Informational Pruning as a Third Path: Final-State Constraints between Consistent Quantum Causes and Quantum Causal Models

Original Paper

Ian Staley
Independent Researcher
ORCID: 0009-0000-8592-3186
E-mail: ian.t.staley@gmail.com

Draft: May 2026

Abstract:

Two recent programs offer competing accounts of quantum causation. Griffiths' Consistent Quantum Causes grounds causal reasoning in the Consistent Histories formalism, treating stochastic quantum time development and noncommuting projectors as the proper foundation for a Pearl-style directed acyclic graph representation. The Costa–Shrapnel Quantum Causal Models program takes the process-matrix formalism as its starting point, generalizing classical interventionist causal discovery to the quantum domain and extending naturally to indefinite causal order. Both programs assume that causal structure is fixed by an initial condition together with forward dynamics. Neither addresses what happens when admissibility is conditioned at both temporal boundaries. In this paper, I argue that final-state-constrained history ensembles, and the informational pruning they induce, provide a third vantage point from which to compare these two programs. I do not claim to resolve the Griffiths–QCM disagreement, nor to construct a new quantum causal formalism from scratch. The more limited claim is that once admissibility is globally filtered by an initial and a final boundary condition, the question of which coarse-grained histories carry the effective causal and computational structure of the process is reframed. The Hartle–Craig rank-product theorem supplies the formal backbone: imposing a final-state projector with rank r_ω restricts the admissible-history family to at most r_α·r_ω nonzero-probability histories, where r_α is the initial-state rank. I illustrate this with the three-box paradox, treated as a joint Consistent Histories and process-matrix scenario under final-state conditioning. I also suggest that an explainability stance, drawn in a methodological sense from explainable artificial intelligence, is useful for asking which admissible histories do the explanatory work under a given endpoint class. I conclude that informational pruning offers a coherent third path that is neither a competitor to Griffiths' approach nor an extension of process-matrix interventionism, but a complementary perspective that sharpens what each program leaves implicit.

Keywords: Consistent histories; Quantum causal models; Final boundary condition; Informational pruning; Two-state vector formalism; Rank-product theorem; Three-box paradox

1. Introduction

Quantum mechanics has generated several distinct programs aimed at extracting a theory of causation from its formalism. Two such programs have matured recently to the point where their differences can be stated sharply. The first, developed by Griffiths, grounds quantum causal reasoning in the Consistent Histories approach, treating quantum time development as inherently stochastic and emphasizing the noncommutation of projectors as the central obstacle to any naïve importation of classical causal graphs [1]. The second, developed by Costa, Shrapnel, and collaborators, takes the process-matrix formalism as its starting point and generalizes classical interventionist causal discovery to the quantum domain [2,3]. Both programs are sophisticated, both have been elaborated technically, and both are now actively contested in the foundations literature.

Despite their differences, the two programs share an unexamined commitment. Each assumes that quantum causal structure is fixed by an initial condition together with forward dynamics, and that the task of a quantum causal theory is to describe how causal influence propagates from earlier to later events under that assumption. Neither program has engaged systematically with dual-boundary conditioning, in which admissibility is determined not by an initial state alone but by an initial state together with a final-state constraint. Yet time-symmetric formulations of quantum theory have a long history [4,5], and recent work has argued that final-state constraints can act as globally significant admissibility filters on history ensembles, not merely as local post-selection devices [6].

The purpose of this paper is to argue that the informational pruning induced by final-state constraints provides a third vantage point on the debate between Consistent Quantum Causes and Quantum Causal Models. I do not propose a new formal theory of quantum causation. I also do not claim to resolve the technical disagreements between Griffiths and the Costa–Shrapnel program. The more limited claim is that each program can be re-examined through the lens of a constrained history-ensemble framework, and that doing so clarifies what each approach leaves implicit about the admissibility conditions under which its causal structure is defined.

The positioning argument must engage two recent papers that approach the CH/QCM terrain from adjacent directions. Ormrod and Barrett [7] synthesize relational quantum mechanics, consistent histories, and quantum causal modelling into a single framework in which preferred projective decompositions are picked out by an internal causal mechanism. Anastopoulos [8] uses final-state-conditioned decoherent histories to address cosmological post-selection. The present paper differs from both: Ormrod–Barrett derive preferred histories from internal causal dynamics rather than from an external final-state projector, and Anastopoulos applies final-state conditioning to cosmological dynamics rather than deploying it as a positioning device between CH and QCM. Adlam [9] provides the philosophical opening that the present positioning aims to occupy, arguing that the process-matrix program identifies a symmetric "causal order" relation rather than asymmetric cause-effect causation. The taxonomy developed by Adlam, Hance, Hossenfelder, and Palmer [10] further clarifies where informational pruning sits relative to retrocausal, future-input-dependent, and atemporal alternatives.

The argument of the paper proceeds in several stages. Section 2 reviews the two programs and the final-state-constrained framework developed in prior work, and summarizes the comparison in Table 1. Section

3 identifies the shared assumption of unilateral temporal conditioning and shows how it is expressed differently in each program. Section 4 develops the positioning argument: informational pruning reframes the question of which coarse-grained histories carry effective causal or computational structure, and it does so without adjudicating the formal dispute between Consistent Histories and process-matrix approaches. Section 5 illustrates the framework with a worked example using the three-box paradox, treated as a joint Consistent Histories and process-matrix scenario, with the Hartle–Craig rank-product theorem as the central formal statement. Section 6 introduces an explainability stance, drawn methodologically from explainable artificial intelligence, and uses it to sharpen the question of which admissible histories carry explanatory weight under a given endpoint class. Section 7 considers objections, including the charge that the positioning is merely rhetorical rather than substantive. Section 8 indicates where the resulting perspective might be developed further. Section 9 concludes.

A note on scope is appropriate at the outset. This paper is a positioning paper, not a theorem paper. It takes as given the formal machinery of Consistent Histories, of Quantum Causal Models, and of the final-state-constrained framework developed in [6], and it argues that a certain relation among them is worth making explicit. Where the paper makes substantive claims, those claims depend on the formal results already established in the cited programs. Where the paper is interpretive, I mark it as such.

2. Three Frameworks for Quantum Causation

The argument of this paper requires a careful statement of what each of the three frameworks under discussion actually commits to. This section reviews the Consistent Quantum Causes program, the Quantum Causal Models program, and the final-state-constrained framework developed in earlier work. The review is not exhaustive. Its aim is to surface the features that matter for the positioning argument developed in later sections. A consolidated comparison is given in Table 1.

2.1. Consistent Quantum Causes

The Consistent Histories approach to quantum mechanics was introduced by Griffiths in 1984 [11] and developed further by Omnès, Gell-Mann, and Hartle [12]. The central idea is that quantum time development should be treated as inherently stochastic, with the Schrödinger equation supplying probabilities for alternative histories rather than a deterministic evolution of a single state vector. A history in this framework is a sequence of quantum events represented by projectors at successive times. A family of histories satisfying a consistency condition supports standard Kolmogorov probability assignment and permits Boolean reasoning within that family. Recent numerical work has demonstrated emergent decoherent histories from first principles in many-body systems [13], establishing that the framework remains a live empirical program and not merely an interpretive stance.

