Quantum Finite Automata and Stinespring Dilation in Lean

Bjørn Kjos-Hanssen and Swarnalakshmi Lakshmanan

May 20, 2026

Abstract

Wiedijk has a famous list of 100 mathematical theorems whose formalization is sought in a proof assistant. David and Palsberg recently announced a similar list restricted to quantum information theory. The latter subject has been intensively formalized by Meiburg. In this article our aim is twofold: (1) to answer the challenge of David and Palsberg by supplying a solution to a finite-dimensional version of their problem #54: formalize the Stinespring dilation theorem; (2) answer an unpublished challenge by Meiburg to formalize quantum finite automata.

1 Introduction

In this project we formalize parts of the article *Unbounded length minimal synchronizing words for quantum channels over qutrits* [5] and notions in quantum information theory.

In particular we formalize the notion of a measure-once one-way general quantum finite automaton (MO-1gQFA) introduced by Li, Qiu, Zou, Li, Wu, Mateus [6] in 2012. As an alternative to "MO-1gQFA" we refer to these simply as Kraus operator automata. In particular we formalize cut-point languages as in [9] who call the automata "RT-QFA" (real-time QFA).

The paper addresses two needs or request by other researchers:

Marco David and Jens Palsberg posted a list of Top 100 Quantum Theorems on March 11, 2026, at `https://marcodavid.net/top100`. We include a matrix version of #54 on their list, the Stinespring Dilation Theorem.

Meiburg, Lessa and Soldati [8] formalized a result from Hayashi and Yamasaki [2] in the proof assistant Lean, giving a corrected proof of the Generalized Quantum Stein's Lemma. The latter is an important result in operational quantum information theory going back to Hiai and Petz [3].

In the course of formalizing this monumental result they build a library `Lean-QuantumInfo` [7] containing large parts of quantum information theory.

One notable missing and sought-after ingredient in Lean-Quantuminfo is the theory of quantum finite automata. In the project `Kraus` [4] we partially remedy this. In particular, we formalize the language of a measure-once real-time QFA [9, 6] in terms of Mathlib's probability mass functions (PMF) and compute some examples related to the Kraus operators in this paper and in Grudka et al.'s paper [1]. Most results are proved for the two fields $\mathbb{R}$ and $\mathbb{C}$, encapsulated in Lean's mathematics library Mathlib4 by the `RCLike` typeclass.

The list collects 100 important theorems in quantum physics, quantum information and quantum computing. It was inspired by the list of the top 100 mathematical theorems of which Freek Wiedijk tracks the formalization progress (Figure 1).

