The Causal Axioms of Algebraic Quantum Field Theory: A Diagnostic

Francisco Calderón[*]

Department of Philosophy, University of Michigan, Ann Arbor, MI, USA

Abstract

Algebraic quantum field theory (AQFT) puts forward three "causal axioms" that aim to characterize the theory as one that implements relativistic causation: the spectrum condition, microcausality, and primitive causality. In this paper, I aim to show, in a minimally technical way, that none of them fully explains the notion of causation appropriate for AQFT because they only capture some of the desiderata for relativistic causation I state or because it is often unclear how each axiom implements its respective desideratum. After this diagnostic, I will show that a fourth condition, local primitive causality (LPC), fully characterizes relativistic causation in the sense of fulfilling all the relevant desiderata. However, it only encompasses the virtues of the other axioms because it is implied by them, as I will show from a construction by Haag and Schroer (1962). Since the conjunction of the three causal axioms implies LPC and other important results in QFT that LPC does not imply, and since LPC helps clarify some of the shortcomings of the three axioms, I advocate for a holistic interpretation of how the axioms characterize the causal structure of AQFT against the strategy in the literature to rivalize the axioms and privilege one among them.

---
[*]fcalder@umich.edu

1 Introduction

Philosophers have often praised algebraic quantum field theory's (AQFT's) mathematically precise, axiomatic framework. The theory's core is to assign to every open, bounded region of spacetime $\mathcal{O}$ a set of observables $\mathcal{A}(\mathcal{O})$ describing the physics in said region. For this paper, every $\mathcal{A}(\mathcal{O})$ is a von Neumann algebra that, in particular cases, can be isomorphic to $\mathcal{B}(\mathcal{H})$, the bounded, linear operators on some Hilbert space $\mathcal{H}$—a structure more familiar from ordinary quantum mechanics (QM). Although there has been much debate on which formulation of QFT we should use to address foundational, philosophical, and interpretive questions (Wallace, 2011; Fraser, 2011; Swanson, 2017), arguably, the fruitfulness of choosing AQFT lies in the rich and robust framework that its axioms provide.

However, as AQFT's advocates know, these axioms are very demanding from a mathematical and physical point of view, so we must be attentive to the role each axiom plays in capturing the physical features we want a QFT to have. It is then puzzling that AQFT lays down multiple axioms that prima facie look like they strive to do the same job—namely, characterizing the causal structure of a theory. The spectrum condition (SC), microcausality (MC), or primitive causality (PC) are all deemed **causal axioms**. From this point of view, it's not surprising that the tendency in the literature on relativistic causation in QFT has been to adopt the following attitude:

> **Atomistic conception of the causal axioms** (ATOMISTIC): One of the axioms implements relativistic causation in (or characterizes the causal relativistic structure of) QFT singlehandedly or in the most exemplary way.

The following passages are evidence of this attitude:

- "[SC] is perhaps the most direct expression of the prohibition of spacelike processes" (Butterfield, 2007, p. 303)

- "The usual causality postulate which is adopted by practically all authors may be expressed in the following way...[MC]" (Haag and Schroer, 1962, p. 250). "the latter axiom [MC] represents the most important specific feature of relativistic systems" (Horuzhy, 1990, p. 14). MC is still the single, most famous constraint among physicists, as most textbooks in conventional QFT can attest.

- "the most fundamental causality requirement in both classical relativistic field theory and relativistic AQFT is the requirement of no superluminal

propagation. [SC] has sometimes been touted as a prohibition on superluminal processes, but we will argue that it is no such thing. The proper expression of the principle of no superluminal propagation in AQFT...is satisfied in models of AQFT that obey [PC]." (Earman and Valente, 2014, p. 2)[1]

The first main upshot of this paper is to reject Atomistic: none of these causal axioms *fully* explains the notion of relativistic causation appropriate for AQFT, so we should not rivalize them and privilege one among them. As I will show, there are two kinds of shortcomings for defending Atomistic. First, the axioms only capture some (and not all) of the desiderata for relativistic causation that I will state in the next section. Second, it is often unclear how each axiom implements its respective desideratum.

The second main upshot is to advocate for the following attitude instead:

> **Holistic conception of the causal axioms** (Holistic): All of the axioms (and the connections between them) are required to implement relativistic causation in (or characterize the causal relativistic structure of) QFT.

The cornerstone for Holistic will be a fourth condition, **local primitive causality** (LPC) (also called the **diamond property** in older literature). LPC will embolden Holistic in two ways. First, LPC *does* fully characterize relativistic causation in the sense of fulfilling all the relevant desiderata. Second, it only encompasses the virtues of the other axioms because it is implied by them. More specifically, I prove that we can derive it from the rest of the axioms, which are known to be logically independent. While some readers might take LPC for granted, doing so would involuntarily assume Haag duality, a problematic assumption from the original derivation of LPC (Haag and Schroer, 1962). Additionally, this would obfuscate the connections between the axioms. So, to reiterate, LPC grants us holism in two ways: by interpreting the axioms as mathematically rigorous but mostly physically intuitive and interpretatively rich constraints on a QFT, but also by taking them as members of an axiomatic *system*.[2] As such, my worries driving this paper are very similar to

---

[1] I have renamed Earman and Valente's core causal condition to fit my terminology here. The reader should bear in mind that the literature on AQFT uses the same terms in different ways.

[2] As far as I know, there are no other explicit advocates of Holistic. Here are the two most similar claims I could find: "The principle of relativistic causality [MC] can be put into various mathematical forms and in terms of various objects: observables, states, elementary observables, corresponding to projection operators. And *it would be hard to say that any particular formulation includes the total contents of this broad physical principle.*" (1990, p. 20, emphasis added). However, Horuzhy is discussing different ways of expressing the content of MC only. The other claim is the following passage from Rédei (endorsed by (Hofer-Szabó and Vecsernyés, 2018, p. 17)): "relativistic locality is *not a single property* a

those of the creators of AQFT: beyond their mathematical consistency, are these physically meaningful and well-motivated requirements? Are they interdependent, and if so, do they codify compatible or rival features of the theory?

The methodology I will adopt to address those questions is that of a diagnostic that scrutinizes what it means to be a *causal* axiom. What this means is that to address a dispute between ATOMISTIC and HOLISTIC, we must go through each axiom aiming to take the "wide frame allowed by the general principles and, narrowing the...frame step by step,...look for the neuralgic spots." (Haag, 2010, p. 243). The main tools at our disposal are, to borrow the terminology from Peskin and Schroeder (1995), the three "themes" of QFT: the concept of field, QM, and special relativity (SR). In other words, an axiom of AQFT should implement the notion of field (or, in other words, make clear what is **local** in the sense of situating phenomena in delimited patches of spacetime, as opposed to **global**) and have a clear input from QM and SR. Now, while Earman and Valente (2014) have presented a similar overview and ended up with a form of ATOMISTIC, another difference between their diagnostic and mine is a deliberate attempt to be minimally technical, partly because of how inaccessible the literature on AQFT is, and partly to make the philosophical stakes of these discussions more salient. But this is not to say that the diagnostic preceding my discussion of LPC is a mere pedagogical rehashing of the literature: aside from assuming less from the reader, I will introduce new difficulties for interpreting the axioms and surface different considerations for thinking about them. With this in mind, I've also relegated the gory details of the proof that the causal axioms imply LPC to an appendix and technical details that might distract readers from the argument to footnotes.

2   Four desiderata for relativistic causation

Before moving on with the diagnostic and seeing which axioms have which symptoms, I should clarify what "relativistic causation" means. At this point, I also depart from the literature on relativistic causation since the following desiderata are trying to track the folk notions of causation used by physicists. Contrast this approach with Butterfield's (2007), which focuses on the notions of causation that are

---

physical theory can in principle have but an intricately interconnected web of features. Each of those individual features expresses some important aspect of relativistic locality, and a physical theory can in principle have some of these features but not others. A physical theory is in full compliance with relativistic locality if it possesses *all* the features in this web" (2014, p. 138). However, the crux of Rédei's holism is a unified framework from category theory. Not only can I make a similar claim with less sophisticated tools, but the holism brought about by LPC partially stems from fulfilling a set of physically motivated desiderata and not merely from belonging to a cohesive mathematical structure.

widespread in the philosophy of science literature. Now, the physics literature has multiple facets of the notion of "causation," so I will lay down multiple desiderata for relativistic causation.

First, as a form of "action," relativistic causation attempts to capture a notion of (i) locality[3] and of (ii) a finite speed of propagation of events:

> For the relative independence of spatially distanced things (A and B), this idea is characteristic: external influences from A have no *immediate* influence on B; this is known as the "principle of proximal action [*Nahewirkung*]," which is only used consistently in field theory. The complete abolition of this principle would make impossible...the formulation of empirically testable laws in our familiar sense. (Einstein, 1948, p. 321-322, my translation)

The worry about the *immediacy* of the propagation of this kind of influence encapsulates (i) and (ii) into the preclusion of superluminal signaling, arguably the most famous notion of relativistic causation. Events are *independent* in a relevant causal sense if we cannot connect them with a light signal. That is, we have merely specified what causal connections *cannot* be. "Relativistic causation" also includes a notion of causal *dependence*, not merely *in*dependence (Earman and Valente, 2014, p. 4). Here is an account from another renowned physicist:

> In physics, causal description...rests on the assumption that the knowledge of the state of a material system at a given time permits the prediction of its state at any subsequent time. (Bohr, 1948, p. 312)

Therefore, (relativistic) causation should also capture an idea of (iii) a deterministic connection between events. Finally, I will add (iv) respecting the metric structure of the underlying spacetime since Geroch (2011) has shown that SR is a viable theory even with superluminal signaling and since SR secures very few of the multiple locality conditions of AQFT (and the ones implementing SR are independent of the rest) (Ruetsche, 2021, p. 314). Additionally, some axioms are restricted to Minkowski spacetime while others only require that the spacetime is globally hyperbolic, so different metric structures avail different resources to interpret the content of the axioms.

---

[3]"Locality" is said in many ways. Specifically, it's sometimes used synonymously with "localization," "separability"/"independence," "no superluminal signaling," or an obverse of Bell non-locality (Earman and Valente, 2014, p. 3). From this list, I will discuss the first three. However, I'm not using the term in the fourth sense, famously absent in quantum theories, especially in QFT.

3 Symptomatology

3.1 The Spectrum Condition

The first axiom we will consider is the spectrum condition (SC):

> **SC**: (a) The joint spectrum of the infinitesimal generators of translations in Minkowski spacetime is confined to the forward lightcone. (b) There is a unique (up to phase factors) **vacuum state**, which is translationally invariant and has zero energy-momentum.