In recent work, Griffiths has extended the Consistent Histories approach explicitly to quantum causation [1]. The argument is that developing a quantum analog of Pearl-style classical causal reasoning requires a genuine theory of stochastic quantum time development, and that the Consistent Histories formalism provides precisely that theory. The noncommutation of quantum projectors is treated not as an obstacle to be worked around but as the essential feature that any serious quantum causal account must confront.

Griffiths applies this framework to simple systems, including beam splitters and Mach-Zehnder interferometers, and uses it to address the Bell inequalities.

Three features of this program are relevant for present purposes. First, it treats causal reasoning as internal to a chosen consistent framework. Multiple incompatible frameworks exist, and questions about causation must be posed within one framework at a time [14]. Second, it assumes a fixed initial condition and time-forward evolution. The histories are constrained by an initial state or initial density operator, and the stochastic process unfolds forward under the Schrödinger dynamics. Third, Griffiths explicitly positions his account against the Quantum Causal Models program, arguing that the process-matrix approach is mathematically elaborate in ways that may obscure rather than clarify the underlying causal structure.

2.2. Quantum Causal Models

The Quantum Causal Models program developed by Costa, Shrapnel, and collaborators takes a different starting point [2,3]. Its foundation is the process-matrix formalism, which describes quantum experiments in terms of operations performed at distinct parties connected by quantum channels, without prior assumptions about causal order. On that foundation, the program defines quantum analogues of classical causal-modelling tools, including the causal Markov condition and faithfulness, and it develops causal discovery algorithms that recover causal structure from process-matrix data [15].

The program has several notable strengths. It extends naturally to indefinite causal order, which is widely studied in the quantum information literature and has led to experimental and theoretical results on causal nonseparability [16,17]. It is interventionist in the Woodward–Pearl sense, treating causal structure as what is revealed by localized interventions. It has been extended to non-Markovian processes and to multi-time structures relevant for noisy quantum devices [18]. It has also generated categorical reformulations, including Markov-categorical approaches that address the quantum no-cloning theorem directly [19].

Adlam [9] argues that what the process-matrix program identifies is a symmetric "causal order" relation rather than an asymmetric cause-effect relation of the kind classical causal analysis recognizes. If that argument is correct, then the QCM program and the Consistent Quantum Causes program are not addressing the same explanatory target, and Griffiths' criticisms of QCM may be partially vindicated while leaving room for a complementary framework that treats boundary structure as primary rather than derivative. The positioning argument of this paper exploits precisely this opening.

For present purposes, the relevant features of QCM are the following. Like the Consistent Quantum Causes program, it assumes that causal structure is defined with respect to an initial condition together with forward dynamics, even when that forward dynamics is generalized to allow indefinite causal order. Interventionist discovery operates by considering what would happen if an intervention were imposed at a given locus, and tracing the implications of that intervention forward to subsequent events. Final-state constraints, in the sense of a globally imposed boundary condition that filters admissibility from outside the dynamics, are not part of the QCM machinery.

2.3. Final-State Constraints and Informational Pruning

A third framework, developed in [6], takes a different starting point again. Rather than building from either Consistent Histories or process matrices, it builds from the time-symmetric formalism of Aharonov, Bergmann, and Lebowitz, in which the probability of an intermediate outcome is conditioned by both a pre-selected initial state and a post-selected final state [4,5]. The Aharonov–Bergmann–Lebowitz rule

$$P(q_i \mid \psi, \phi) = |\langle\phi|\Pi_i|\psi\rangle|^2 / \Sigma_j \, |\langle\phi|\Pi_j|\psi\rangle|^2 \qquad (1)$$

gives the conditional probability of an intermediate outcome q_i given both boundary conditions, and serves as the formal entry point by which final-state constraints enter the space of admissible histories. In [6], the universe's possible macroscopic development is modelled by a directed acyclic history graph G whose vertices are coarse-grained macroscopic configurations and whose edges are allowed transitions. Imposing a final boundary condition $\langle\phi|$ induces an admissible subgraph $G' \subseteq G$ containing only those histories compatible with both temporal boundaries.

Two features of that framework are relevant here. The first is informational: under a history-ensemble encoding, final-state conditioning may reduce the effective description length of realized histories relative to an unconstrained branching representation. Where K denotes effective description length in the Kolmogorov sense, one considers the question of whether

$$K(G'_\phi) + K(\phi) < K(G) \qquad (2)$$

holds for a given endpoint class. If it does, the constrained representation is informationally preferable under the chosen encoding. The argument in [6] distinguishes endpoint classes by their structural complexity, with trivial, structured, and computational terminal states inducing progressively stronger pruning. Computational terminal states provide the strongest selectivity and, under a structural rather than psychological definition of agency, support an agentic interpretive reading.

The second feature is that the framework is explicitly retrocausal in boundary architecture but non-interventionist in mechanism. Histories incompatible with both boundary conditions are excluded from the admissible ensemble rather than revised after the fact. The framework does not require operational backward signaling, nor does it claim that a future state reaches backward to overwrite a fixed past. It is a consistency filter imposed by global admissibility, not a dynamical intervention. In the taxonomy of [10], this places informational pruning within the boundary-conditioned rather than the interventionist-retrocausal category — a distinction that matters for the positioning argument, since it dissolves a class of objections that would apply to interventionist retrocausality but not to global boundary conditioning.

2.4. Two Closely Related Recent Frameworks

Two 2024 papers approach territory adjacent to the present argument and require explicit differentiation.

Ormrod and Barrett [7] propose a synthesis of relational quantum mechanics, consistent histories, and quantum causal modelling, in which events obtain relative to a set of systems and correspond to projectors picked out by causal structure. Their mechanism for selecting preferred consistent families is internal: dynamic causal relations among systems determine which projective decompositions avoid "interference influence" with later measurements. The present framework differs at the level of what does the selecting. For Ormrod and Barrett, consistent families are selected by internal causal-dynamical structure. For the framework here, consistent families compatible with a final-state projector are selected by an external boundary condition imposed on the history ensemble. Both approaches reduce the space of admissible histories, but they reduce it for different reasons and they appeal to different foundational resources. Ormrod and Barrett do not invoke final-state constraints, and the present framework does not invoke event relativity.

Anastopoulos [8] applies final-state-conditioned decoherent histories to quantum cosmology, deriving effective equations for an FRW universe under post-selection and arguing that cosmic acceleration can be read as a quantum post-selection effect. The technical formalism overlaps substantially with what is used in [6] and in the present paper, which is useful evidence that final-state-conditioned decoherent histories is an active, not merely proposed, research tool. The difference is application scope and interpretive layer. Anastopoulos deploys the formalism for cosmological prediction; the present paper deploys it for positioning between CH and QCM, with informational pruning and an explainability stance rather than cosmological dynamics as the interpretive content.

2.5. Consolidated Comparison

Table 1 consolidates the review of the three frameworks across seven dimensions. The table makes visible what the prose establishes: the Consistent Quantum Causes and Quantum Causal Models programs differ on several important dimensions, but they agree on the structural absence of a final-state admissibility condition. The final-state-constrained framework is the only one of the three that makes dual-boundary conditioning a structural element. This agreement-by-omission between the first two programs is what the positioning argument of Sections 3 and 4 develops.

