1–23.
[2] J. Avigad, E. Dean, and J. Rute, 'A metastable dominated convergence theorem', *J. Log. Anal.* **4**(3) (2012), 1–19.
[3] J. Avigad, P. Gerhardy, and H. Towsner, 'Local stability of ergodic averages', *Trans. Amer. Math. Soc.* **362**(1) (2010), 261–288.
[4] J. Avigad and J. Iovino, 'Ultraproducts and metastability', *New York J. Math.* **19** (2013), 713–727.
[5] M. Bezem, 'Strongly majorizable functionals of finite type: a model for bar recursion containing discontinuous functionals', *J. Symb. Log.* **50** (1985), 652–660.
[6] K.P.S. Bhaskara Rao and M. Bhaskara Rao, 'Theory of charges', in *Pure and Applied Mathematics*. vol. 109 (Academic Press, London, 1983).
[7] P.J. Daniell, 'A general form of integral', *Ann. Math.* **19**(4) (1918), 279–294.
[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*, vol. 1 (Chapman and Hall/CRC, New York, 2017), 131–187.
[9] D. Egoroff, 'Sur les suites des fonctions mesurables', *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, **152** (1911), 244–246.
[10] F. Ferreira, L. Leuştean, and P. Pinto, 'On the removal of weak compactness arguments in proof mining', *Adv. Math.* **354** (2019), 106728.
[11] F. Ferreira and P. Oliva, 'Bounded functional interpretation', *Ann. Pure Appl. Logic* **135** (2005), 73–112.
[12] P. Gerhardy and U. Kohlenbach, 'General logical metatheorems for functional analysis', *Trans. Amer. Math. Soc.* **360** (2008), 2615–2660.
[13] K. Gödel, 'Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes', *Dialectica* **12** (1958), 280–287.
[14] I. Goldbring and H. Towsner, 'An approximate logic for measures', *Israel J. Math.* **199**(2) (2014), 867–913.
[15] A. Grzegorczyk, 'Some classes of recursive functions', in *Rozprawny Matematyczne*, (Instytut Matematyczny Polskiej Akademi Nauk, Warsaw, 1953).
[16] D. Günzel and U. Kohlenbach, 'Logical metatheorems for abstract spaces axiomatized in positive bounded logic', *Adv. Math.* **290** (2016), 503–551.
[17] P.R. Halmos, *Measure Theory*, Graduate Texts in Mathematics (Springer, New York, NY, 1974).
[18] D. Hilbert, 'Über das Unendliche', *Math. Ann.* **95**(1) (1926), 161–190.
[19] W.A. Howard, 'Hereditarily majorizable functionals of finite type', in A.S. Troelstra, editor, *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis* vol. 344 of *Lecture Notes in Mathematics* (Springer, New York, 1973), 454–461.
[20] U. Kohlenbach. *Theorie der majorisierbaren und stetigen Funktionale und ihre Anwendung bei der Extraktion von Schranken aus inkonstruktiven Beweisen: Effektive Eindeutigkeitsmodule bei besten Approximationen aus ineffektiven Eindeutigkeitsbeweisen*. PhD thesis, (Goethe-Universität Frankfurt am Main, 1990).
[21] U. Kohlenbach, 'Effective bounds from ineffective proofs in analysis: an application of functional interpretation and majorization', *J. Symb. Log.* **57**(1992), 1239–1273.
[22] U. Kohlenbach, 'Pointwise hereditary majorization and some applications', *Arch. Math. Log.* **31** (1992), 227–241.
[23] 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)* (Oxford University Press, Oxford, 1996), 225–260.
[24] U. Kohlenbach, 'Mathematically strong subsystems of analysis with low rate of growth of provably recursive functionals', *Arch. Math. Log.* **36** (1996), 31–71.
[25] U. Kohlenbach, 'Some logical metatheorems with applications in functional analysis', *Trans. Amer. Math. Soc.* **357**(1) (2005), 89–128.
[26] U. Kohlenbach, 'Applied Proof Theory: Proof Interpretations and their Use in Mathematics', Springer Monographs in Mathematics (Springer-Verlag, Berlin Heidelberg, 2008).
[27] U. Kohlenbach, 'Recent progress in proof mining in nonlinear analysis', *IFCoLog J. Log. App.* **10**(4) (2017), 3361–3410.
[28] U. Kohlenbach, 'Proof-theoretic Methods in Nonlinear Analysis', in B. Sirakov, P. Ney de Souza, and M. Viana, editors, *Proc. ICM 2018*, vol. 2 (World Scientific, Singapore, 2019), 61–82.
[29] U. Kohlenbach, 'Local formalizations in nonlinear analysis and related areas and proof-theoretic tameness', in P. Weingartner and H.-P. Leeb, editors, *Kreisel's Interests. On the Foundations of Logic and Mathematics*, vol. 41 of *Tributes* (College Publications, 2020), 45–61.
[30] U. Kohlenbach and A. Nicolae, 'A proof-theoretic bound extraction theorem for CAT(κ)-spaces', *Stud. Log.* **105** (2017), 611–624.
[31] U. Kohlenbach and P. Oliva, 'Proof Mining: A systematic way of analysing proofs in mathematics', *Proceedings of the Steklov Institute of Mathematics* **242** (2003), 136–164.
[32] A.N. Kolmogorov, 'Sur la loi forte des grands nombres', *Comptes rendus de l'Académie des Sciences* **191** (1930), 910–912.
[33] G. Kreisel, 'On the interpretation of non-finitist proofs—Part I', *J. Symb. Log.* **16**(4) (1951), 241–267.
[34] G. Kreisel, 'On the interpretation of non-finitist proofs–Part II. Interpretation of number theory. Applications', *J. Symb. Log.* **17**(1) (1952), 43–58.
[35] A. Kreuzer, 'Measure theory and higher order arithmetic', *Proc. Amer. Math. Soc.*, **143**(12), (2015), 5411–5425.