In other words, we will only consider states with positive energy-momentum (plus the vacuum). SC is usually supplemented with a local "counterpart," the **axiom of covariance**, which dictates how the algebra of a given region transforms under the unitary representations of Lorentz transformations. Since SC and the axiom of covariance cover the demands of the Poincaré group of SR, including both Lorentz transformations and spacetime translations, the axiom of covariance is usually taken as one of the causal axioms of the theory. Now, from this perspective of symmetries, they are both straightforward applications of Wigner's theorem: to obtain the quantum analog of the symmetry group of a classical theory (in this case, the Poincaré group), take the group's unitary representations. Thanks to Stone's theorem, those unitary representations can be further expressed in terms of their infinitesimal generators, which, for the translation subgroup, are the key ingredient for SC. So, the way in which SC implements QM and SR only codifies the geometry of Minkowski spacetime. That covers the fourth desideratum: SC respects the structure of (Minkowski) spacetime. What about the other desiderata? And is it not a shortcoming of SC as a causal axiom that we surely want to be able to claim that there is some form of "relativistic causation" in AQFT in curved spacetimes?

Let's start with the first question. SC is fairly silent about determinism, and although it's not local, the axiom of covariance is. As for forbidding superluminal signaling, its role (if it has one) needs to be clarified. SC does not hold for every QFT since the energy (density) of a *quantum* field need not be positive (Epstein et al., 1965). On top of that, the energy-momentum operators do not necessarily coincide with the energy-momentum (tensor) of the quantum field in question (Earman and Valente, 2014). Now, SC doesn't hold for every QFT, but quantum fields obtained by quantizing classical relativistic fields exhibit no superluminal signaling (Earman and Valente, 2014, pp. 17-18). Maybe there's something about field theories in general precluding superluminal signaling? It doesn't seem that way since the classical analogs of SC are not sufficient conditions to prohibit superluminal signaling (Earman, 2014, pp. 103-105). As for the second question, SC can be stated at every

point of a curved spacetime (Haag, 1996, p. 338), but (non-unique) generalizations of restrictions on the energy-momentum of the fields (Curiel, 2017) are extensions from the classical analogs of SC I just mentioned are not sufficient to prohibit superluminal signaling.

If SC (a) does not fulfill the goal we would want for *a single* causal axiom to have or to have in the most exemplary way, what about SC (b)? SC (b) is sometimes taken not as a characterization of relativistic causation in AQFT but, ironically, as one of the ingredients needed to question AQFT's status as a relativistic theory. One of the consequences of SC (when supplemented with a condition of "additivity" saying that observables can be expressed in terms of observables of arbitrarily small regions) is the **Reeh-Schlieder theorem**. The theorem claims that if we act on the vacuum defined by SC (b) with an element of $\mathcal{A}(\mathcal{O})$ for some region $\mathcal{O}$, we can generate any other state. This result is unsettling because it seems that we can measure the energy of the vacuum in a lab and yet approximate any other quantum state in any other region of spacetime, even states that *do* have energy! There are interpretive strategies and more sophisticated mathematical tools to get out of this problem (Halvorson, 2001; Wallace, 2006; Valente, 2014). Still, they do not change that SC (b) does not implement the prohibition of superluminal signaling, even if we manage to show that it does not impede it.

3.2 Microcausality

The second axiom we will consider is the microcausality (MC), Einstein causality, or local commutativity condition:

> **MC**: $\mathcal{A}(\mathcal{O}_1)$ commutes with $\mathcal{A}(\mathcal{O}_2)$ if $\mathcal{O}_1$ and $\mathcal{O}_2$ are space-like separated.

Let us unpack this condition: since no physical process can occur along a spacelike trajectory, no measurement in $\mathcal{O}_1$ can disturb the outcomes of a measurement carried out in $\mathcal{O}_2$, and vice-versa. The notion of spacelike separation is tracking the metric structure of spacetime (and can be extended to locally flat tangent spaces in curved spacetimes (Haag, 1996, p. 338)), and this condition is clearly local. Now, MC is often rephrased to say that there would be statistical correlations if we could connect $\mathcal{O}_1$ and $\mathcal{O}_2$ with a superluminal signal. However we want to read MC, it's an attempt to transcribe to the language of operator algebras the passage I cited from Einstein in the introduction. For that reason, MC is considered a condition of **independence** or **separability** between quantum systems in regions $\mathcal{O}_1$ and $\mathcal{O}_2$.

Prima facie, MC is a straightforward implementation of relativistic causation in the sense of no superluminal signaling. That is, MC attempts to capture the first two desiderata of relativistic causation that I had laid out. Aside from only capturing

some of the desiderata for relativistic causation in AQFT, this section will show that MC's attempt to prohibit superluminal signaling comes with a "morass of recalcitrant interpretational issues." (Earman and Valente, 2014, p. 16).

Since the way in which this axiom implements its causal desiderata comes off badly from my diagnostic, I want to stress its importance for QFT in general before moving on. First, MC can be easily corroborated for the most common field theories, or it is a crucial assumption in their construction. Consider a Klein-Gordon field, the simplest relativistic free field. Then, the commutator of quantized fields operators at spacetime points x and y vanishes for equal times x^0 and y^0 (for any spatial separation $\vec{x} - \vec{y}$) and spacelike values of $x - y$. Second, MC can be adapted and modified to include fields with spin or charge, like the Dirac field, via the spin-statistics theorem. It is even a sufficient condition to define a Lorentz-covariant scattering matrix since it is connected to important analyticity properties of the fields (Duncan, 2012; Weinberg, 1995). It is, then, no surprise that physicists follow Einstein's suit in putting MC at the heart of relativistic causation in QFT.

However, MC is not as straightforward as it seems. For it to be a QM-informed axiom, we must now bite the bullet of not knowing what "measurement" means in a quantum setting (Earman and Valente, 2014, p. 11). Whatever "measurement" means, we need observables and states to obtain expectation values. One way to obtain the theory's observables is through a quantization procedure of a classical field theory.[4] Then, we need a state $|\Psi\rangle$ such that $AB\,|\Psi\rangle = BA\,|\Psi\rangle$, where $A \in \mathcal{A}(\mathcal{O}_1)$ and $B \in \mathcal{A}(\mathcal{O}_2)$ for $\mathcal{O}_1$ and $\mathcal{O}_2$ are spacelike separated. Since $\mathcal{O}_1$ and $\mathcal{O}_2$ are subregions of the background spacetime $\mathcal{M}$, $\mathcal{A}(\mathcal{O}_1)$ and $\mathcal{A}(\mathcal{O}_2)$ are subalgebras of $\mathcal{A}(\mathcal{M})$. Given this property of **isotony** and since $\mathcal{O}_1$ and $\mathcal{O}_2$ are spacelike separated, $\mathcal{A}(\mathcal{M})$ should admit a state. However, it would be absurd that local operations like $A\,|\Psi\rangle$ or $B\,|\Psi\rangle$ would alter a physical state for all space and time (Ruetsche, 2011, p. 110).

Despite these problems, MC still seems to be a reasonable axiom for a field theory in implementing some "separability" of physical subsystems, which, additionally, has testable consequences in terms of the measurements we can perform in each subsystem. A reason to probe this further is that, following Einstein's quotation, the separability of physical subsystems and the possibility of communicating them are closely connected, so we need a better grasp of that connection before diagnosing MC as a causal axiom. I will address those topics separately.

---

[4]Since the fields are operator-valued distributions, we also need a sufficiently smooth test function to smear them before using them in calculations.

3.2.1 Separability/Independence conditions

This subsection aims to nail down some interpretive difficulties of diagnosing MC as an axiom in AQFT. Since MC strives to make the separability of quantum field theoretical systems compatible with the prohibition of superluminal signaling between them, this subsection sets the stage for my discussion of causation in the following subsection.

Consider $\mathcal{A}(\mathcal{O}_1) \vee \mathcal{A}(\mathcal{O}_2)$, the smallest algebra generated by $\mathcal{A}(\mathcal{O}_1)$ and $\mathcal{A}(\mathcal{O}_2)$, where $\mathcal{O}_1$ and $\mathcal{O}_2$ are spacelike separated. A state ω in AQFT is a linear, positive-definite functional over the operator algebra. For our purposes, $\omega(A)$ is the expectation value of the observable A belonging to some algebra $\mathcal{A}$. Now, consider a pair of states ω_j acting on $\mathcal{A}(\mathcal{O}_j)$ for $j = 1, 2$. Then, we should be able to find a state ω acting on $\mathcal{A}(\mathcal{O}_1) \vee \mathcal{A}(\mathcal{O}_2)$ such that its restriction to $\mathcal{A}(\mathcal{O}_j)$ is just ω_j. To give an illustration more familiar to readers coming from QM, if the algebras are matrix algebras, the composition $\vee$ can be equivalent to a tensor product, and the expectation value $\omega(A)$ can be calculated using the trace prescription for a density matrix ρ_ω. In that case, the "restriction" from the global state to the local states would be taking the partial trace over either one of the subsystems.

The most famous form[5] of independence in the spirit of my last paragraph is **statistical independence**, which is merely taking the expected values of ω_1 and ω_2 as probabilistically independent, that is, $\omega(C) = \omega_1(C)\omega_2(C)$ for some observable $C \in \mathcal{A}(\mathcal{O}_1) \vee \mathcal{A}(\mathcal{O}_2)$. There is no resulting statistical mixture of the individual states because they are localized in causally disjoint regions. This form of independence looks like a straightforward rendition of what we would want the compatibility of measurements carried out in spacelike separated regions to be since we are breaking potential correlations between the measurements in both regions. However, for statistical independence to hold, we need an additional assumption about the nature of the algebras called the **split property**, which in turn depends on the fields' energy densities having certain properties (Fewster, 2016) different from SC that, though reasonable, are not always fulfilled. That is, it can be shown that statistical independence can be derived from MC, but it needs additional assumptions. On their own, MC and statistical independence are logically independent (Earman and Valente, 2014, p. 11). Moreover, more general forms of independence that avoid superluminal signaling do not imply MC (Halvorson, 2007, pp. 753-755).

Still, MC does seem to implement some "separability" of physical subsystems, even if it is not because of their independent statistical predictions. One suggestion is to see MC more as a form of mereology of the structure of the algebras than a claim about locality (Ruetsche, 2011, pp. 112-113); the operation $\vee$ in $\mathcal{A}(\mathcal{O}_1) \vee \mathcal{A}(\mathcal{O}_2)$

---

[5]Earman and Valente (2014, §4-5) survey other ways to formulate independence conditions.

should incorporate and respect MC. However, consider the composition of physical subsystems of standard QM: take two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, and construct their tensor product, $\mathcal{H}_1 \otimes \mathcal{H}_2$. Then the algebras $\mathcal{B}(\mathcal{H}_1) \otimes \mathrm{Id}_2$ and $\mathrm{Id}_1 \otimes \mathcal{B}(\mathcal{H}_2)$ commute, and yet QM is not a relativistic theory. This example shows that the composition of bigger subsystems from commuting algebras is not a feature of relativistic theories (Halvorson, 2007, p. 752). So even if MC seems reasonable, the way in which it combines QM and SR brings to light multiple difficulties.

3.2.2 MC and causation

Aside from statistical independence, the **no-signaling theorem** is a result closely connected to MC, but that shifts our focus from separability to relativistic causation. In this subsection, I will finally show that the problems of separability and measurements in the two previous ones disrupt the possibility that MC fully characterizes relativistic causation in AQFT.