Table 1. Comparison of Consistent Quantum Causes

Dimension	Consistent Quantum Causes (Griffiths)	Quantum Causal Models (Costa–Shrapnel)	Informational Pruning (Staley)
Formal starting point	Consistent Histories framework; projectors at successive times	Process-matrix formalism; operations at distinct parties	Two-state vector formalism; history graph with admissible subgraph
Initial condition	Specified initial state or density operator	Specified input state; forward process matrix	Specified initial boundary condition $\|\psi\rangle$
Final condition	Not a structural element; unilateral forward dynamics	Not a structural element; no globally imposed terminal projector	Final boundary condition $\langle\phi\|$ as global admissibility filter
Causal structure	Internal to a chosen consistent framework; framework-dependent	Interventionist in Woodward–Pearl sense; recovered via discovery algorithms	Not proposed; inherited from either CH or QCM as applied to admissible ensemble
Indefinite causal order	Not addressed as separate topic; framework-selection handles adjacency	Natural extension via process-matrix formalism	Orthogonal to the framework; compatible with either CH or QCM treatment

Dimension	Consistent Quantum Causes (Griffiths)	Quantum Causal Models (Costa–Shrapnel)	Informational Pruning (Staley)
Treatment of noncommutation	Essential feature; frames the entire program	Handled via process-matrix structure; no-cloning addressed in categorical reformulation	Inherited from underlying framework (CH or QCM); not independently addressed
Retrocausal commitment	Explicitly non-retrocausal	Non-retrocausal; indefinite causal order is not backward causation	Retrocausal in boundary architecture; non-interventionist in mechanism

Table 1. *Comparison of Consistent Quantum Causes, Quantum Causal Models, and the final-state-constrained framework across seven dimensions. The shared assumption of unilateral temporal conditioning in the first two programs, contrasted with the dual-boundary structure of the third, is the formal basis for the positioning argument developed in Section 4.*

Figure 1. Three Frameworks for Quantum Causation

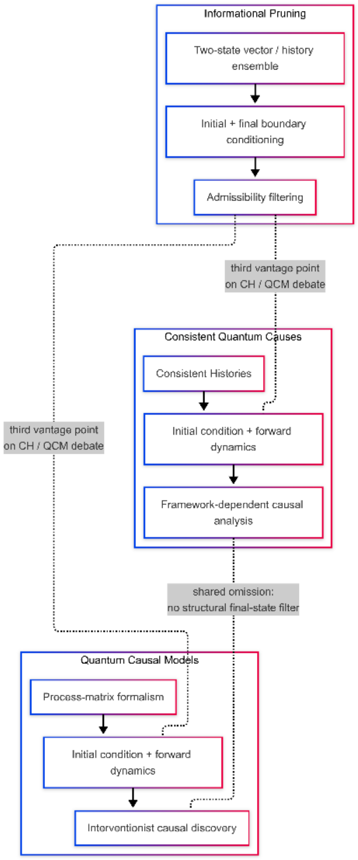

Figure 1. *Comparison of the three frameworks discussed in the paper. Consistent Quantum Causes and Quantum Causal Models differ in formal starting point and causal methodology, but both assume unilateral temporal conditioning. Informational pruning differs by treating final-state admissibility as a structural element.*

3. The Shared Assumption of Unilateral Temporal Conditioning

Having reviewed the three frameworks, I now turn to the observation that motivates the positioning argument of the paper. The Consistent Quantum Causes and Quantum Causal Models programs disagree on many things. They disagree about the right formal starting point for a theory of quantum causation. They disagree about whether process matrices or quantum histories provide the more fundamental structure. They disagree about the status of indefinite causal order. And, in the explicit framing of Griffiths' 2024 paper, they disagree about whether the QCM program's machinery helps or obscures the underlying physics [1].

What they agree on, though largely without saying so, is that quantum causal structure is determined by an initial condition together with forward dynamics. This shared assumption is not trivial. It reflects the deep influence of forward-evolving state-vector descriptions on how quantum foundations has been conducted, and it is inherited by both programs from the wider tradition in which they sit.

Figure 2. Unilateral vs dual-boundary conditioning

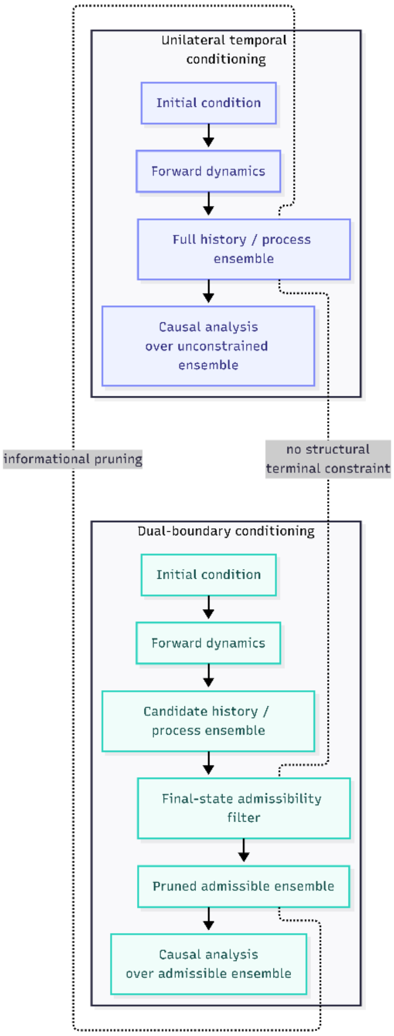

Figure 2. *Contrast between unilateral temporal conditioning and dual-boundary conditioning. Under unilateral conditioning, causal analysis is performed on the unconstrained history or process ensemble. Under dual-boundary conditioning, a final-state admissibility condition prunes the ensemble before causal analysis, yielding informational pruning relative to the unconstrained case..*

3.1. Unilateral Conditioning in Consistent Quantum Causes

In the Consistent Histories formalism as Griffiths develops it, a family of histories is defined with respect to an initial state or initial density operator. The consistency condition required for Boolean probability assignment is a condition on the joint evolution of projectors under forward unitary dynamics. Causal relations, when introduced, are relations between events at different times within such a family, read forward from earlier to later. There is no formal analogue, within Consistent Quantum Causes as currently developed, of a final-state constraint that acts as a global admissibility filter on the history ensemble.

This is not an oversight. The Consistent Histories approach is compatible with time-symmetric formulations and has been extended in that direction by some authors, most notably in the Hartle treatment of time-neutral boundary conditions [20]. The point is rather that the current Consistent Quantum Causes program does not make use of final-state admissibility as a structural element of its causal analysis. Histories are selected for consistency and analyzed for causal structure under unilateral temporal conditioning.

3.2. Unilateral Conditioning in Quantum Causal Models

The QCM program is unilateral in a somewhat different way. Its process-matrix formalism is explicitly designed to be agnostic about background causal order, and it admits indefinite causal order as a natural generalization. In that sense the program is not committed to a simple past-to-future reading of causal influence. Nonetheless, interventionist causal discovery as implemented in the program works by considering localized interventions at specified loci and tracing their implications under the dynamics encoded in the process matrix. The structure of the process matrix is itself determined by the preparation of the input state and the channels connecting operations. Final-state constraints, treated as globally imposed boundary conditions distinct from any prepared input, do not enter.