1. The Spectral Theorem
2. Eigenstates in an infinite square well
3. The Quantum Harmonic Oscillator
4. Ehrenfest's Theorem
5. Quantum tunneling through a finite potential barrier
6. Conservation of the probability (4-)current
7. Eigenstates of a -distribution well
8. Heisenberg Uncertainty Relation
9. Baker–Campbell–Hausdorff Formula
10. Trotter Product Formula (Lie–Trotter–Kato)
11. Simultaneous diagonalization of commuting observables
12. Noether's Theorem
13. Wigner's Theorem
14. Stone–von Neumann Theorem
15. Spectrum of the hydrogen atom (Coulomb potential)
16. Addition of angular momentum (e.g. Clebsch–Gordan)
17. Wigner–Eckart Theorem
18. Power-series spectrum of a perturbed time-independent Hamiltonian
19. Dyson series for time-dependent perturbation theory
20. Correctness of the Rayleigh–Ritz method
21. Correctness of the Born–Oppenheimer approximation
22. The Adiabatic Theorem
23. Berry phase (Geometric phase)
24. Superdense coding
25. Quantum teleportation with shared entanglement
26. No-Cloning Theorem / No-Broadcast Theorem / No-Teleportation Theorem
27. No-Communication Theorem
28. No-Deleting Theorem
29. Schrödinger–HJW Theorem (purification of mixed states)
30. Schmidt decomposition for Hilbert spaces
31. Bell's Theorem (QM violates CHSH inequality)
32. Monogamy of entanglement
33. Tsirelson bound
34. Correctness of the Deutsch–Jozsa algorithm
35. Correctness of Shor's algorithm
36. Correctness of Grover's algorithm
37. The Bennett–Bernstein–Brassard–Vazirani Theorem (BBBV)
38. Quantum key distribution by BB84
39. Impossibility of quantum bit commitment (Mayers–Lo–Chau)
40. Solovay–Kitaev Theorem
41. Universality of sets of two-qubit quantum gates
42. Universality of the Toffoli and the Hadamard
43. Five two-qubit gates are necessary and sufficient to implement a Toffoli gate
44. Six two-qubit linear nearest-neighbor gates are necessary and sufficient to implement a Toffoli gate
45. 3n CX gates are necessary to implement a multi-control Toffoli gate with n-1 controls
46. Adiabatic and gate-based quantum computing are equivalent up to polynomial factors
47. The Threshold Theorem (quantum fault-tolerance)
48. Petz recovery map
49. Choi's Theorem on completely positive maps
50. Choi–Jamiołkowski isomorphism (channel-state duality)
51. Gelfand–Naimark–Segal construction
52. Gelfand–Naimark Theorem
53. Krein–Milman Theorem
54. **Stinespring's Factorization Theorem / Naimark's Dilation Theorem**
55. Continuous Functional Calculus
56. Gleason's Theorem
57. Holevo's Theorem
58. Pusey–Barrett–Rudolph Theorem
59. Kochen–Specker Theorem
60. The Spectral Theorem for PVMs
61. Strong subadditivity of quantum entropy
62. Entanglement-assisted classical capacity of quantum channels
63. Quantum state discrimination for two states (Ivanović–Dieks–Peres Limit)
64. The Lloyd–Shor–Devetak Theorem for quantum channel capacity
65. Eastin–Knill Theorem
66. Gottesman–Knill Theorem
67. Magic state distillation
68. Correctness of distillation of Bell pairs
69. 1D gapped Hamiltonians have area-law ground states
70. The Lieb–Robinson bound
71. Onsager's solution to the 2D Ising model
72. MIP* = RE
73. BQP $\subseteq$ PP
74. QIP = PSPACE
75. Quantum Stein Lemma
76. Quantum 3-SAT is QMA1-complete
77. Computing the Jones polynomial at roots of unity is BQP-hard
78. Uniqueness of 4-dimensional representations of the Clifford algebra (up to unitary equivalence)
79. Spin-statistics Theorem
80. Reeh–Schlieder Theorem
81. Wick's Theorem
82. Elitzur's Theorem
83. 2D TQFTs are equivalent to Frobenius algebras
84. Chern–Simons theory can compute the Jones polynomial
85. The CFT central charge is its entanglement entropy
86. Osterwalder–Schrader Reconstruction Theorem
87. 4D Gaussian free field theory satisfies the Osterwalder-Schrader axioms
88. Bisognano–Wichmann Theorem
89. Wood–Spekkens–Bell Theorem
90. Correctness of block encoding
91. Correctness of Hamiltonian simulation by Trotterization
92. Correctness of Hamiltonian simulation by Linear Combination of Unitaries
93. Correctness of Hamiltonian simulation by Qubitization
94. Qubit reuse is NP-complete
95. Correctness of a quantum circuit optimizer
96. Correctness of the toric code
97. Correctness of entanglement-assisted stabilizer codes
98. The Knill–Laflamme conditions
99. Correctness of approximate circuit synthesis
100. Correctness of Simon's algorithm

Figure 1: The Top 100 Quantum Theorems.

2 Formalizing QFAs

The following definitions and lemmas were generated by the Gemini 3 agent when asked to translate our Lean code from [4] into LaTeX and English.

First, operator "sandwiches".

Definition 1 (Kraus Map)**.** Let R be a star-ring with an additive commutative monoid structure. Given a family of matrices $\{K_i\}_{i=1}^{r} \in M_q(R)$, the **Kraus application map** $\Phi : M_q(R) \to M_q(R)$ is defined as the completely positive map:

$$\Phi(\rho) = \sum_{i=0}^{r-1} K_i \rho K_i^\dagger \tag{1}$$

where $K_i^\dagger$ denotes the conjugate transpose of the i-th Kraus operator K_i.