Of the many formulations of the no-signaling theorem, I will follow (Earman and Valente, 2014, §4.2) to avoid further technicalities. Consider an observable $A \in \mathcal{A}(\mathcal{O}_1)$ with a (countable) spectral decomposition $A = \sum_i a_i E_i^A$, where the E_is are projectors and the a_is A's eigenvalues. We can then define a map $T^A(\cdot) = \sum_i E_i^A \cdot E_i^A$ over $\mathcal{A}(\mathcal{O}_1) \vee \mathcal{A}(\mathcal{O}_2)$ that is explicitly related to an A-measurement. Then, for a state ω and an observable $B \in \mathcal{A}(\mathcal{O}_2)$, we have:

$$\omega(T^A(B)) = \omega(B)$$

This means that states acting on the algebras of spacelike separated regions are unaffected by measurements in the other region. The moral is clear: MC allows the statistical predictions of one of the subsystems to be preserved even after performing measurements in the other.

However, a few complications stop us from proclaiming victory over superluminal signaling using MC. First, the preservation of the statistics of the outcomes of a measurement may not require commutativity after all (Rédei and Valente, 2010), and the no-signaling theorem only works for a restricted class of operations (Ruetsche, 2011, p. 111). Second, even if we have no-signaling theorems, the statistical independence of subsystems can get bypassed by measurement protocols like the one from figure 1. For this setup, successive observations of the same state in $\mathcal{O}_1$, $\mathcal{O}_2$, and $\mathcal{O}_3$ exhibit correlations between the measurement carried out in $\mathcal{O}_1$ and that in $\mathcal{O}_3$ (Sorkin, 1993). This result seems to imply that there is some superluminal signaling in QFT. Although Sorkin's paradox just turned 30, it has been significantly overlooked by most people working on measurements in QFT until a
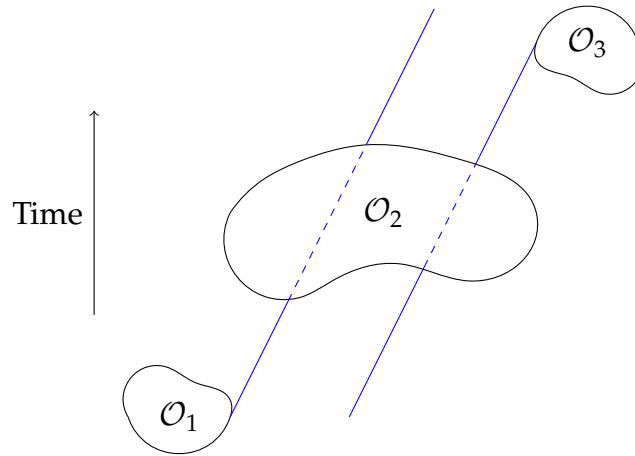
Figure 1: $\mathcal{O}_1$ and $\mathcal{O}_3$ are spacelike separated.

couple of years.[6] Some papers claim to have found a solution (Bostelmann et al., 2021), but whether it is satisfactory or not, it is more contrived than assuming MC.

In conclusion, although MC seems to shed some light on relativistic causation in AQFT, multiple difficulties impede us from claiming we have achieved a complete, clear characterization. First, as an axiom stated in terms of quantum measurements, we fall into the difficulties of understanding what "measurement" means and, especially, what it means for spacelike measurements to be compatible. We tried to implement MC as a form of separability for this purpose. Still, statistical independence needs more than MC to work, and some other forms of independence are logically independent of MC or satisfied in non-relativistic cases. Even having failed to characterize what form of separability MC implies, the no-signaling theorem made us hopeful that it would still prohibit superluminal signaling. However, the theorem only works for a particular case of measurements, it can be proved without MC, and it can be bypassed in setups like Sorkin's. The takeaway is that we still need a good characterization of relativistic causation in AQFT. Every time MC seemed to lead us in the right direction, interpretive difficulties or other pitfalls stopped us from clearly seeing the role of MC in prohibiting superluminal signaling.

3.3 Primitive Causality

From my initial characterization of relativistic causation, SC and MC have aimed to implement the role of locality, subluminality, and the metric structure of (Minkowski)

---

[6]For a recent exception, see (Papageorgiou and Fraser, 2023).

spacetime. However, I have not discussed determinism. Primitive causality (PC), or the time slice axiom, will fill this gap. First, we need some definitions. Consider two spacetime regions $\mathcal{O}_1$ and $\mathcal{O}_2$. $\mathcal{O}_2$ **depends causally** on $\mathcal{O}_1$ if every light ray in the backward (or forward) light cone originating from any point in $\mathcal{O}_2$ intersects $\mathcal{O}_1$. Then we have:
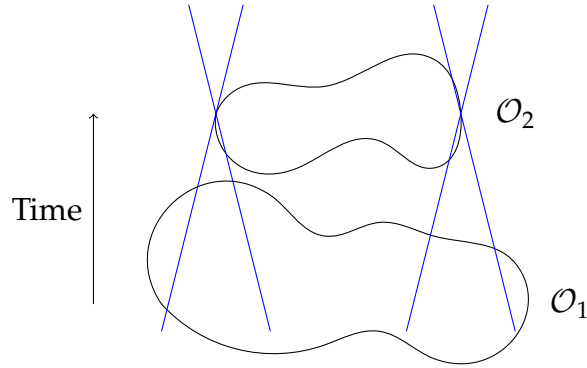

Figure 2: $\mathcal{O}_2$ depends causally on $\mathcal{O}_1$.

PC: If $\mathcal{O}_2$ depends causally on $\mathcal{O}_1$, $\mathcal{A}(\mathcal{O}_2) \subseteq \mathcal{A}(\mathcal{O}_1)$ (cf. figure 2).

Notice that if $\mathcal{O}_2 \subset \mathcal{O}_1$, $\mathcal{O}_2$ depends causally on $\mathcal{O}_1$, so the property of isotony can be derived from PC. Additionally, since $\mathcal{M}$ depends causally on any time slice and any time slice is isotonically included in $\mathcal{M}$, PC is usually presented in the literature in an alternative way:

> Consider a **time slice** in Minkowski spacetime.[7] That is, a region infinitely extended in space but restricted to a time interval of size τ:
> $\mathcal{O}_\tau := \{x \in \mathcal{M} : |x^0 - t| < \tau\}$. Then $\mathcal{A}(\mathcal{O}_\tau) = \mathcal{A}(\mathcal{M}) \; \forall \tau$.

This version of the axiom makes determinism à la Bohr even more salient since the standard interpretation of PC is that there should be a "dynamical law" that allows us to determine the values of the fields at any given time with the values of the fields at a time slice (Haag, 1996, p. 48). This claim can be re-stated in the following way: the time slice provides a region of evaluation of an initial value problem, and the "dynamical law" is deterministic (Bogolubov et al., 1987). Clearly, the exact nature of the law depends on the dynamics of the fields (Earman and Valente, 2014, p. 21).

---

[7]More generally, the spacetime manifold should be globally hyperbolic, and the time slice is some Cauchy surface.

Since deterministic laws have causal-like behavior, PC is the core of Earman and Valente's Atomistic view. This section will argue that PC's determinism is insufficient to characterize relativistic causation in AQFT.

As mentioned, PC is local, and its time slice version is global. Aside from this point, I will treat both versions of PC interchangeably. Dimock proved the relevance of PC for QFT since it holds for the Klein-Gordon field (1980) and the Dirac field (1982). However, as an initial value problem, it is trivially fulfilled. The Lagrangians of field theories are formulated in terms of the fields and their derivatives at a time slice, and if their derivatives are of a sufficiently higher order, we can recover the field values in all spacetime (Bogolubov et al., 1987, pp. 330-331). As such, PC is an entirely reasonable assumption for AQFT to make, but only because this form of "determinism" is a truism for QFT.[8] PC is not, then, a feature of relativistic causation *in AQFT* but in field theories in general.

Now, the QFTs considered by Dimock should satisfy PC since their equations of motion are hyperbolic partial differential equations (PDEs), which are characterized by having unique solutions within their domain of dependence (Geroch, 2011) and by having a finite propagation speed (Bär et al., 2007). The notion of **domain of dependence** has a precise definition within the realm of PDEs, but, in our case, it is merely the locus of causally dependent points to some region $\mathcal{O}$, denoted by $D(\mathcal{O})$. Another example of a hyperbolic PDE is the wave equation describing every undulatory phenomenon. From this point of view, PC is not a requirement from SR since it appears in non-relativistic phenomena, nor QM since it appears in classical ones. It is also not a sufficient characterization of relativistic causation since wave equations like the Klein-Gordon one would only be relativistic because their finite speed of propagation is the speed of light, but that does not rule out any higher speeds! These are the same reasons why PC is not a good causal axiom: the claim that the dynamics of the field are deterministic does not mean that the theory is causal in the sense that the dynamics should *also* respect the light cone structure of the underlying spacetime (Hofer-Szabó and Vecsernyés, 2018, p. 19).

Additionally, as Earman (2014) has claimed, determinism should apply not only to the time evolution of the field observables but also to the states. If states ω_1 and ω_2 give the same predictions with the observables in $\mathcal{A}(\mathcal{O})$, they should also provide the same ones with the observables in $\mathcal{A}(D(\mathcal{O}))$. Although I agree with Earman's suggestion, he bases it on taking $\mathcal{A}(\mathcal{O}) = \mathcal{A}(D(\mathcal{O}))$ as a result of Haag and Schroer (1962). However, this equation is not valid in general. We only obtain $\mathcal{A}(\mathcal{O}) \subseteq \mathcal{A}(D(\mathcal{O}))$ thanks to PC, and that is different from what Haag and Schroer

---

[8]Within AQFT, there are technical difficulties that complicate claiming that we can restrict the global dynamics to those of a time slice. So, if it is a truism in Lagrangian QFT, it is hardly obvious that it holds in AQFT. Thanks to Noel Swanson for this remark.

prove, which is a consequence of the local primitive causality that I will discuss in the next section.

3.4 Taking stock: against Atomistic

In the past subsections, we have seen the shortcomings and interpretive problems with the three causal axioms. Some of their drawbacks stem from not implementing QM, SR, and locality in a way amenable to codifying the desiderata of relativistic causation I laid out or from implementing them in ways that imported interpretive difficulties from each of those themes. And yet, someone who endorses ATOMISTIC might be undeterred: if none of these causal axioms do the job, let us look for a condition that puts all those advantages in one place! In the next section, I will argue that local primitive causality (LPC) does that for us, but I will show that it does so because it is implied by the conjunction of SC, MC, and PC, thus going against ATOMISTIC.

4 LPC as a defense of Holistic

To state LPC, we first need to define the **causal complement** of a region $\mathcal{O}$, $\mathcal{O}'$, as the set of points outside $D(\mathcal{O})$. Then we have:

  LPC: For a region $\mathcal{O} \subset \mathcal{M}$,[9] $\mathcal{A}(\mathcal{O}) = \mathcal{A}(\mathcal{O}'')$.