It is possible that the process-matrix formalism could be extended to include final-state boundary conditions, and there is philosophical work on retrocausality that touches on causal-model framing [21,22]. But the QCM program as currently developed is not organized around final-state conditioning. Interventions at a locus propagate causal consequences through the process matrix forward from that locus, and what is revealed by such interventions is the causal structure as it exists under the prepared input and forward dynamics.

3.3. What Final-State Conditioning Adds

The final-state-constrained framework developed in [6] differs from both programs on this point. It does not deny the validity of either unilateral framework within its own domain. What it adds is a further admissibility filter imposed at the global level, by which histories are excluded from the admissible ensemble if they are incompatible with a specified final boundary condition. This is not a revision of the dynamics internal to either Consistent Histories or process matrices. It is an additional condition on which histories are treated as admissible, imposed from outside the forward dynamics of either framework.

The key point is structural. Under unilateral conditioning, the space of histories available for causal analysis is fixed by initial preparation and forward dynamics. Under dual-boundary conditioning, that space is filtered further by compatibility with a terminal condition. Both Consistent Quantum Causes and Quantum

Causal Models can, in principle, be applied to histories within the admissible ensemble G' rather than to histories within the unconstrained graph G. The difference is not a difference in the formal machinery of either program. It is a difference in what is being analyzed: the space of histories prior to final-state filtering, or the space of histories after it.

4. Informational Pruning as a Third Vantage Point

The observation developed in Section 3 is straightforward but has a substantive consequence. Once the shared assumption of unilateral temporal conditioning is made explicit, it becomes possible to ask what each program looks like when the histories it analyzes are first filtered by a final-state admissibility condition. This section develops that question into a positioning claim.

4.1. The Positioning Claim

The positioning claim is this. Informational pruning, as developed in the final-state-constrained framework, provides a third vantage point from which to evaluate the Consistent Quantum Causes and Quantum Causal Models programs. It is not a competitor to either program at the level of formal machinery. It does not propose new consistency conditions for quantum histories, nor does it propose new interventionist algorithms on process matrices. What it proposes is that the space of histories over which either program operates can be filtered by a final-state admissibility condition, and that doing so changes which histories count as explanatorily central.

This framing yields three distinct claims. The first is that applying final-state admissibility to the Consistent Histories formalism shifts the target of causal analysis from the full set of consistent histories to the admissible subset compatible with the terminal condition. Within that admissible subset, the noncommutation considerations and framework selections emphasized by Griffiths continue to operate, but they operate on a smaller and structurally distinct space. The second claim is that applying final-state admissibility to the Quantum Causal Models framework similarly restricts the space of processes over which interventionist discovery is performed. The process-matrix machinery continues to apply, but the processes under consideration are those whose structure is compatible with a specified terminal condition. The third claim is that these two restrictions, though they operate on different formal objects, are related at the informational level: both reduce the effective description length of the analyzed ensemble relative to its unconstrained counterpart, and both shift the interpretive weight toward histories that carry the structure required for terminal-condition realization.

4.2. Formal Backbone: The Rank-Product Theorem

The positioning claim has a formal backbone in the time-neutral Consistent Histories literature. Craig [23] showed that the maximum number of histories with nonzero probability in a decoherent family, given an initial density operator ρ_α of rank r_α and a final density operator ρ_ω of rank r_ω, is bounded by the product r_α · r_ω. Hartle [20] subsequently showed that the existence of a decoherent set with both nontrivial initial and final conditions requires time-asymmetric boundaries, and that the rank structure of the boundary density operators directly controls which families can be made to decohere.

This result can be read as follows. Without a final-state constraint, the admissible history family is bounded in complexity only by the initial density operator and the dynamics. With a final-state projector of rank r_ω, the admissible family has at most $r_\alpha \cdot r_\omega$ nonzero-probability histories. This is a precise quantitative statement of the pruning effect. It is also a formal result in the Consistent Histories tradition itself, established independently of the final-state-constrained framework developed in [6]. The informational argument of [6] can therefore be read as an interpretive overlay on an established formal result rather than as a freestanding claim that requires independent technical justification.

The rank-product theorem also connects the informational pruning framework to the Dowker–Kent set-selection problem. Consistency underdetermines the choice of history family; many mutually incompatible consistent families coexist for any given initial condition. The Dowker–Kent challenge is that standard Consistent Histories provides no principled reason to prefer one consistent family over another. Informational pruning offers a partial response: among the consistent families compatible with both an initial and a final boundary condition, the Hartle–Craig bound provides a quantitative measure of selectivity, and the description-length inequality from [6] provides a preference ordering on terminal-condition classes. This does not solve the set-selection problem, but it supplies an additional selection criterion that becomes available once dual-boundary conditioning is taken seriously.

4.3. Coarse-Grained Agency under Dual-Boundary Admissibility

A further consequence follows when the terminal condition is structured or computational in the sense developed in [6]. Under trivial terminal states, the admissible ensemble remains nearly as large as the unconstrained space, and the positioning claim adds little beyond a minor restriction on which histories count. Under structured or computational terminal states, the admissible ensemble is substantially pruned, and the histories that remain are those that support the organizational structure required for terminal-condition realization. Table 2 summarizes this hierarchy.

Table 2. Hierarchy of Terminal-Condition Classes

Endpoint class	**Specification cost $K(\phi)$**	**Pruning strength**	**Admissible ensemble size $\|\Omega_\phi\|$**
Trivial	Low (e.g., simple equilibrium or uniform final state)	Weak; few histories excluded	Nearly as large as $\|\Omega\|$; compression ratio $R_\phi \approx 1$
Structured	Moderate; encodes stable organization or differentiated structure	Stronger; histories lacking organizational features excluded	Substantially smaller than $\|\Omega\|$; $R_\phi > 1$
Computational	High; encodes information-processing and recursive self-favoring structure	Strongest; histories lacking computational organization excluded	Sharply pruned subset; $R_\phi \gg 1$ for sufficiently selective ϕ

Table 2. *Hierarchy of terminal-condition classes under the final-state-constrained framework developed in [6]. The compression ratio $R_\phi = K(G) / [K(G'_\phi) + K(\phi)]$ quantifies the informational work done by a given endpoint class. Computational terminal conditions impose the strongest selectivity, and under the structural definition of agency adopted in [6], support an agentic reading of the admissible-history ensemble.*

Within such a pruned ensemble, the question of coarse-grained agency acquires a distinct character. Under a computational terminal state, the admissible histories are precisely those in which information-processing structures, self-replicating systems, and explicit computational organization appear. From the Consistent Quantum Causes perspective, these are the histories within which the framework-dependent causal analysis has its most interpretive traction. From the Quantum Causal Models perspective, they are the processes for which interventionist causal discovery yields the most organized structural output. In both cases, the coarse-grained agential character of the admissible ensemble is not imposed from outside either program. It is a consequence of the terminal-condition complexity class.