Lemma 2 (Kraus Map Preserves PSD)**.** *Let $\rho \in M_q(R)$ be a positive semi-definite matrix ($\rho \geq 0$). Then for any Kraus operators $\{K_i\}$, the result of the map*

$$\Phi(\rho) = \sum_i K_i \rho K_i^\dagger$$

is also positive semi-definite.

Proof. For each i, since $\rho \geq 0$, the conjugation $K_i \rho K_i^\dagger$ is positive semi-definite. Because the sum of positive semi-definite matrices is also positive semi-definite, it follows that $\sum_i K_i \rho K_i^\dagger \geq 0$. □

Quantum State and Operation Definitions

Definition 3 (Quantum Channel / TP Condition)**.** A set of Kraus operators $\{K_i\}_{i=1}^{r} \subset M_q(\mathbb{C})$ defines a **quantum channel** (or a trace-preserving map) if they satisfy the completeness relation:

$$\sum_{i=1}^{r} K_i^\dagger K_i = I_q \tag{2}$$

where I_q is the $q \times q$ identity matrix.

Definition 4 (Quantum Operation)**.** A set of Kraus operators $\{K_i\}_{i=1}^{r} \subset M_q(\mathbb{C})$ defines a **quantum operation** if they satisfy:

$$\sum_{i=1}^{r} K_i^\dagger K_i \leq I_q \tag{3}$$

In this context, $\leq$ denotes the Loewner order, meaning $I_q - \sum K_i^\dagger K_i$ is a positive semi-definite matrix.

Definition 5 (Density Matrix)**.** A matrix $\rho \in M_q(\mathbb{C})$ is a **density matrix** (or a normalized state) if it satisfies:

1. $\rho \geq 0$ (Positive semi-definite)
2. $\mathrm{Tr}(\rho) = 1$ (Normalization)

Definition 6 (Sub-normalized Density Matrix)**.** A matrix $\rho \in M_q(\mathbb{C})$ is a **sub-normalized density matrix** if it satisfies:

1. $\rho \geq 0$ (Positive semi-definite)
2. $\mathrm{Tr}(\rho) \leq 1$

These often represent states after a non-deterministic measurement.

Lemma 7 (Trace Preservation of Quantum Channels)**.** *Let $\Phi : M_q(\mathbb{C}) \to M_q(\mathbb{C})$ be a map defined by a set of Kraus operators $\{K_i\}_{i=1}^r$ satisfying the quantum channel condition $\sum_i K_i^\dagger K_i = I_q$. Then for any $\rho \in M_q(\mathbb{C})$, the trace is preserved:*

$$\mathrm{Tr}(\Phi(\rho)) = \mathrm{Tr}(\rho) \tag{4}$$

Proof. By the linearity and cyclic property of the trace, $\mathrm{Tr}(\sum_i K_i \rho K_i^\dagger) = \sum_i \mathrm{Tr}(K_i^\dagger K_i \rho) = \mathrm{Tr}((\sum_i K_i^\dagger K_i)\rho) = \mathrm{Tr}(I_q \rho) = \mathrm{Tr}(\rho)$. □

Proposition 8 (Trace-Preserving Property)**.** *Let Φ be a quantum channel defined by Kraus operators $\{K_i\}$ such that $\sum_i K_i^\dagger K_i = I$. If ρ is a normalized matrix with $\mathrm{Tr}(\rho) = 1$, then the output of the channel is also normalized:*

$$\mathrm{Tr}(\Phi(\rho)) = 1 \tag{5}$$

Definition 9 (Quantum Channel Map on Density Matrices)**.** Let $\mathcal{D}_q$ denote the set of $q \times q$ density matrices (positive semi-definite matrices with unit trace). A quantum channel Φ induces a well-defined map:

$$\Phi : \mathcal{D}_q \to \mathcal{D}_q \tag{6}$$

mapping $\rho \mapsto \sum_i K_i \rho K_i^\dagger$. This map is well-defined because it preserves both the positive semi-definite property and the unit trace of the input state.