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press

[36] S. Kuroda, 'Intuitionistische Untersuchungen der formalistischen Logik', *Nagoya Math. J.* **2** (1951), 35–47.
[37] L. Leuştean, 'Proof mining in $\mathbb{R}$-trees and hyperbolic spaces', *Electronic Notes in Theoretical Computer Science* **165** (2006), 95–106.
[38] L. Leuştean, 'An application of proof mining to nonlinear iterations', *Ann. Pure App. Logic* **165**(9) (2014), 1484–1500.
[39] N. Luzia, 'A simple proof of the strong law of large numbers with rates', *Bull. Austral. Math. Soc.* **97**(3) (2018), 513–517.
[40] A. Macintyre, 'The mathematical significance of proof theory', *Phil. Trans. Royal Soc.* **363** (2005), 2419–2435.
[41] A. Macintyre, 'The impact of Gödel's incompleteness theorems on mathematics', in M. Baaz, C.H. Papadimitriou, H.W. Putnam, D.S. Scott, and jr. C.L. Harper, editors, *Kurt Gödel and the foundations of mathematics: Horizons of truth*, (Cambridge University Press, Cambridge, 2011), 3–25.
[42] M. Neri, 'Quantitative strong laws of large numbers', *Electronic Journal of Probability* **30**(20) (2024), 1–22.
[43] M. Neri, 'A finitary Kronecker's lemma and large deviations in the strong law of large numbers', *Ann. Pure Appl. Logic* **176**(6) (2025), 103569, 1–31.
[44] 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', *ArXiv e-prints*, 2024. arXiv, math.OC, 2504.12922
[45] 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*, vol. 15764 of *LNCS* (2025), 333–348.
[46] M. Neri and T. Powell, 'A quantitative Robbins-Siegmund theorem', *Annals of Applied Probability*, 2025. to appear.
[47] M. Neri and T. Powell, 'On quantitative convergence for stochastic processes: Crossings, fluctuations and martingales', *Transactions of the American Mathematical Society, Series B*, **12** (2025), 974–1019.
[48] P. Oliva, Personal communication, November 2023.
[49] L. Paunescu and A. Sipoş, 'A proof-theoretic metatheorem for tracial von Neumann algebras', *Math. Log. Q.* **69**(1) (2023), 63–76.
[50] N. Pischke, 'Logical Metatheorems for Accretive and (Generalized) Monotone Set-Valued Operators', *J. Math. Logic* **24**(2) (2024), 2350008, 1–59.
[51] N. Pischke, 'Proof Mining for the dual of a Banach space with extensions for uniformly Fréchet differentiable functions', *Trans. Amer. Math. Soc.* **377**(10) (2024), 7475–7517.
[52] N. Pischke, 'A proof-theoretic metatheorem for nonlinear semigroups generated by an accretive operator and applications', *Sel. Math*. **31**(2) (2025), 32, 1–67.
[53] N. Pischke and T. Powell, 'Asymptotic regularity of a generalised stochastic Halpern scheme', *ArXiv e-prints*, 2024. arXiv, math.OC, 2411.04845.
[54] M. Schönfinkel, 'Über die Bausteine der mathematischen Logik', *Math. Ann.* **92** (1924), 305–316.
[55] D. Siegmund, 'Large deviation probabilities in the strong law of large numbers', *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **31**(2) (1975), 107–113.
[56] A. Sipoş, 'Proof mining in L^p spaces', *J. Symb. Logic* **84**(4) (2019), 1612–1629.
[57] E. Specker, 'Nicht konstruktiv beweisbare Sätze der Analysis', *J. Symb. Logic* **14**(3) (1949), 145–158.
[58] 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, Proceedings of Symposia in Pure Mathematics*, vol. 5 (AMS, Providence, 1962), 1–27.
[59] T. Tao, 'Norm convergence of multiple ergodic averages for commuting transformations' *Ergod. Theory Dyn. Syst.* **28** (2008), 657–688.
[60] A.S. Troelstra, editor. *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, vol. 344 of *Lecture Notes in Mathematics* (Springer, Berlin, 1973).
[61] A.S. Troelstra and D. van Dalen, editors. 'Constructivism in Mathematics, vol. 1', vol. 121 of *Studies in Logic and the Foundations of Mathematics* (North-Holland, Amsterdam, 1988).

https://doi.org/10.1017/fms.2025.10138 Published online by Cambridge University Press