This is where the positioning argument becomes substantive rather than merely organizational. A criticism of the final-state-constrained framework has always been that its agentic reading relies on considerations that float above the formal machinery of quantum mechanics. The positioning developed here shows that the agentic reading is connected, through the shared analysis of admissible histories, to the internal causal structures identified by both Consistent Quantum Causes and Quantum Causal Models. What the final-state-constrained framework adds is the admissibility filter. What Consistent Histories and QCM add, within that filter, is the causal analysis of the histories that survive it.

4.4. Why This Is Not a Synthesis

It is important to be explicit about what the positioning argument does not claim. It does not claim to synthesize Consistent Quantum Causes and Quantum Causal Models into a unified framework. Griffiths' objections to the QCM program, including those concerning the mathematical elaboration of the process-matrix formalism and its treatment of fine-tuning, are not addressed here. Nor does the argument resolve the Costa–Shrapnel response to Wood and Spekkens on fine-tuning and faithfulness [3]. These remain live disagreements within the quantum foundations literature.

What the positioning does do is identify a level of description at which both programs agree about something substantive, namely that unilateral temporal conditioning fixes the space of histories over which their analyses operate, and at which a third framework intervenes. The third framework does not choose between Consistent Histories and process matrices. It imposes an admissibility condition prior to either, and it invites both to be applied to the restricted ensemble. Whether one prefers Consistent Quantum Causes or Quantum Causal Models for the subsequent analysis is a question on which the final-state-constrained framework remains neutral.

5. Worked Example: The Three-Box Paradox Under Dual-Boundary Admissibility

To make the positioning claim concrete, I work through a standard scenario in the time-symmetric literature: the three-box paradox introduced by Aharonov and Vaidman [24]. This example is chosen for three reasons. First, it is already formulated in terms of pre- and post-selection, so the final-state constraint is not artificially imposed. Second, it has been treated within the full Consistent Histories framework [14,25], which makes the comparison with Consistent Quantum Causes direct. Third, it can be recast in process-matrix language, which makes the comparison with Quantum Causal Models meaningful.

5.1. The Setup

A particle is prepared at time t_1 in the pre-selected state

$$|\psi\rangle = (|1\rangle + |2\rangle + |3\rangle) / \sqrt{3} \tag{3}$$

where |i⟩ denotes the particle being in box i. At time t_3, the particle is post-selected on

$$|\phi\rangle = (|1\rangle + |2\rangle - |3\rangle) / \sqrt{3}. \tag{4}$$

Between t_1 and t_3, one may ask about the outcome of an intermediate measurement at time t_2 that determines which box the particle occupies. The Aharonov–Bergmann–Lebowitz rule yields probability 1 that the particle is in box 1 if one opens only box 1 (and probability 1 that it is in box 2 if one opens only box 2), while the particle is never in box 3 if only box 3 is opened. Read naively, this appears to say that the particle is in two places simultaneously.

The standard Consistent Histories resolution is that the three questions correspond to three incompatible frameworks, and no single consistent family contains projectors for both "in box 1" and "in box 2" at t_2. The framework-dependence is not a bug but the essential feature. The three-box paradox thus provides a clean test case for examining how final-state admissibility interacts with framework selection.

5.2. Final-State Conditioning as Admissibility Filter

In the unconstrained case, the initial state |ψ⟩ at t_1 and unitary evolution to t_3 would admit many consistent families over intermediate times, including families that assign nonzero probabilities to sequences inconsistent with the post-selection outcome |ϕ⟩. Imposing the final-state projector Π_ϕ = |ϕ⟩⟨ϕ| as an admissibility condition at t_3 restricts the admissible family to those consistent sequences whose terminal projector has nonzero overlap with |ϕ⟩.

By the Hartle–Craig rank-product theorem [20,23], the initial pure state has r_α = 1 and the final-state projector has r_ω = 1, so the admissible family has at most r_α · r_ω = 1 nonzero-probability history per framework. This is a strong restriction. Within each consistent framework compatible with both boundary conditions, a single admissible history remains. The three apparent paradoxical conclusions correspond to three different choices of consistent framework, each of which survives admissibility filtering with exactly one history. Table 3 summarizes the probability assignments across the three framework choices under both unconstrained and dual-boundary-conditioned analysis.

Table 3. Intermediate-Time Probability

Framework choice at t_2	Unconstrained probability (Born rule)	ABL probability (pre- and post-selected)	Admissible histories after rank-product filtering
Open only box 1	1/3 (in box 1); 2/3 (not in box 1)	1 (in box 1); 0 (not in box 1)	1 history per framework (r_α · r_ω = 1 · 1 = 1)
Open only box 2	1/3 (in box 2); 2/3 (not in box 2)	1 (in box 2); 0 (not in box 2)	1 history per framework (r_α · r_ω = 1 · 1 = 1)

Framework choice at t_2	Unconstrained probability (Born rule)	ABL probability (pre- and post-selected)	Admissible histories after rank-product filtering
Open only box 3	1/3 (in box 3); 2/3 (not in box 3)	0 (in box 3); 1 (not in box 3)	1 history per framework (r_α · r_ω = 1 · 1 = 1)

Table 3. *Intermediate-time probability assignments in the three-box paradox under three incompatible framework choices, with pre-selection* $|\psi\rangle = (|1\rangle+|2\rangle+|3\rangle)/\sqrt{3}$ *and post-selection* $|\phi\rangle = (|1\rangle+|2\rangle-|3\rangle)/\sqrt{3}$*. The three frameworks are mutually incompatible (no consistent family contains projectors for both "in box 1" and "in box 2" at t_2). By the Hartle–Craig rank-product theorem, each admissible family under dual-boundary conditioning contains exactly one nonzero-probability history, since* $r_\alpha = r_\omega = 1$ *for the pure-state boundaries.*

The framework-dependence that Griffiths emphasizes is therefore preserved under dual-boundary conditioning. What final-state conditioning contributes is an admissibility cut within each framework: the family is not only required to be consistent but also required to be compatible with the terminal projector. The Consistent Quantum Causes analysis of the three-box scenario continues to operate, but it operates on a strictly smaller family — one admissible history per framework — rather than on the full unconstrained consistent family.