Definition 10 (Quantum Word Map)**.** Let Σ be an alphabet and $\mathcal{K} : \Sigma \to (\mathrm{Fin}\ r \to M_q(R))$ be a function mapping each symbol to a Kraus family. For a word w of length n, the iterated application map $\delta^* : (\mathrm{Fin}\ n \to \Sigma) \to M_q(R) \to M_q(R)$ is defined recursively by:

$$\delta^*(w, \rho) = \begin{cases} \rho & \text{if } n = 0 \\ \Phi_{w(n-1)}\left(\delta^*(\mathrm{init}(w), \rho)\right) & \text{if } n > 0 \end{cases} \tag{7}$$

where $\Phi_a(\sigma) = \sum_{i \in \mathrm{Fin}\ r} K_{a,i} \sigma K_{a,i}^\dagger$ is the completely positive map associated with the symbol $a \in \Sigma$, and $\mathrm{init}(w)$ denotes the prefix of the word of length $n-1$.

Theorem 11 (Stability of Quantum Word Processing)**.** *Let Σ be an alphabet where each symbol $a \in \Sigma$ is associated with a quantum channel Φ_a (satisfying $\sum_i K_{a,i}^\dagger K_{a,i} = I$). For any word $w \in \Sigma^n$ and any initial density matrix ρ, the state after processing the word, $\rho_w = \delta^*(w, \rho)$, satisfies:*

1. *$\rho_w \geq 0$ (Positive Semi-definiteness)*
2. *$\mathrm{Tr}(\rho_w) = 1$ (Trace Preservation)*

In other words, $\delta^*(w, \cdot)$ is a well-defined map from $\mathcal{D}_q$ to $\mathcal{D}_q$.

Proof. By induction on the length of the word n:

- **Base case ($n = 0$):** $\delta^*(\epsilon, \rho) = \rho$, which is a density matrix by hypothesis.
- **Inductive step:** Assume $\rho_{w'}$ is a density matrix for a word of length m. For a word $w = w'a$, $\rho_w = \Phi_a(\rho_{w'})$. Since Φ_a is a quantum channel, it preserves both the PSD property and the unit trace of $\rho_{w'}$.

□

Proceeding this way, eventually we get to defining positive operator valued measure in terms of a PMF (probability mass function) where the probability of outcome e_i is the trace of the product of the density matrix and the corresponding pure state:

Definition 12 (Computational Basis Measurement PMF)**.** Let $\rho \in M_k(\mathbb{C})$ be a density matrix ($\rho \geq 0$ and $\mathrm{Tr}(\rho) = 1$). Let $\{\mathbf{e}_i\}_{i=1}^k$ be the standard basis vectors in $\mathbb{C}^k$, where $\mathbf{e}_i$ has a 1 at index i and 0 elsewhere. The **pure state projection** associated with outcome i is defined as:

$$P_i = \mathbf{e}_i \mathbf{e}_i^\top \tag{8}$$

The resulting probability mass function (PMF) over the set of outcomes $\{1, \ldots, k\}$ is defined by the assignment:

$$\mathrm{Prob}(i) = \mathrm{Re}(\mathrm{Tr}(P_i \rho)) \tag{9}$$

Next we turn this into a Bernoulli probability measure by declaring that a specific e_{acc} is the accepted subspace:

Definition 13 (Binary Acceptance PMF)**.** Let $\rho \in M_k(\mathbb{R})$ be a density matrix, and let $\mathrm{acc} \in \{1, \ldots, k\}$ be the designated acceptance index. The probability distribution over the outcomes $\{0, 1\}$ (Reject, Accept) is given by the PMF:

$$\Pr(i) = \begin{cases} \mathrm{Tr}(P_{\mathrm{acc}} \rho) & \text{if } i = 1 \text{ (Accept)} \\ \mathrm{Tr}((I - P_{\mathrm{acc}})\rho) & \text{if } i = 0 \text{ (Reject)} \end{cases} \tag{10}$$

where $P_{\mathrm{acc}} = \mathbf{e}_{\mathrm{acc}} \mathbf{e}_{\mathrm{acc}}^\top$ is the projection onto the accepting basis state.