  I will use LPC to defend HOLISTIC in two ways. First, contrary to the causal axioms, I will show that it satisfies all the desiderata for relativistic causation. In summary, LPC is local since its dependence on spacetime regions is explicit; it is a crucial ingredient in showing that concrete experimental setups exhibit no superluminal signaling (Buchholz and Yngvason, 1994; Yngvason, 2005); it is the key assumption in theories of local observables that attempt to ensure that the determinism of PC is compatible with the temporal evolution of field values in the light cone of their regions (Hofer-Szabó and Vecsernyés, 2018); and it emphasizes the metric structure of $\mathcal{M}$ underlying the definition of "causal complement." I will only elaborate on how LPC rules out superluminal signaling and improves upon PC's attempt at capturing determinism.

  However, LPC's role of unifying the desiderata does not warrant promoting it to the ring to rule all the causal conditions of AQFT. The second defense of HOLISTIC

---

[9](Hofer-Szabó and Vecsernyés, 2018, p. 19) give a more general definition that I want to give here. See also their footnote 2.

will come as proof that the three causal axioms imply LPC, thus showing that LPC crystallizes how all the axioms hang together. Although the first version of this proof was first suggested by Haag and Schroer (1962) and I use their diamond construction, I fill in the missing steps and amend it in the appendix by not using Haag duality, which people involuntarily commit to when using LPC, but it is not satisfied by some of our best field theories and only works in particular spacetime regions (Haag, 1996, pp. 145-147), thus restricting greatly the kind of physical models we can consider.

4.1 LPC and determinism

The time slice version of PC implements determinism too crudely: if we want to know whether some event involving quantum fields causes another one, we specify the values of the fields (and probably some of its derivatives) at *a whole time slice*, and the pay-off is *every possible operation on spacetime*. As argued by Hofer-Szabó and Vecsernyés (2018), the axiom of covariance allows us to define a time-evolution for the fields (similarly to how we implement time translations unitarily in QM) $\alpha_{\tau',\tau}$ taking the fields from time slice τ to τ'. And yet, PC can't guarantee that $\alpha_{\tau',\tau}(\mathcal{A}(\mathcal{O}_\tau)) \subseteq \mathcal{A}(\mathcal{O}_{\tau'} \cap \mathcal{D}(\mathcal{O}_\tau))$! That merely symbolizes what I mentioned in the PC section: the dynamical law does not confine the time evolution of the fields to their light cones.

LPC grants us a more fine-grained notion of determinism, especially since it's formulated for finite regions. Since we assume that the underlying spacetime is globally hyperbolic, there will be a Cauchy surface transversing every open, bounded region $\mathcal{O}$. What LPC tells us, then, is that specifying the values of the fields, not in a whole Cauchy surface but in the chunk intersecting $\mathcal{O}$ suffices to know that operations in $\mathcal{A}(\mathcal{O})$ can influence operations in $\mathcal{A}(\mathcal{O}'')$. Figure 4 below will help illuminate this point: since the causal future of $\mathcal{C}$ intersects the causal future of points in $\mathcal{C}'$, claiming that $\mathcal{A}(\mathcal{C}) = \mathcal{A}(\mathcal{C}'')$ tells us how much can we enlarge the operations we do in $\mathcal{C}$ while still being insulated from possible influences from operations in $\mathcal{C}'$. So, fulfilling Hofer-Szabó and Vecsernyés's aspirations of a condition that is *both* deterministic and restricted to the light cones of finite regions (as opposed to infinite time slices), LPC is the cornerstone of a local, causal dynamical evolution. In the limit where $\mathcal{C}$ extends infinitely spatially, a full Cauchy surface transverses it, and the diamond $\mathcal{C}''$ becomes the full spacetime, as expected from PC!

4.2 LPC and superluminal signaling: The Fermi's Two-Atom System

In this subsection, I will introduce (but skip most of the details) an experimental setup devised by Fermi (1932) to check whether the quantum theory of radiation that

he was studying was compatible with SR and translated in (Yngvason, 2005) to an AQFT language. Consider two atoms, a and b, separated by a distance R. At time $t = 0$, a is in its ground state, and b is in an excited state. When b emits radiation, there is a non-zero probability that a is excited for $t > 0$. Since the effect of b's decay cannot be superluminal, a should be in its ground state for $t < R/c$, where c is the speed of light. (See figure 3.)
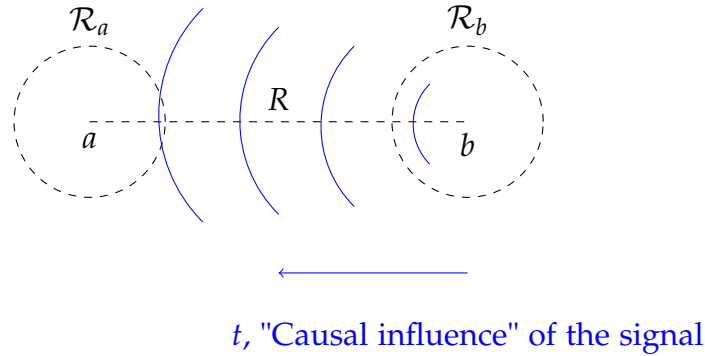
t, "Causal influence" of the signal

Figure 3: Fermi's two-atom system.

Yngvason improved upon Fermi's result of the transition probability from a's ground to a's excited state using the observables available in QFT but not in QM. More specifically, since QFT does not have "minimal projectors," we can't make a measurement that merely acts like a switch from "unexcited" to "excited."[10] However, we can still know of changes in the state of a by purely local operations.[11]

---

[10]Technical aside: Let $\mathcal{A}$ be a von Neumann algebra. Two orthogonal projectors $P_1, P_2 \in \mathcal{A}$ are called equivalent with respect to $\mathcal{A}$, denoted $P_1 \sim P_2$, if there exists an operator $V \in \mathcal{A}$ such that $P_1 = V^*V$ and $P_2 = VV^*$. Projector equivalence is an equivalence relation on the projector lattice on the Hilbert space $\mathcal{H}$ on which the von Neumann algebra is represented. Denote $P_1 \leq P_2$ when $P_1 \sim P_1'$ and $P_1'\mathcal{H} \subseteq P_2\mathcal{H}$. If $\mathcal{A}$ is a von Neumann factor, $\leq$ is a total order (Haag, 1996, Theorem 2.1.9). A projector is said to be finite if it is not equivalent to any proper subprojector. The notion of finiteness, together with the total ordering, allows us to define a projector's relative dimension d, a positive number (possibly infinite) where $d(P) = 0$ if and only if $P = 0$, $d(P_1) \leq d(P_2)$ if and only if $P_1 \leq P_2$, $d(P_1 + P_2) = d(P_1) + d(P_2)$ if $P_1 \perp P_2$ etc. (See (Moretti, 2017, § 7.6) for details). Then, a minimal projector P has no non-zero subprojectors, so it is used to normalize $d(P) = 1$. Equivalently, minimal projectors are the atoms of the projector lattice. That is the precise sense in which they are "switch-like" in footnote 11. Only type I factors contain minimal projectors.

[11] Technical aside: Yngvason's "strong local preparability" only needs that the algebras are type III von Neumann factors, but we also want to prepare the states with arbitrary precision for any measurement to be as "switch-like" as possible. There are at least two strategies for doing this. One is Yngvason's, which uses the subclassification of type III von Neumann factors: type III_1 factors help specify further *how* states can be prepared (Connes and Størmer, 1978, Theorem 4). Another one is

The crux of the setup is to define two states, one in which atom b is excited in $\mathcal{R}_b$ at $t = 0$ and a state in which atom b is in its ground state (or absent, since a only "knows" that some radiation came into $\mathcal{R}_a$, not the state of b) in $\mathcal{R}_b$ at $t = 0$. Since these are joint states of the atoms, doing a measurement outside of $\mathcal{R}_b$ at $t = 0$ cannot distinguish them from the ground state of a; one can only check whether b is in its ground state or excited state at $t = 0$ *from inside* $\mathcal{R}_b$.

Then, to check whether a became excited by the signal emitted by b *measured in* $\mathcal{R}_a$ is to consider how much these states deviate from each other at some other time t with observables in $\mathcal{O}_a$ (where $\mathcal{O}_a$ is the spacetime region made up of the spatial region $\mathcal{R}_a$ and an infinitesimal interval around $t = 0$). Since $\mathcal{O}_a''$ includes the region that the light rays have not yet reached, the two states are indistinguishable for $t < R/c$ if LPC holds. As soon as b de-excites itself, it is irrelevant what happens to it from a's point of view; it would not be able to detect any effect before the arrival of the light rays stemming from b's de-excitation anyway. So LPC guarantees that a doesn't know of b's presence before it detects the radiation it emitted!

Notice that $\mathcal{A}(\mathcal{O}_a) = \mathcal{A}(\mathcal{O}_a'')$ implies that we have "enlarged" the region $\mathcal{O}_a$ to $\mathcal{O}_a''$. Since the light rays emitted from b are moving, the region causally dependent on $\mathcal{O}_b$ is growing. Only when the rays reach $\mathcal{R}_a$ will an observer be able to measure whether the state of a changed. The light rays from PC's notion of causal dependence need not originate from points at rest at the reference frame at which their light cones are drawn, so here we see how a region flips in time from causally independent to causally dependent. So, aside from showing how LPC rules out superluminal signaling, Fermi's experiment helps to see how LPC finesses the notion of causal dependence, as argued in the previous subsection.

4.3 Further advantages of LPC

There's more. First, LPC can be generalized to curved spacetimes and used in concrete measurement protocols (Fewster and Verch, 2020), thus alleviating some of

---

to use the non-local type I algebra interpolating local type III algebras satisfying the split property to approximate local operations (Kitajima, 2018, Corollary 1 of Theorem 4). Although the operational significance of this strategy has been questioned in the literature (Ruetsche and Earman, 2011, §5.2.5), it's not clear to me that it's totally unviable. Even if we need to be careful about using type I algebras to study relativistic causation (Hegerfeldt, 1994), results like Kitajima's or those in (Busch, 1999) show that rejecting them altogether would be too rash, and even Yngvason made a mistake in the argument ruling out using type I factors. Probably inspired by MC, he associates the orthogonality of a's ground and excited states with the fact that their respective orthogonal projectors will localize them in causally independent regions. But, if A is a self-adjoint operator of a von Neumann algebra $\mathcal{A}$ (independently of its type), all of A's spectral projections belong to $\mathcal{A}$ (Murphy, 1990, Theorem 4.1.11), including orthogonal projections. Thanks to Aleksandr Pinzul for bringing this theorem to my attention.

the worries I had raised about SC and MC.[12] Second, in keeping this paper minimally technical, I have avoided any talk on the details of the mathematical structure of observable algebras because my arguments are independent of these considerations. However, these details are vital for relativistic causation. For example, superluminal signaling is unavoidable if the algebras are the matrix algebras of QM, which do have minimal projectors (Hegerfeldt, 1994). From this point of view, Buchholz and Yngvason (1994; 2005) have shown how LPC and the specific type of von Neumann algebra relevant for AQFT come together in implementing relativistic causation. (Cf. footnote 11 for further discussion).