Figure 3. Three-box paradox under dual-boundary admissibility

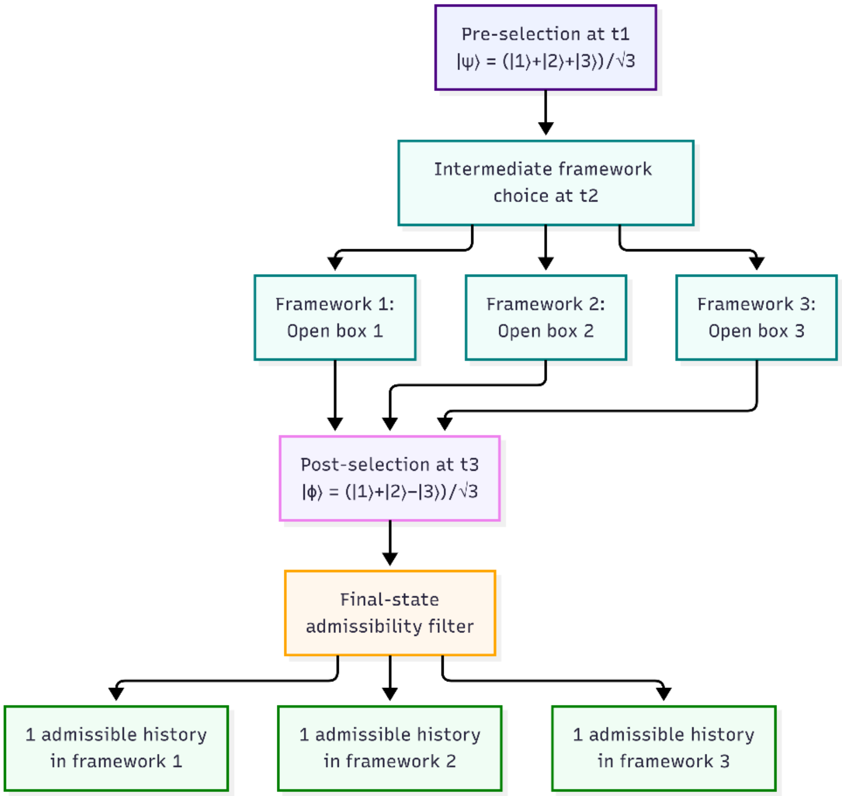

Figure 3. *Three-box paradox represented as a pre-selected, intermediate-measurement, and post-selected scenario. Under dual-boundary admissibility, each compatible framework is pruned to a single nonzero-probability history, consistent with the Hartle–Craig rank-product bound for pure initial and final states.*

5.3. The Same Scenario in Process-Matrix Language

The three-box scenario can be recast in process-matrix language using the pre- and post-selection-to-process-matrix correspondence that has been developed in recent work on post-selected quantum processes. The initial preparation, unitary evolution, and intermediate measurement define operations at distinct "parties" in the process-matrix sense, and the post-selection |ϕ⟩ at t_3 enters as a boundary condition on the process. The Costa–Shrapnel interventionist discovery algorithm [15] applied to the unconditioned process would identify a causal structure connecting the intermediate measurement locus to the final outcome. Applied to the post-selected process, the algorithm operates on a restricted class of process matrices, namely those compatible with |ϕ⟩.

The restriction parallels the one on the Consistent Histories side. Informational pruning does not choose between the two formal languages. It imposes a compatibility condition that restricts the objects each language describes. In the Consistent Histories language, the restriction is on history families. In the process-matrix language, the restriction is on process matrices. Both restrictions are induced by the same terminal projector, and both reduce the effective description length of the analyzed ensemble relative to its unconstrained counterpart.

I do not claim to provide a full technical development of post-selection in the QCM framework; that is a research program in its own right and one that the QCM literature has not yet pursued systematically. What the three-box example illustrates is that the positioning claim has concrete content on both sides of the CH/QCM divide: the same scenario can be analyzed through both lenses, and final-state conditioning exerts a structurally analogous pruning effect on each.

6. An Explainability Stance for Admissible Histories

The preceding sections developed the positioning argument at a formal level. This section introduces a methodological frame that has been useful elsewhere in science and philosophy and that has, I argue, a clarifying application here. The frame is the explainability stance that has become familiar in the literature on explainable artificial intelligence [26,27].

6.1. The Stance, Not the Methods

In the explainable artificial intelligence literature, the stance behind a variety of methods is that predictive success is not sufficient for understanding. A model that predicts outcomes well may still be opaque with respect to the internal structure that generates those outcomes, and various methods, including feature attribution, counterfactual explanation, and mechanistic probing, have been developed to characterize that structure. I emphasize that I am not importing these specific methods into quantum foundations. The methods are calibrated to classical machine learning models and would not transfer straightforwardly. What I am importing is the underlying stance: the insistence that producing the right probabilities is not equivalent to identifying which internal structures are doing the explanatory work.

A related line of inquiry has begun to apply explainability methods to quantum machine learning models [28,29]. That work is distinct in scope from the present paper. Its object is quantum machine learning itself, not quantum foundations. The stance, however, is portable.

6.2. Application to Final-State-Constrained Histories

Applied to the final-state-constrained framework, the explainability stance sharpens a question that is otherwise easy to leave implicit. Given an admissible-history ensemble Ω_ϕ induced by a terminal condition ϕ, not all admissible histories contribute equally to the realization of ϕ. Some histories are such that their admissibility is structurally required, in the sense that the terminal condition cannot be realized without histories of their type. Other histories are admissible but play no structurally necessary role; they are along for the ride. The explainability stance asks which histories are doing the work.

This question has a precise analogue in both Consistent Quantum Causes and Quantum Causal Models. In the Consistent Histories framework, the question corresponds to which projector-defined events within an admissible history carry the causal weight for the outcomes one wishes to explain. In the QCM framework, it corresponds to which loci in the process would, if intervened upon, change the terminal outcome class. Both programs have machinery for addressing these questions within their respective formalisms. The contribution of the explainability stance, combined with the final-state-constrained framework, is to ask these questions about the admissible ensemble rather than about the unconstrained space.

Concretely, this suggests a methodological move. Rather than asking what the causal structure of a quantum process looks like under forward dynamics alone, one asks what that structure looks like when restricted to histories admissible under a specified terminal condition, and one asks which histories within that restricted ensemble are structurally load-bearing. The first question is addressed by the final-state-constrained framework. The second is addressed by the causal machinery of Consistent Quantum Causes or Quantum Causal Models, applied within the admissible ensemble.

6.3. Scope and Limits of the Stance

The explainability stance, as used here, is a methodological frame rather than a technical apparatus. It does not supply new formal tools. It highlights a distinction, already implicit in all three frameworks, between predictive equivalence and structural identification. And it suggests that the positioning argument of Section 4 is not merely an organizational observation but a recommendation about how causal analysis might be conducted within admissible ensembles.

I want to be careful about overreach. The explainability stance does not by itself show that one framework is correct and another is not. It does not establish that final-state conditioning is empirically warranted. It does not resolve the disputes between Griffiths and the Costa–Shrapnel program. What it does is identify a question that each framework can address within the admissible ensemble, and that becomes sharper when the ensemble is filtered by a terminal condition with nontrivial structural complexity.

7. Objections and Replies

Several objections to the positioning argument are foreseeable. This section states them and responds.

7.1. Is the Positioning Merely Rhetorical?

A first objection holds that the positioning argument is rhetorical rather than substantive. It does not prove new theorems, it does not resolve existing disputes, and it does not supply new predictions. It merely rearranges existing programs under a new heading. Why should anyone in quantum foundations take this as progress?

The reply is that positioning papers have a legitimate role when a field contains programs that address adjacent questions without addressing each other. The Consistent Quantum Causes and Quantum Causal Models programs do not cite each other frequently, and when they do the framing is competitive. The final-state-constrained framework has not yet been integrated with either. Identifying a shared assumption, namely unilateral temporal conditioning, and a vantage point at which a third framework intervenes, clarifies what the programs are doing and what they are not doing. This is not a substitute for formal progress. It is preparatory work for such progress. Moreover, the Hartle–Craig rank-product theorem discussed in Section 4.2 supplies a formal anchor that the positioning can be measured against: the claim that final-state conditioning prunes the admissible ensemble is not merely rhetorical, it is quantitatively constrained.