Finally we get to the measure-once language accepted by our quantum finite automaton, with cutpoint 1/2:

Definition 14 (MO-QFA Language Acceptance)**.** Let $\mathcal{A} = (\Sigma, \mathcal{K}, \rho_0, P_{\mathrm{acc}})$ be a Measure-Once Quantum Finite Automaton. The **language accepted** by $\mathcal{A}$ with cutpoint $\lambda = \frac{1}{2}$ is the set of words:

$$L(\mathcal{A}) = \left\{ w \in \Sigma^* \mid \mathrm{Tr}(P_{\mathrm{acc}}\, \delta^*(w, \rho_0)) > \frac{1}{2} \right\} \tag{11}$$

where:

- $\delta^*(w, \rho_0)$ is the state after processing word w.
- $P_{\mathrm{acc}} = \mathbf{e}_{\mathrm{acc}} \mathbf{e}_{\mathrm{acc}}^\top$ is the projection onto the accepting state.
- The index 1 in the probability mass function represents the "Accept" outcome.

Lemma 15. *A probability mass function on two orthogonal projections $P, P^\perp$, with measure equal to the trace of the density matrix ρ times the projection.*

Definition 16. Transition function δ^* corresponding to a word over an alphabet, where each symbol is mapped to a completely positive map in Kraus form, of rank at most r.

Some other formal lemmas, that we only describe stenographically, include:

Lemma 17. *A projection-valued measure, with measure equal to the trace of the density matrix ρ times the projection.*

Lemma 18. *If A and B are PSD then AB has nonnegative trace.*

Lemma 19. *If P is a projection and ρ is PSD then the trace of $P\rho$ is nonnegative.*

Definition 20. Projection-valued measure.

Definition 21. Quantum channel.

Lemma 22. *Kraus operators preserve the PSD property.*

Lemma 23. *An automaton based on a quantum channel maps density matrices to density matrices while reading a single letter.*

Lemma 24. *An automaton based on a quantum channel maps density matrices to density matrices while reading a word.*

Lemma 25. *A basis state density matrix has trace one.*

Lemma 26. *A pure state is PSD.*

3 Stinespring dilation

We now switch gears to consider the Stinespring dilation, in matrix form.

Stinespring Dilation and Partial Trace

Definition 27 (Partial Trace)**.** Let $\rho \in M_{mn}(\mathbb{C})$ be a density matrix on a composite system $\mathcal{H}_A \otimes \mathcal{H}_B$. The **partial trace** over the right subsystem ($\mathcal{H}_B$) is the map $\mathrm{Tr}_B : M_{mn}(\mathbb{C}) \to M_m(\mathbb{C})$ defined by the components:

$$[\mathrm{Tr}_B(\rho)]_{ij} = \sum_{k=1}^{n} \rho_{(i,k),(j,k)} \tag{12}$$

Definition 28 (Stinespring Operator)**.** Given a Kraus family $\{K_i\}_{i=1}^r \subset M_m(\mathbb{C})$, the **Stinespring operator** $V \in M_{mr\times m}(\mathbb{C})$ is defined as:

$$V = \sum_{i=1}^{r} K_i \otimes \mathbf{e}_i \tag{13}$$

where $\mathbf{e}_i$ are the standard basis vectors of the environment space $\mathbb{C}^r$.

Definition 29 (Stinespring Dilation)**.** The **Stinespring dilation** of a state ρ relative to the operator V is the matrix in the enlarged space:

$$\mathrm{Stinespring}(\rho, V) = V\rho V^\dagger \tag{14}$$

It follows that the original Kraus map is recovered by $\Phi(\rho) = \mathrm{Tr}_B(V\rho V^\dagger)$.

Here is the main result towards answering Question #54 of David and Palsberg.

Theorem 30 (Equivalence of Kraus and Stinespring Forms)**.** *Let $\{K_i\}_{i=1}^r \subset M_m(\mathbb{C})$ be a family of Kraus operators and $\rho \in M_m(\mathbb{C})$ be a density matrix. Let $V = \sum_i K_i \otimes \mathbf{e}_i$ be the Stinespring operator. Then:*

$$\mathrm{Tr}_B(V\rho V^\dagger) = \sum_{i=1}^{r} K_i \rho K_i^\dagger \tag{15}$$

where Tr_B denotes the partial trace over the environment space $\mathbb{C}^r$.