4.4   Holistic redux: LPC is implied by the causal axioms

The past subsections could have been the ingredients for a defense of ATOMISTIC: LPC should supersede the three causal axioms as the most direct expression of relativistic causation in AQFT since it captures all the desiderata. However, this attempt at privileging one condition or even having it replace the rest would be too rash for two reasons. First, there are results important for the foundations of QFT, like the CPT theorem, that can be derived from the three causal axioms but that are logically independent of LPC (Haag, 1996, §5.1). So even if LPC has some additional advantages compared to the three causal axioms *from the point of view of the causal structure of the theory*, the axioms have more physical content than that which goes into codifying relativistic causation. Plus, I have been discussing in this paper which axioms do the causal heavy lifting for the theory, not which ones are necessary to derive the core theorems of QFT or to have a consistent theory.

   However, more importantly for my purposes, we can derive LPC from the three causal axioms (and some other physically reasonable conditions), as I will sketch below as a defense of HOLISTIC. Consider the construction of (Haag and Schroer, 1962) from figure 4. It is now possible to prove that $\mathcal{A}(\mathcal{D}) = \mathcal{A}(\mathcal{C})$, which is exactly

---

[12]However, I am not claiming that LPC solves problems like Sorkin's paradox since that would require a different conception of measurements in QFT from the one I have been assuming in this paper. The advantages I have presented from Buchholz, Yngvason, and Fewster and Verch rely on specific measurement protocols or experimental setups. Fewster and Verch do not assume LPC explicitly, but they obtain it for free (and with it, its advantages) because their framework relies heavily on *causally convex* (i.e., diamond-like) regions, for which LPC is trivially true. Taking all regions to be diamonds is not esoteric; Haag uses it to formulate a lattice of causally complete regions that makes the connection between causal complements and algebraic commutants salient (1996, §III.4.1) and to study superselection sectors (1996, p. 156). However, Fewster and Verch's version of LPC is more general than the one I presented here because they work with local *-algebras, not von Neumann algebras, which gives them less topological structure than the one the derivation of LPC from the other axioms that I cite in the main text can exploit.
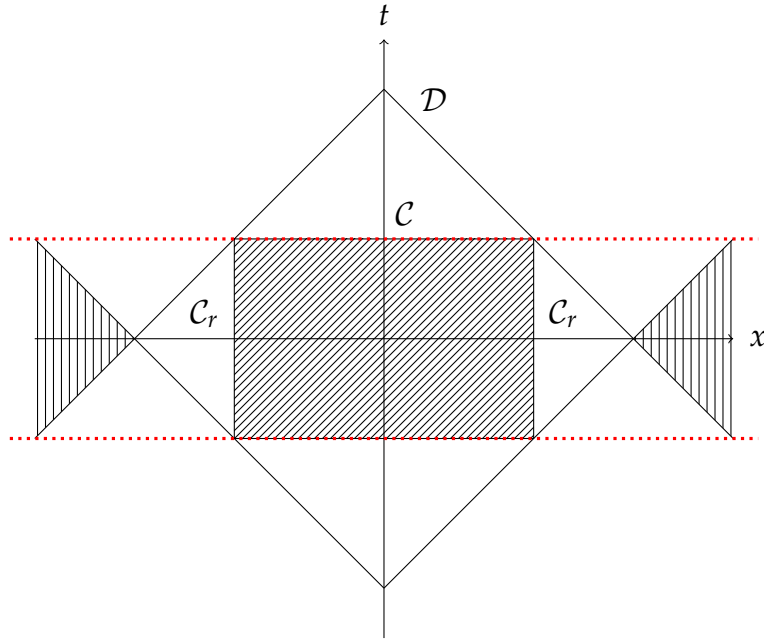
Figure 4: The region shaded with slanted lines is $\mathcal{C}$. The region shaded with vertical lines is $\mathcal{C}'$, the causal complement of $\mathcal{C}$. The diamond $\mathcal{D}$, containing $\mathcal{C}$, is the causal complement of $\mathcal{C}'$. The red dotted lines extend spatially to form a time slice of the temporal size of $\mathcal{C}$. Call $\mathcal{C}_r$ the two caps outside of $\mathcal{C}$ that extend in the spatial direction.

LPC since the diamond $\mathcal{D}$ satisfies $\mathcal{D} = \mathcal{C}''$. All the pieces of the proof have been put together in the appendix, but here, I will only sketch the main steps. From SC, one can prove that the domain of analyticity of the fields in $\mathcal{C}$ can be extended to $\mathcal{C} \cup \mathcal{C}_r$ (Borchers, 1961; Streater and Wightman, 2000). From PC and MC, we can build a time slice from $\mathcal{C} \cup \mathcal{C}_r$ and $\mathcal{C}'$, where $\mathcal{C}$ and $\mathcal{C}'$ localize separated subsystems. The fact that the observables at a time slice generate those in all spacetime gives some nice closure properties that display the intimate connection between the local algebras and the regions in which they are supported. Finally, from MC, the observables from $\mathcal{C}'$ commute with those of $\mathcal{C}$ but also with those of $\mathcal{C}''(= \mathcal{D})$. The takeaway from this construction is that even if SC, MC, and PC individually failed to characterize relativistic causation in AQFT, LPC captures all the desiderata of relativistic causation in AQFT *because* it is a fruit of their conjunction.

5 Outlook

Throughout this paper, we saw that SC, MC, and PC only give partial characterizations of relativistic causation. Even if this is unsatisfactory for the advocates of Atomistic, the need for multiple axioms should not be startling anymore precisely because they highlight different aspects of the causal framework of the theory. Additionally, LPC encapsulates each axiom's advantages for the most widely used regions and theories in AQFT, which was the key idea behind my defense of Holistic. Assuming LPC has become a widespread move in more technical literature in AQFT. Still, its motivations are rarely stated, its role as a causal axiom is left uninterpreted, and its interdependence with the other causal axioms is ignored.

However, many additional worries emerged in analyzing this wide variety of axioms. First, an appropriate interpretation of MC would require tackling the measurement problem in QFT. One way to address this problem is to introduce a more sophisticated account of local operations in AQFT, which is where some of the more technical literature has been heading (Okamura and Ozawa, 2016; Kitajima, 2018; Drago and Moretti, 2020). However, from an interpretational point of view, it is unclear whether considering a more widespread class of measurements would solve the problem of answering what "measurement" means. Even if looking for an account of measurements in QFT seems orthogonal to a project on relativistic causation, I do think they are connected through the worries that philosophers have raised about the "operationalist" views of the founders of AQFT and the "algebraic imperialist" (Ruetsche, 2011) tendencies of the mathematical physicists working on this formulation of QFT. Additionally, revising the connections between causation and measurements would require further examining the local algebras' mathematical structure and the states that can act on them. The need to appeal to those tools can be seen not only in the problems of measurements in, e.g., MC but in the interplay between the different formulations of PC as a requirement demanding that we have a compatible notion of local dynamics and global determinism. That is one of the problems that even promoting LPC to an additional assumption in AQFT leaves unsettled.

The second problem is that any account of relativistic causation in AQFT is challenged by entanglement, which is ubiquitous and maximal in AQFT, even in spacelike separated regions (Summers and Werner, 1988). Entanglement *itself* is not something we should worry about since it is just a key feature of QM and QFT. Instead, the origin of the uneasiness about entanglement is that it is often conceived as a form of non-locality or non-separability, which conflicts with, e.g., conceiving MC as a form of independence between spacelike separated subsystems. The pull of

entanglement as a problem for causation in AQFT relies heavily on assumptions outside of the theory's axiomatic framework about measurements and operations in QFT (Ruetsche, 2021) and, as such, it is connected to my first remaining worry. These problems are, to the best of my knowledge, unsettled. I hope this diagnostic points to some of the issues that need to be addressed and the importance of doing so.

Acknowledgements

Appendix: Causal diamond proof using essential duality

Preliminaries: notation, some definitions, and re-stating the axioms

Let $\mathcal{O}$ be an open, bounded subregion of Minkowski spacetime $\mathcal{M}$, and $\mathcal{A}(\mathcal{O})$ a von Neumann algebra whose elements are taken to be the observables describing the physics in $\mathcal{O}$. I'll take the algebras to be "concretely represented," which means that $\mathcal{A}(\mathcal{O}) \subseteq \mathcal{B}(\mathcal{H})$, the bounded, linear operators on some Hilbert space $\mathcal{H}$. I'll also assume that $\mathcal{A}(\mathcal{M}) = \mathcal{B}(\mathcal{H})$. Denote by $\mathcal{A}(\mathcal{O})'$ the **commutant** of $\mathcal{A}(\mathcal{O})$—the set of all operators commuting with those of $\mathcal{A}(\mathcal{O})$. A **von Neumann factor** is a von Neumann algebra satisfying $\mathcal{A}(\mathcal{O}) \cup \mathcal{A}(\mathcal{O})' = \mathcal{B}(\mathcal{H})$ or, equivalently, $\mathcal{A}(\mathcal{O})' \cap \mathcal{A}(\mathcal{O}) = \mathbb{C} \operatorname{Id}_{\mathcal{B}(\mathcal{H})}$.

Consider two spacetime regions $\mathcal{O}_1$ and $\mathcal{O}_2$. $\mathcal{O}_2$ **depends causally** on $\mathcal{O}_1$ if every light ray in the backward (or forward) light cone originating from any point in $\mathcal{O}_2$ intersects $\mathcal{O}_1$. If $\mathcal{O}_1 \subseteq \mathcal{O}_2$, **isotony** tells us that $\mathcal{A}(\mathcal{O}_1) \subseteq \mathcal{A}(\mathcal{O}_2)$. Let $D(\mathcal{O})$ denote the **domain of dependence** of $\mathcal{O}$, the locus of causally dependent points to $\mathcal{O}$, and let $\mathcal{O}'$, the **causal complement** of $\mathcal{O}$, be the set of points outside $D(\mathcal{O})$. **Essential duality** (Halvorson, 2007, p. 841) then tells us that $\mathcal{A}(\mathcal{O}')' = \mathcal{A}(\mathcal{O}'')$.

The following **causal axioms** follow the notation and terminology from the main text.

- **Spectrum Condition** (SC): (a) The joint spectrum of the infinitesimal generators of translations in Minkowski spacetime is confined to the forward lightcone. (b) There is a unique (up to phase factors) **vacuum state**, which is translationally invariant and has zero energy-momentum.

- **Microcausality** (MC): $\mathcal{A}(\mathcal{O}_1)$ commutes with $\mathcal{A}(\mathcal{O}_2)$ if $\mathcal{O}_1$ and $\mathcal{O}_2$ are space-like separated. In symbols, $\mathcal{A}(\mathcal{O}) \subseteq \mathcal{A}(\mathcal{O}')'$.

- **Primitive Causality** (PC): Consider a **time slice** in $\mathcal{M}$. That is, a region infinitely extended in space but restricted to a time interval of size τ: $\mathcal{O}_\tau := \{x \in \mathcal{M} : |x^0 - t| < \tau\}$. Then $\mathcal{A}(\mathcal{O}_\tau) = \mathcal{A}(\mathcal{M}) \; \forall \tau$.