7.2. Does Final-State Conditioning Presuppose What It Purports to Explain?

A second objection concerns the structural complexity of terminal conditions. The positioning argument claims that under computational terminal states, the admissible histories are precisely those that support information-processing organization. One might worry that this is circular: the terminal condition is defined in terms of computational structure, and the admissible histories therefore trivially support that structure. Nothing is explained.

The reply is that the circularity is apparent rather than real, and is addressed at length in [6]. The positioning argument inherits the reply and does not depend on solving the circularity afresh. The argument of [6] is that endpoint classes differ in their informational work. Trivial terminal states impose weak selectivity. Structured terminal states impose stronger selectivity. Computational terminal states impose the strongest selectivity. The hierarchy is a claim about pruning power, not a definitional stipulation. Whether the hierarchy is correct is a separate question from whether the positioning argument presupposes it; and the positioning argument remains informative even under weaker versions of the hierarchy, since the basic observation that unilateral conditioning is shared by both Consistent Quantum Causes and QCM does not depend on any particular endpoint-complexity claim.

7.3. Is Final-State Conditioning Retrocausal?

A third objection is that the final-state-constrained framework is covertly committed to retrocausal mechanisms, and that importing it into a discussion of Consistent Quantum Causes or Quantum Causal Models simply relocates the metaphysical problems. Griffiths' framework is explicitly non-retrocausal. The QCM framework admits indefinite causal order but does not require backward dynamical influence. Introducing final-state conditioning seems to smuggle in retrocausal machinery through the back door.

The reply is that the framework in [6] is retrocausal in boundary architecture and non-interventionist in mechanism. It does not claim that a future state reaches backward to overwrite a fixed past. Histories incompatible with both boundary conditions are excluded from the admissible ensemble rather than revised. This is a global admissibility condition, not a dynamical intervention. In the taxonomy developed by Adlam, Hance, Hossenfelder, and Palmer [10], the distinction between boundary-conditioned and interventionist-retrocausal frameworks is explicit, and informational pruning falls on the boundary-conditioned side. Whether this qualifies as retrocausality in a problematic sense is itself contested in the literature [30], but the positioning argument does not require a stronger commitment than global admissibility.

7.4. How Does the Positioning Differ from Ormrod–Barrett?

A fourth objection observes that Ormrod and Barrett [7] have already synthesized relational quantum mechanics, consistent histories, and quantum causal modelling in a single framework, and asks what the present paper adds beyond relabeling.

The reply is that Ormrod and Barrett select preferred consistent families by an internal causal mechanism, in which dynamic causal structure among systems picks out which projective decompositions avoid interference influence with later measurements. The present framework selects admissible consistent families by an external boundary condition imposed on the history ensemble. These are two different mechanisms for restricting admissibility, and they appeal to different foundational resources. Ormrod and Barrett do not invoke final-state constraints. The present framework does not invoke event relativity. A joint paper comparing the two mechanisms — which is a different paper from either — might be productive, but the present paper is distinct at the level of what does the selecting and why.

7.5. Does the Explainability Stance Do Real Work?

A fifth objection asks whether the explainability stance adds anything beyond relabeling. If the stance amounts to asking which admissible histories are structurally load-bearing, one might object that Consistent Histories analysis and Quantum Causal Models analysis already ask this question within their own frameworks. What has been gained?

The reply is that the stance does not supplant these analyses. It locates them within the admissible ensemble, and it makes explicit the distinction between predictive adequacy and structural identification. The stance is methodological rather than technical, and I have been explicit about that in Section 6. Whether it amounts to more than a useful reframing will depend on what future work does with it. The present paper positions the stance as a frame, not as a theorem.

8. Directions for Further Development

The positioning argument suggests several directions in which further work could be pursued. I do not undertake any of these here; the purpose of this section is to indicate where the argument points.

First, a formal analysis could be conducted of how Consistent Histories consistency conditions interact with final-state admissibility beyond the rank-product bound. In particular, one could ask whether the families of consistent histories compatible with a given terminal condition form a natural subfamily of the unconstrained consistent families, and whether framework-selection considerations within that subfamily differ from those in the unconstrained case. The connection to the Dowker–Kent set-selection problem noted in Section 4.2 suggests that this could yield substantive new material rather than only organizational clarification.

Second, a similar analysis could be conducted for the QCM framework. Process matrices compatible with a final-state constraint form a restricted class, and one could ask how interventionist causal discovery on that class differs from discovery on the unconstrained class. The de Finetti results recently developed for quantum causal structures [18] suggest that this class might have structural regularities worth studying, and the three-box example worked through in Section 5 suggests that the restriction is nontrivial in even simple cases.

Third, the relation between admissible-history pruning and complexity-theoretic measures of agential or computational structure could be sharpened. The framework in [6] defines agency structurally, in terms of selectivity, computational organization, and recursive self-favoring. These properties admit more precise formulation, and the pruning analysis can be developed along quantitative rather than qualitative lines, with effective complexity measures [31] as a natural starting point.

Fourth, empirical development of the sort indicated in [6] could be pursued in dialogue with the causal-model literature. Intervention protocols in laboratory quantum systems [32,33] might be extended to include post-selection structures designed to test endpoint-sensitive conditioning. This is speculative. The point is that the positioning developed in this paper opens avenues for combined analyses that neither the Consistent Quantum Causes program nor the Quantum Causal Models program has thus far pursued.

Fifth, the relationship between informational pruning and the Ormrod–Barrett event-relativity framework deserves direct treatment. Both mechanisms restrict admissibility, but through different selection principles. A paper analyzing how the two frameworks interact when both an external boundary condition and an internal causal-dynamical selection principle are imposed would be a natural extension of both programs.

9. Conclusion

Two programs currently dominate formal work on quantum causation. The Consistent Quantum Causes program, developed by Griffiths, grounds causal reasoning in the Consistent Histories formalism and emphasizes the irreducible role of noncommuting projectors and stochastic time development. The Quantum Causal Models program, developed by Costa, Shrapnel, and collaborators, grounds causal reasoning in the process-matrix formalism and extends classical interventionist causal discovery to the quantum domain. Both programs are sophisticated. Both are technically mature. And they disagree with each other on questions of formal starting point and interpretive emphasis.

They also share an assumption that goes largely unexamined in their own literatures. Each takes quantum causal structure to be fixed by an initial condition together with forward dynamics. Neither engages systematically with dual-boundary conditioning, in which admissibility is determined by an initial state together with a final-state constraint. The final-state-constrained framework developed in prior work treats such conditioning seriously and shows that, under a history-ensemble encoding, terminal-condition structure can induce substantial informational pruning of admissible histories.

I have argued that this framework provides a third vantage point on the Consistent Quantum Causes and Quantum Causal Models programs. It is not a competitor to either program. It does not propose new consistency conditions or new causal-discovery algorithms. What it proposes is that the space of histories over which either program operates can be filtered by a terminal-condition admissibility condition, and that doing so changes what the programs are analyzing. The Hartle–Craig rank-product theorem supplies a quantitative measure of this pruning, and the three-box paradox provides a concrete illustration of how it operates in a scenario that can be analyzed through both the Consistent Histories and the process-matrix lenses.

I have also suggested that an explainability stance, drawn methodologically from explainable artificial intelligence, is useful for asking which admissible histories do the structural work. The stance is a frame, not a theorem. It does not supplant the causal machinery of either program. It locates that machinery within the admissible ensemble and sharpens the question of structural versus merely admissible contribution.

The argument of this paper is positioning rather than construction. It does not resolve the Griffiths–QCM disagreement, and it does not establish the empirical warrant of final-state conditioning. Its more limited aim is to show that informational pruning, as a third vantage, is coherent, non-trivial, and productive of further questions. A constrained history-ensemble approach does not yet establish a unified theory of quantum causation, but it does offer a perspective from which the two leading programs can be re-examined and, potentially, brought into closer conversation.

References

1. R. B. Griffiths, Consistent Quantum Causes, arXiv:2303.13617 [quant-ph] v3 (2024).
2. F. Costa and S. Shrapnel, Quantum causal modelling, New Journal of Physics 18(6): 063032 (2016). doi:10.1088/1367-2630/18/6/063032
3. S. Shrapnel, Discovering Quantum Causal Models, British Journal for the Philosophy of Science 70(1): 1–25 (2019). doi:10.1093/bjps/axx044
4. Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz, Time symmetry in the quantum process of measurement, Physical Review 134(6B): B1410–B1416 (1964). doi:10.1103/PhysRev.134.B1410
5. Y. Aharonov and L. Vaidman, Properties of a quantum system during the time interval between two measurements, Physical Review A 41(1): 11–20 (1990). doi:10.1103/PhysRevA.41.11
6. I. Staley, Final-State Constraints and Informational Pruning in Quantum Histories, International Journal of Quantum Foundations 12: 719–737 (2026).
7. N. Ormrod and J. Barrett, Quantum influences and event relativity, arXiv:2401.18005 [quant-ph] (2024).
8. C. Anastopoulos, Final States in Quantum Cosmology: Cosmic Acceleration as a Quantum Post-Selection Effect, arXiv:2401.07662 [gr-qc] v3 (2024).
9. E. Adlam, Is there causation in fundamental physics? New insights from process matrices and quantum causal modelling, Synthese 201: 152 (2023). doi:10.1007/s11229-023-04160-z
10. E. Adlam, J. R. Hance, S. Hossenfelder, and T. N. Palmer, Taxonomy for physics beyond quantum mechanics, Proceedings of the Royal Society A 480(2294): 20230779 (2024). doi:10.1098/rspa.2023.0779
11. R. B. Griffiths, Consistent histories and the interpretation of quantum mechanics, Journal of Statistical Physics 36(1–2): 219–272 (1984). doi:10.1007/BF01015734
12. M. Gell-Mann and J. B. Hartle, Classical equations for quantum systems, Physical Review D 47(8): 3345–3382 (1993). doi:10.1103/PhysRevD.47.3345
13. P. Strasberg, T. Reinhard, and J. Schindler, First Principles Numerical Demonstration of Emergent Decoherent Histories, Physical Review X 14: 041027 (2024). doi:10.1103/PhysRevX.14.041027
14. R. B. Griffiths, Consistent Quantum Theory, Cambridge University Press, Cambridge (2002).
15. C. Giarmatzi and F. Costa, A quantum causal discovery algorithm, npj Quantum Information 4: 17 (2018). doi:10.1038/s41534-018-0062-6
16. O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, Nature Communications 3: 1092 (2012). doi:10.1038/ncomms2076
17. L. A. Rozema, T. Strömberg, H. Cao, Y. Guo, B.-H. Liu, and P. Walther, Experimental aspects of indefinite causal order in quantum mechanics, Nature Reviews Physics 6(8): 483–499 (2024). doi:10.1038/s42254-024-00739-8
18. F. Costa, J. Barrett, and S. Shrapnel, A de Finetti theorem for quantum causal structures, Quantum 9: 1628 (2025). doi:10.22331/q-2025-02-11-1628

19. J. Barrett, I. Friend, and A. Kissinger, Quantum and Classical Markovian Graphical Causal Models and Their Identification, in 33rd EACSL Annual Conference on Computer Science Logic (CSL 2025), LIPIcs 326: 48:1–48:23, Schloss Dagstuhl (2025). doi:10.4230/LIPIcs.CSL.2025.48

20. J. B. Hartle, Arrows of Time and Initial and Final Conditions in the Quantum Mechanics of Closed Systems Like the Universe, arXiv:2002.07093 [gr-qc] (2020).

21. P. W. Evans, Quantum Causal Models, Faithfulness, and Retrocausality, British Journal for the Philosophy of Science 69(3): 745–774 (2018). doi:10.1093/bjps/axw037

22. H. Price and K. Wharton, Dispelling the Quantum Spooks: A Clue That Einstein Missed?, in C. Bouton and P. Huneman (eds.), Time of Nature and the Nature of Time, Boston Studies in the Philosophy and History of Science 326, Ch. 7, Springer, Cham (2017). doi:10.1007/978-3-319-53725-2_7

23. D. A. Craig, Geometry of Consistency: Decohering Histories in Generalized Quantum Theory, arXiv:quant-ph/9704031 (1997).

24. Y. Aharonov and L. Vaidman, Complete description of a quantum system at a given time, Journal of Physics A 24(10): 2315–2328 (1991). doi:10.1088/0305-4470/24/10/018

25. R. E. Kastner, The Three-Box "Paradox" and Other Reasons to Reject the Counterfactual Usage of the ABL Rule, Foundations of Physics 29(6): 851–863 (1999). doi:10.1023/A:1018865010735

26. M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you?": Explaining the predictions of any classifier, in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1135–1144 (2016). doi:10.1145/2939672.2939778

27. C. Molnar, Interpretable Machine Learning: A Guide for Making Black Box Models Explainable, 2nd ed., Leanpub (2022).

28. L. Pira and C. Ferrie, On the interpretability of quantum neural networks, Quantum Machine Intelligence 6: 52 (2024). doi:10.1007/s42484-024-00191-y

29. R. Heese, T. Gerlach, S. Mücke, S. Müller, M. Jakobs, and N. Piatkowski, Explaining quantum circuits with Shapley values: towards explainable quantum machine learning, Quantum Machine Intelligence 7: 27 (2025). doi:10.1007/s42484-025-00254-8

30. M. S. Leifer and M. F. Pusey, Is a time symmetric interpretation of quantum theory possible without retrocausality?, Proceedings of the Royal Society A 473: 20160607 (2017). doi:10.1098/rspa.2016.0607

31. M. Gell-Mann and S. Lloyd, Effective complexity, in Nonextensive Entropy: Interdisciplinary Applications, eds. M. Gell-Mann and C. Tsallis, Oxford University Press, Oxford, pp. 387–398 (2004).

32. K. Ried, M. Agnew, L. Vermeyden, D. Janzing, R. W. Spekkens, and K. J. Resch, A quantum advantage for inferring causal structure, Nature Physics 11: 414–420 (2015). doi:10.1038/nphys3266

33. G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Quantum computations without definite causal structure, Physical Review A 88: 022318 (2013). doi:10.1103/PhysRevA.88.022318