Proof

Consider the following cylinder $\mathcal{C} := \{x \in \mathcal{M} : ||\vec{x}|| < a, |x^0| < \tau\}$ for some a and τ, shaded with diagonal lines in figure 5. Its causal complement $\mathcal{C}'$ is shaded with vertical lines. Call the caps between the cylinder and the vertices of the diamond $\mathcal{C}_t$ and $\mathcal{C}_r$, according to their orientation toward the time or space axis. Let $\mathcal{D}$ be the diamond $\mathcal{C} \cup \mathcal{C}_t \cup \mathcal{C}_r$, which is equivalent to $\mathcal{C}''$. This proof aims to show that $\mathcal{A}(\mathcal{C}) = \mathcal{A}(\mathcal{D})$. Although this result is only suited for regions such as the diamond, $\mathcal{A}(\mathcal{O}) = \mathcal{A}(\mathcal{O}'')$ is usually lifted to a general property. See the main text for further discussion.

Here's a sketch of the proof, where every step will have its own subproof.

1. From SC, $\mathcal{A}(\mathcal{C} \cup \mathcal{C}_r) = \mathcal{A}(\mathcal{C})$.

2. From PC and 1, $(\mathcal{A}(\mathcal{C} \cup \mathcal{C}_r) \cup \mathcal{A}(\mathcal{C}'))'' = (\mathcal{A}(\mathcal{C}) \cup \mathcal{A}(\mathcal{C}'))'' = \mathcal{B}(\mathcal{H})$.

3. From MC and 2, $\mathcal{A}(\mathcal{C})$ is a von Neumann factor.

$\therefore$ From MC and essential duality, $\mathcal{A}(\mathcal{C}) = \mathcal{A}(\mathcal{D})$.

Proof of 1

The proof of this step is the most elaborate one, and it relies crucially on a lemma found in (Borchers, 1961). However, Borchers doesn't clarify what the role of SC is, so, in (Calderón Ossa, 2019, pp. 22-31), where much of this appendix comes from, I supplemented Borchers' proof with results found in (Streater and Wightman, 2000). I've also added more details to some of the proofs I adapted from them.
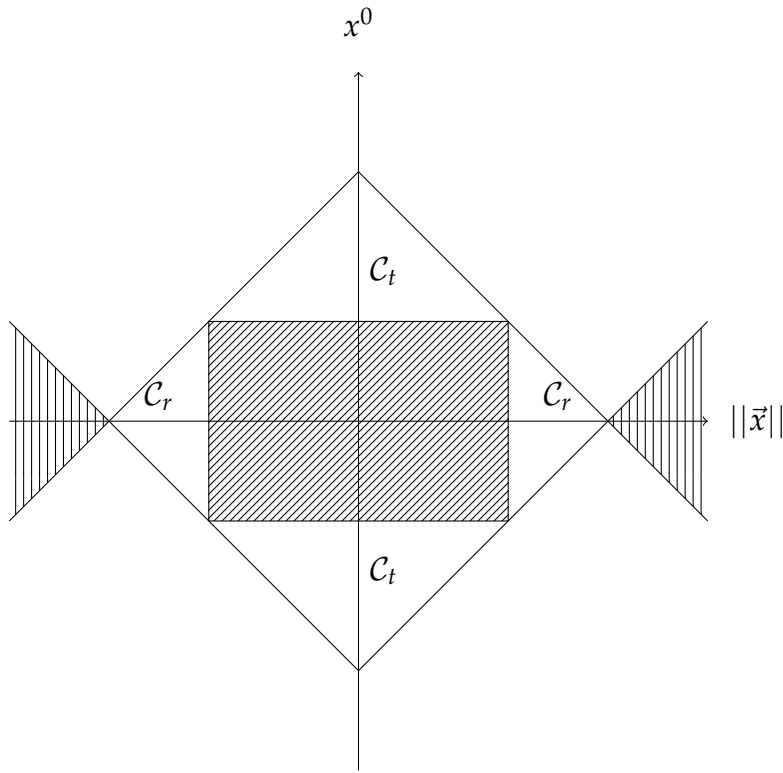
Figure 5: Causal diamond.

Streater and Wightman's Theorem 2-13: Let F_1 be a function holomorphic in an open set D_1 in the upper half-plane with an open interval $a < x < b$ as part of its boundary. Let F_2 be holomorphic in an open set D_2 in the lower half-plane with $a < x < b$ as part of its boundary. Suppose $F_1(x) = \lim_{y \to 0^+} F_1(x + iy)$ and $F_2(x) = \lim_{y \to 0^+} F_2(x - iy)$ exist uniformly, in $a < x < b$, are continuous and satisfy $F_1(x) = F_2(x)$ for $a < x < b$. Then, F_1 and F_2 are holomorphic on $a < x < b$ and are the same holomorphic function.

The theorem, a special case of Bogolyubov's Edge of the Wedge theorem (Rudin, 1971; Bogolubov et al., 1987; Vladimirov et al., 1994), states the condition under which the analytic continuation of two holomorphic functions is such that they coincide in their real boundaries. We'll use this result below and, as Streater and Wightman discuss, this result can be extended for distributions whose real part is compactly supported.

Let V_+ be the set of all real momentum four-vectors p such that $p^2 > 0$ and $p^0 > 0$

on some reference frame, and let $\bar{V}_+$ be its closure (the set of ps such that $p^2 \geq 0$ and $p^0 \geq 0$). Define V_- and $\bar{V}_-$ analogously but for negative energies.

> **Borchers' lemma:** Let $F^\pm(x + iy)$ be analytic functions with boundary values $F^\pm(x) = \lim_{\substack{y \to 0 \\ y \in \bar{V}_\pm}} F^\pm(x + iy)$ that exist as tempered distributions. Then, if $F^+(x) = F^-(x)$ for $x \in \mathcal{C}$, $F^+(x) = F^-(x)$ for $x \in \mathcal{C} \cup \mathcal{C}_r$.

Sketch of the proof. From the special case of the Edge of the Wedge theorem from above, $F^+(\zeta)$ and $F^-(\zeta)$ coincide in their domain of analyticity, which contains the cylinder $\mathcal{C}$. Then, a re-parametrization of ζ within allows us to extend the domain of analyticity of $F^\pm$, and we can prove that the new ζ lies in $\mathcal{C} \cup \mathcal{C}_r$.

Before proving Streater and Wightman's Theorem 3-2(b), the last ingredient we're missing to prove my premise 1, we need an intermediate result that uses SC.

> **Intermediate result (Streater and Wightman, 2000, pp. 91-92):** Let $U(a, \Lambda)$ be a unitary representation of the Poincaré group on a Hilbert space $\mathcal{H}$, where a is a four-vector and Λ is a Lorentz transformation. Let $\Phi, \Psi \in \mathcal{H}$ and $\langle | \rangle$ be $\mathcal{H}$'s inner product. Then, $\int \mathrm{d}a\, e^{-ip \cdot a} \langle \Phi | U(a, \mathrm{Id}) | \Psi \rangle = 0$ unless p belongs to the spectrum of the energy-momentum operators.

Proof. Let us write $\langle \Phi | \Psi \rangle$ in terms of a basis[13] of momentum eigenstates $|Q_\alpha\rangle$: $\langle \Phi | \Psi \rangle = \sum_\alpha \int \mathrm{d}Q\, \langle \Phi | Q_\alpha \rangle \langle Q_\alpha | \Psi \rangle$. Then, we have that

$$
\int \mathrm{d}a\, e^{-ip \cdot a} \langle \Phi | U(a, \mathrm{Id}) | \Psi \rangle = \sum_\alpha \int \mathrm{d}Q \int \mathrm{d}a\, e^{-ip \cdot a} \langle \Phi | Q_\alpha \rangle \langle Q_\alpha | U(a, \mathrm{Id}) | \Psi \rangle
$$

$$
= \sum_\alpha \int \mathrm{d}Q \int \mathrm{d}a\, e^{-i(p - Q) \cdot a} \langle \Phi | Q_\alpha \rangle \langle Q_\alpha | \Psi \rangle
$$

$$
= (2\pi)^4 \sum_\alpha \int \mathrm{d}Q\, \delta(Q - p) \langle \Phi | Q_\alpha \rangle \langle Q_\alpha | \Psi \rangle
$$

The second equality follows from the fact that $U(a, \mathrm{Id}) = e^{i\hat{P}^\mu a_\mu}$, where $\hat{P}$ is the energy-momentum operator, and that $|Q_\alpha\rangle$ is one of its eigenvectors. Then, since Φ and Ψ are arbitrary states, $\int \mathrm{d}a\, e^{-ip \cdot a} \langle \Phi | U(a, \mathrm{Id}) | \Psi \rangle = 0$ unless p is equal to some $Q \in \bar{V}_+$. $\qquad\square$

---

[13]Streater and Wightman recognize that this derivation might not be rigorous (p. 91), but it can be made rigorous with different techniques (p. 92). I'll stick to this version since it's more intuitive and to me and the role of SC is clearer.

Streater and Wightman's Theorem 3-2(b): Let $\varphi_1, \varphi_2, ... \varphi_n$ be any components of any fields and recall the definition of a **Wightman function**: $W^{(n)}(x_1, ..., x_n) := \langle 0 | \varphi_1(x_1)...\varphi_n(x_n) | 0 \rangle$. Then there are tempered distributions $\mathcal{W}(\xi_1, ..., \xi_{n-1})$ depending on the coordinates $\xi_j = x_j - x_{j+1}$ for $j = 1, 2, ..., n-1$ that satisfy

$$W^{(n)}(x_1, ..., x_n) = \mathcal{W}(\xi_1, ..., \xi_{n-1}).$$

Their Fourier transforms are tempered distributions defined by

$$\tilde{W}^{(n)}(p_1, ..., p_n) = \int dx_1...dx_n \; e^{i \sum_{j=1}^{n} p_j x_j} W^{(n)}(x_1, ..., x_n)$$

$$\tilde{\mathcal{W}}(q_1, ..., q_{n-1}) = \int d\xi_1...d\xi_{n-1} \; e^{i \sum_{j=1}^{n-1} q_j \xi_j} \mathcal{W}(\xi_1, ..., \xi_{n-1})$$

and are related by

$$\tilde{W}^{(n)}(p_1, ..., p_n) = (2\pi)^4 \delta \left(\sum_{j=1}^{n} p_j \right) \tilde{\mathcal{W}}(p_1, p_1 + p_2, ..., p_1 + p_2 + ... + p_{n-1}).$$

Further, we have

$$\tilde{\mathcal{W}}(q_1, ..., q_{n-1}) = 0$$

if any q is not in the energy-momentum spectrum of the states.

Proof. The existence of $\mathcal{W}$ follows automatically from the fact that the Wightman functions are themselves tempered distributions (Wightman, 1996) and that they're translationally invariant (Streater and Wightman, 2000, pp. 38-40). From their translational invariance, their Fourier transforms, straightforwardly generalized from functions to distributions, are well-defined.

As for the relationship between their Fourier transforms, it follows from the Fourier transform of the Dirac delta we already used in the intermediate result above and from the following trick:
$\sum_{j=1}^{n} p_j x_j = p_1(x_1 - x_2) + (p_1 + p_2)(x_2 - x_3) + ... + (p_1 + ... + p_{n-1})(x_{n-1} - x_n) + x_n \sum_{j=1}^{n} p_j = p_1 \xi_1 + (p_1 + p_2)\xi_2 + ... + (p_1 + ... + p_{n-1})\xi_{n-1} + x_n \sum_{j=1}^{n} p_j$.

As for the last result, let us use the intermediate result from above, with $\Phi \to |0\rangle$ and $U(a, \text{Id})\Psi \to \varphi_1(x_1)...\varphi_j(x_j)U(-a, \text{Id})\varphi_{j+1}(x_{j+1})\varphi_n(x_n) |0\rangle = \varphi_1(x_1)...\varphi_j(x_j + a)\varphi_{j+1}(x_{j+1})...\varphi_n(x_n) |0\rangle$. That way, the inner product of Φ and Ψ is directly related to the Wightman functions. The "new" Ψ is also a unitary operator acting on a state, but we "moved" the operator using its unitarity and the

translational invariance of the Wightman functions to obtain the following general result for $p \notin \bar{V}_+$

$$\int \mathrm{d}a \; e^{ip \cdot a} \mathcal{W}(\xi_1, ..., \xi_j + a, ..., \xi_{n-1}) = 0 \; , \; j = 1, 2, ..., n.$$

Since this result holds for any $j = 1, 2, ..., n$ and a is an arbitrary translation, $\tilde{\mathcal{W}}(q_1, ..., q_{n-1}) = 0$ unless *each* q_j lies in the physical spectrum. Using the relationship between the original Wightman functions and the translated ones, their Fourier transform vanishes unless $p_1, p_1 + p_2, ..., p_1 + p_2 + ... + p_n \in \bar{V}_+$. □

We can now go back to Borchers' paper (1961, pp. 791-792) and prove my premise 1.

Rest of the proof. Let C be a bounded operator and let P and Q be eigenstates of the energy-momentum operator such that $[\varphi(x), C] = 0$ for $\varphi(x)$ a scalar field and $x \in \mathcal{C}$. Consider the tempered distributions $F^+(x) = \langle P | \varphi(x) C | Q \rangle$ and $F^-(x) = \langle P | C\varphi(x) | Q \rangle$. (The distributions will only be tempered if $|P_0|, |Q_0| < \infty$, but this is expected from SC). From Streater and Wightman's theorem 3-2(b), their Fourier transforms satisfy $F^+(k) \neq 0$ only for $k + P \in \bar{V}_+$ and $F^-(k) \neq 0$ only for $k - Q \in \bar{V}_-$. Since $\varphi(x)$ and C commute in $\mathcal{C}$, $F^+(x) = F^-(x)$ for $x \in \mathcal{C}$, but this implies $F^+(x) = F^-(x)$ for $x \in \mathcal{C} \cup \mathcal{C}_r$ due to Borchers' lemma from above. It follows that $[\varphi(x), C] = 0$ for $x \in \mathcal{C} \cup \mathcal{C}_r$. Since the fields are the generators of the local algebras $\mathcal{A}(\mathcal{O})$ and since any bounded operator that commutes with fields in $\mathcal{C}$ also commutes with any field in $\mathcal{C} \cup \mathcal{C}_r$, $\mathcal{A}(\mathcal{C})' \subseteq \mathcal{A}(\mathcal{C} \cup \mathcal{C}_r)'$. Now, since $\mathcal{A}_1 \subseteq \mathcal{A}_2$ implies $\mathcal{A}_2' \subseteq \mathcal{A}_1'$ for any two von Neumann algebras (Haag, 1996, p. 114), we have that $\mathcal{A}(\mathcal{C})'' \supseteq \mathcal{A}(\mathcal{C} \cup \mathcal{C}_r)''$. Since these are von Neumann algebras, $\mathcal{A}'' = \mathcal{A}$, so $\mathcal{A}(\mathcal{C}) \supseteq \mathcal{A}(\mathcal{C} \cup \mathcal{C}_r)$.[14] Due to the isotony property we have that $\mathcal{A}(\mathcal{C}) \subseteq \mathcal{A}(\mathcal{C} \cup \mathcal{C}_r)$. Combining these results, we finally obtain $\mathcal{A}(\mathcal{C} \cup \mathcal{C}_r) = \mathcal{A}(\mathcal{C})$, which is precisely premise 1.

Proof of 2

The proof of this step comes directly from (Haag and Schroer, 1962). Let us take a time slice of the size of the cylinder, as shown in figure 6.

From PC, the algebra of the time slice generates $\mathcal{A}(\mathcal{M})$, which contains all the bounded operators. Now, notice that the time slice is generated from $\mathcal{C}$, $\mathcal{C}_r$, a subset of $\mathcal{C}'$, and the caps between $\mathcal{C}'$ and $\mathcal{C} \cup \mathcal{C}_r$ shaded in the diagram above. Those caps, following (Haag and Schroer, 1962), can be neglected if we take the smallest von

---

[14]In the following pages, I will use the properties of von Neumann algebras from this sentence and the previous one without warning.
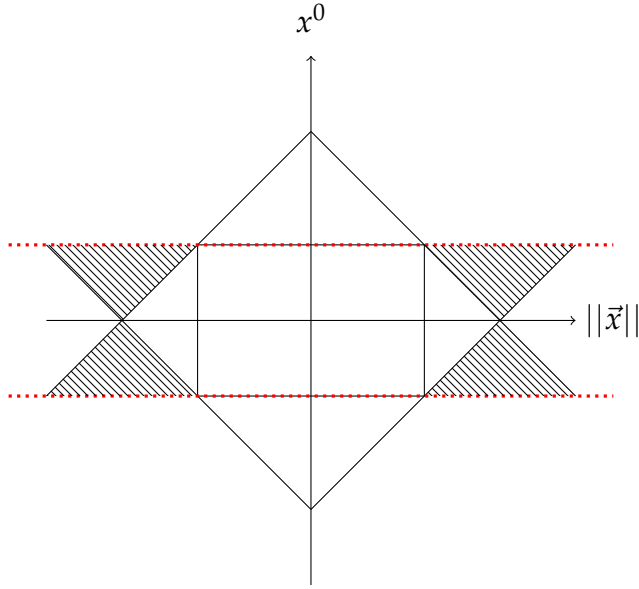
Figure 6: Time slice of the size of the cylinder delimited by the red dotted lines.

Neumann algebra generated by $\mathcal{A}(\mathcal{C} \cup \mathcal{C}_r)$ and $\mathcal{A}(\mathcal{C}')$, $(\mathcal{A}(\mathcal{C} \cup \mathcal{C}_r) \cup \mathcal{A}(\mathcal{C}'))''$. We take the double commutant of the union (or, equivalently, their weak closure, following von Neumann's bicommutant theorem) since the union of two von Neumann algebras is not necessarily a von Neumann algebra. Since this is the smallest von Neumann algebra containing them, this choice should minimize the effect of ignoring the caps. Then, $(\mathcal{A}(\mathcal{C} \cup \mathcal{C}_r) \cup \mathcal{A}(\mathcal{C}'))'' = \mathcal{B}(\mathcal{H})$, which is precisely premise 2. From premise 1, $\mathcal{A}(\mathcal{C} \cup \mathcal{C}_r) = \mathcal{A}(\mathcal{C})$, so we conclude that $(\mathcal{A}(\mathcal{C}) \cup \mathcal{A}(\mathcal{C}'))'' = \mathcal{B}(\mathcal{H})$.

Proof of 3

The proof of this step follows automatically from MC and premise 2—so much so that Haag and Schroer don't include a proof! However, here is one. From MC, $\mathcal{A}(\mathcal{C}') \subseteq \mathcal{A}(\mathcal{C})'$. We can use this inclusion and premise 2 to prove that $\mathcal{A}(\mathcal{C})$ is a von Neumann factor.

Proof.

$$\mathcal{A}(\mathcal{C}') \subseteq \mathcal{A}(\mathcal{C})'$$
$$\mathcal{A}(\mathcal{C}) \cup \mathcal{A}(\mathcal{C}') \subseteq \mathcal{A}(\mathcal{C}) \cup \mathcal{A}(\mathcal{C})'$$
$$(\mathcal{A}(\mathcal{C}) \cup \mathcal{A}(\mathcal{C}'))' \supseteq (\mathcal{A}(\mathcal{C}) \cup \mathcal{A}(\mathcal{C})')'$$
$$(\mathcal{A}(\mathcal{C}) \cup \mathcal{A}(\mathcal{C}'))'' \subseteq (\mathcal{A}(\mathcal{C}) \cup \mathcal{A}(\mathcal{C})')''$$
$$\mathcal{B}(\mathcal{H}) = (\mathcal{A}(\mathcal{C}) \cup \mathcal{A}(\mathcal{C}'))'' \subseteq (\mathcal{A}(\mathcal{C}) \cup \mathcal{A}(\mathcal{C})')'' \subseteq \mathcal{B}(\mathcal{H})$$
$$(\mathcal{A}(\mathcal{C}) \cup \mathcal{A}(\mathcal{C})')'' = \mathcal{B}(\mathcal{H})$$

□

Since the set of von Neumann algebras on a Hilbert space forms an orthocomplemented lattice (Haag, 1996, p. 114), taking the double commutant on the left-hand side of the last equation is the identity operation, so $\mathcal{A}(\mathcal{C}) \cup \mathcal{A}(\mathcal{C})' = \mathcal{B}(\mathcal{H})$, which is precisely premise 3.

Proof of the conclusion

The proof of this final step was prepared as a supplement for the argument in the main text as an alternative ending from the version of the proof I had done in 2019 using Haag duality. I owe Noel Swanson the suggestion to try to use essential duality instead.

To prove that $\mathcal{A}(\mathcal{C}) = \mathcal{A}(\mathcal{C}'')$, I will show the double inclusion.

- One side follows from MC and essential duality: $\mathcal{A}(\mathcal{C}) \subseteq \mathcal{A}(\mathcal{C}')' = \mathcal{A}(\mathcal{C}'')$.

- For the other side, our starting point is premise 3. Since $\mathcal{A}(\mathcal{C})$ is a von Neumann factor, $\mathcal{A}(\mathcal{C}) \cup \mathcal{A}(\mathcal{C})' = \mathcal{B}(\mathcal{H})$ or, equivalently, that $\mathcal{A}(\mathcal{C})' \cap \mathcal{A}(\mathcal{C}) = \mathbb{C} \operatorname{Id}_{\mathcal{B}(\mathcal{H})}$. Then, since $\mathcal{A}(\mathcal{C}')' \subseteq \mathcal{B}(\mathcal{H}) = \mathcal{A}(\mathcal{C}) \cup \mathcal{A}(\mathcal{C})'$, the elements of $\mathcal{A}(\mathcal{C}')'$ are in either $\mathcal{A}(\mathcal{C})$, which is what we want, or in $\mathcal{A}(\mathcal{C})'$. So we want to prove they are not in $\mathcal{A}(\mathcal{C})'$. Assume, for *reductio*, that $\mathcal{A}(\mathcal{C}')' \subseteq \mathcal{A}(\mathcal{C})'$, which implies that $\mathcal{A}(\mathcal{C})'' \subseteq \mathcal{A}(\mathcal{C}')''$. Since these are von Neumann algebras, $\mathcal{A}(\mathcal{C})'' \subseteq \mathcal{A}(\mathcal{C}')''$ is equivalent to $\mathcal{A}(\mathcal{C}) \subseteq \mathcal{A}(\mathcal{C}')$. Then, for some $A \in \mathcal{A}(\mathcal{C})$, $A \in \mathcal{A}(\mathcal{C}')$. Since $\mathcal{A}(\mathcal{C})$ is a von Neumann factor and $A \in \mathcal{A}(\mathcal{C})$, $A \notin \mathcal{A}(\mathcal{C})'$ unless $A = c \operatorname{Id}_{\mathcal{B}(\mathcal{H})}$ for some $c \in \mathbb{C}$. For non-trivial operators, if $A \notin \mathcal{A}(\mathcal{C})'$, $A \notin \mathcal{A}(\mathcal{C}')''$ because $\mathcal{A}(\mathcal{C})' \subseteq \mathcal{A}(\mathcal{C}')''$ by MC. But, since these are von Neumann algebras, $\mathcal{A}(\mathcal{C}')'' = \mathcal{A}(\mathcal{C}')$, so $A \notin \mathcal{A}(\mathcal{C}')$, which is a contradiction. Therefore, $\mathcal{A}(\mathcal{C}')' \not\subseteq \mathcal{A}(\mathcal{C})'$. Since $\mathcal{A}(\mathcal{C}')' \subseteq \mathcal{A}(\mathcal{C}) \cup \mathcal{A}(\mathcal{C})'$, we conclude that $\mathcal{A}(\mathcal{C}')' \subseteq \mathcal{A}(\mathcal{C})$. Using essential duality once again, $\mathcal{A}(\mathcal{C}'') \subseteq \mathcal{A}(\mathcal{C})$.

References

Bogolubov, N. N., A. A. Logunov, A. I. Oksak, and I. Todorov (1987). *General Principles of Quantum Field Theory*. Mathematical Physics and Applied Mathematics. Springer Netherlands.

Bohr, N. (1948). ON THE NOTIONS OF CAUSALITY AND COMPLEMENTARITY. *Dialectica 2*(3/4), 312–319.

Borchers, H. J. (1961). Über die Vollständigkeit lorentzinvarianter Felder in einer zeitartigen Röhre. *Il Nuovo Cimento (1955-1965) 19*(4), 787–793.

Bostelmann, H., C. J. Fewster, and M. H. Ruep (2021). Impossible measurements require impossible apparatus. *Physical Review D 103*(2), 025017.

Buchholz, D. and J. Yngvason (1994). There are No Causality Problems for Fermi's Two Atom System. *Physical Review Letters 73*(5), 613–616.

Busch, P. (1999). Unsharp localization and causality in relativistic quantum theory. *Journal of Physics A: Mathematical and General 32*(37), 6535–6546.

Butterfield, J. (2007). Reconsidering Relativistic Causality. *International Studies in the Philosophy of Science 21*(3), 295–328.

Bär, C., N. Ginoux, and F. Pfäffle (2007). *Wave Equations on Lorentzian Manifolds and Quantization*. European Mathematical Society.

Calderón Ossa, F. (2019). A survey of causality in algebraic relativistic quantum field theory [Undergraduate thesis, Universidad de los Andes].

Connes, A. and E. Størmer (1978). Homogeneity of the state space of factors of type III$_1$. *Journal of Functional Analysis 28*(2), 187–196.

Curiel, E. (2017). A Primer on Energy Conditions. In D. Lehmkuhl, G. Schiemann, and E. Scholz (Eds.), *Towards a Theory of Spacetime Theories*, pp. 43–104. Springer.

Dimock, J. (1980). Algebras of local observables on a manifold. *Communications in Mathematical Physics 77*(3), 219–228.

Dimock, J. (1982). Dirac quantum fields on a manifold. *Transactions of the American Mathematical Society 269*(1), 133–147.

Drago, N. and V. Moretti (2020). The notion of observable and the moment problem for *-algebras and their GNS representations. *Letters in Mathematical Physics 110*(7), 1711–1758.

Duncan, A. (2012). Dynamics IV: Aspects of locality: clustering, microcausality, and analyticity. In *The Conceptual Framework of Quantum Field Theory*, pp. 132–170. Oxford University Press.

Earman, J. (2014). No superluminal propagation for classical relativistic and relativistic quantum fields. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 48*, 102–108.

Earman, J. and G. Valente (2014). Relativistic Causality in Algebraic Quantum Field Theory. *International Studies in the Philosophy of Science 28*, 1–48.

Einstein, A. (1948). QUANTEN-MECHANIK UND WIRKLICHKEIT. *Dialectica 2*(3/4), 320–324.

Epstein, H., V. Glaser, and A. Jaffe (1965). Nonpositivity of the energy density in quantized field theories. *Il Nuovo Cimento (1955-1965) 36*(3), 1016–1022.

Fermi, E. (1932). Quantum Theory of Radiation. *Reviews of Modern Physics 4*(1), 87–132.

Fewster, C. J. (2016). The split property for quantum field theories in flat and curved spacetimes. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg 86*(2), 153–175.

Fewster, C. J. and R. Verch (2020). Quantum Fields and Local Measurements. *Communications in Mathematical Physics 378*(2), 851–889.

Fraser, D. (2011). How to take particle physics seriously: A further defence of axiomatic quantum field theory. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 42*(2), 126–135.

Geroch, R. (2011). Faster Than Light? In *Advances in Lorentzian Geometry: Proceedings of the Lorentzian Geometry Conference in Berlin*, Volume 49 of *AMS/IP Studies in Advanced Mathematics*, pp. 59–69. American Mathematical Society.

Haag, R. (1957/2010). Discussion of the 'axioms' and the asymptotic properties of a local field theory with composite particles. *The European Physical Journal H 35*(3), 243–253.

Haag, R. (1996). *Local Quantum Physics: Fields, Particles, Algebras* (2nd ed.). Springer.

Haag, R. and B. Schroer (1962). Postulates of Quantum Field Theory. *Journal of Mathematical Physics 3*(2), 248–256.

Halvorson, H. (2001). Reeh-Schlieder Defeats Newton-Wigner: On Alternative Localization Schemes in Relativistic Quantum Field Theory. *Philosophy of Science 68*(1), 111–133.

Halvorson, H. (2007). Algebraic Quantum Field Theory (with an Appendix by Michael Müger). In J. Butterfield and J. Earman (Eds.), *Philosophy of Physics*, pp. 731–864. North-Holland.

Hegerfeldt, G. C. (1994). Causality problems for Fermi's two-atom system. *Physical Review Letters 72*(5), 596–599.

Hofer-Szabó, G. and P. Vecsernyés (2018). Locality and Causality Principles. In *Quantum Theory and Local Causality*, pp. 17–23. Springer.

Horuzhy, S. S. (1990). *Introduction to Algebraic Quantum Field Theory*. Kluwer Academic Publishers.

Kitajima, Y. (2018). Local Operations and Completely Positive Maps in Algebraic Quantum Field Theory. In M. Ozawa, J. Butterfield, H. Halvorson, M. Rédei, Y. Kitajima, and F. Buscemi (Eds.), *Reality and Measurement in Algebraic Quantum Theory*, pp. 83–95. Springer Singapore.

Moretti, V. (2017). *Spectral Theory and Quantum Mechanics : Mathematical Foundations of Quantum Theories, Symmetries and Introduction to the Algebraic Formulation* (2nd ed.). Springer.

Murphy, G. J. (1990). *C*–Algebras and Operator Theory*. San Diego: Academic Press.

Okamura, K. and M. Ozawa (2016). Measurement theory in local quantum physics. *Journal of Mathematical Physics 57*(1), 015209.

Papageorgiou, M. and D. Fraser (2023). Eliminating the "impossible": Recent progress on local measurement theory for quantum field theory. arXiv:2307.08524.

Peskin, M. E. and D. V. Schroeder (1995). *An Introduction to Quantum Field Theory*. Westview Press.

Rudin, W. (1971). *Lectures on the Edge-of-the-Wedge Theorem*, Volume 6 of *CBMS Regional Conference Series in Mathematics*. Providence, Rhode Island: American Mathematical Society.

Ruetsche, L. (2011). *Interpreting Quantum Theories*. Oxford University Press.

Ruetsche, L. (2021). Locality in (Axiomatic) Quantum Field Theory: A Minority Report. In E. Knox and A. Wilson (Eds.), *The Routledge Companion to Philosophy of Physics*, pp. 311–322. Routledge.

Ruetsche, L. and J. Earman (2011). Interpreting Probabilities in Quantum Field Theory and Quantum Statistical Mechanics. In C. Beisbart and S. Hartmann (Eds.), *Probabilities in Physics*, pp. 263–290. Oxford University Press.

Rédei, M. (2014). A categorial approach to relativistic locality. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 48*, 137–146.

Rédei, M. and G. Valente (2010). How local are local operations in local quantum field theory? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 41*(4), 346–353.

Sorkin, R. (1993). Impossible Measurements on Quantum Fields. In B. Hu and T. Jacobson (Eds.), *"Directions in General Relativity", Proceedings of the 1993 International Symposium, Maryland, Vol. II: Papers in honor of Dieter Brill*, pp. 293–305. Cambridge UP.

Streater, R. F. and A. S. Wightman (1964/2000). *PCT, Spin and Statistics, and All That*. Princeton University Press.

Summers, S. J. and R. Werner (1988). Maximal violation of Bell's inequalities for algebras of observables in tangent spacetime regions. *Annales de l'I.H.P. Physique théorique 49*(2), 215–243.

Swanson, N. (2017). A philosopher's guide to the foundations of quantum field theory. *Philosophy Compass 12*(5), e12414.

Valente, G. (2014). Does the Reeh–Schlieder theorem violate relativistic causality? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 48*, 147–155.

Vladimirov, V. S., V. V. Zharinov, and A. G. Sergeev (1994). Bogolyubov's "edge of the wedge" theorem, its development and applications. *Russian Mathematical Surveys 49*(5), 51.

Wallace, D. (2006). In Defence of Naiveté: The Conceptual Status of Lagrangian Quantum Field Theory. *Synthese 151*(1), 33–80.

Wallace, D. (2011). Taking particle physics seriously: A critique of the algebraic approach to quantum field theory. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 42*(2), 116–125.

Weinberg, S. (1995). Quantum Fields and Antiparticles. In *The Quantum Theory of Fields*, Volume 1, pp. 191–258. Cambridge University Press.

Wightman, A. S. (1996). How It Was Learned that Quantized Fields Are Operator-Valued Distributions. *Fortschritte der Physik/Progress of Physics 44*(2), 143–178.

Yngvason, J. (2005). The Role of Type III Factors in Quantum Field Theory. *Reports on Mathematical Physics 55*(1), 135–147.