Proof. By the definition of the Stinespring operator V, we have:

$$V\rho V^\dagger = \left(\sum_i K_i \otimes \mathbf{e}_i\right) \rho \left(\sum_j K_j^\dagger \otimes \mathbf{e}_j^\top\right) = \sum_{i,j} (K_i \rho K_j^\dagger) \otimes (\mathbf{e}_i \mathbf{e}_j^\top)$$

Applying the partial trace Tr_B to each term in the sum:

$$\mathrm{Tr}_B((K_i \rho K_j^\dagger) \otimes (\mathbf{e}_i \mathbf{e}_j^\top)) = (K_i \rho K_j^\dagger)\,\mathrm{Tr}(\mathbf{e}_i \mathbf{e}_j^\top)$$

Since $\mathrm{Tr}(\mathbf{e}_i \mathbf{e}_j^\top) = \delta_{ij}$ (the Kronecker delta), the double sum collapses to:

$$\sum_i K_i \rho K_i^\dagger$$

which is exactly the Kraus representation $\Phi(\rho)$. □

We end with a sophisticated construction addressing a common problem in quantum information: if you have a Quantum Operation (which is trace-decreasing, $\sum K_i^\dagger K_i \leq I$), how do you "complete" it into a full Quantum Channel (which is trace-preserving, $\sum K_j^\dagger K_j = I$)?

Definition 31 (Kraus Completion)**.** Let $\{K_i\}_{i=1}^r \subset M_m(\mathbb{C})$ be a family of operators satisfying the trace-non-increasing (TNI) condition $\sum_i K_i^\dagger K_i \leq I$. The **orthogonal CPTP completion** is the operator $V_{comp} \in M_{m(r+1)\times m}(\mathbb{C})$ defined block-wise as:

$$V_{comp} = \begin{pmatrix} K_1 \\ \vdots \\ K_r \\ \Delta \end{pmatrix} \quad \text{where} \quad \Delta = \sqrt{I - \sum_{i=1}^{r} K_i^\dagger K_i} \tag{16}$$

In Lean notation, this is represented by branching on the index $x.2$: using the Stinespring operator for the first r indices and the deficit root Δ for the last index.

Theorem 32 (Isometry Property)**.** *If the initial operators satisfy $\sum_i K_i^\dagger K_i \leq I$, then the completion V_{comp} is an isometry, i.e.,*

$$V_{comp}^\dagger V_{comp} = I \tag{17}$$

4 Conclusion

We have chosen a somewhat concrete approach to formalization of quantum information theory using matrices. This has the advantage that it is reasonably feasible to construct examples (such as those used to obtain synchronizing words in [5]). On the other hand, for eventual inclusion in Lean's Mathlib, a higher level of generality is needed. A natural next step for the Stinespring dilation is to formalize the unitary extension, also known as "going to the church of the larger Hilbert space". For quantum finite automata there is a lot that can be done including formalizing a version of quantum automatic complexity.

Acknowledgments

This work was partially supported by grants from the Simons Foundation (#704836 to Bjørn Kjos-Hanssen) and Majesco Inc. (University of Hawai‘i Foundation Account #129-4770-4).

References

[1] Andrzej Grudka, Marcin Karczewski, Paweł Kurzyński, Jedrzej Stempin, Jan Wójcik, and Antoni Wójcik. Quantum synchronizing words: Resetting and preparing qutrit states, 2025.

[2] Masahito Hayashi and Hayata Yamasaki. Generalized quantum Stein's lemma and second law of quantum resource theories, 2025.

[3] Fumio Hiai and Dénes Petz. The proper formula for relative entropy and its asymptotics in quantum probability. *Communications in Mathematical Physics*, 143(1):99–114, 1991.

[4] Bjørn Kjos-Hanssen. Kraus: formalization of quantum finite automata, 2026. Project website (https://bjoernkjoshanssen.github.io/kraus/).

[5] Bjørn Kjos-Hanssen and Swarnalakshmi Lakshmanan. Unbounded length minimal synchronizing words for quantum channels over qutrits, 2026.

[6] Lvzhou Li, Daowen Qiu, Xiangfu Zou, Lvjun Li, Lihua Wu, and Paulo Mateus. Characterizations of one-way general quantum finite automata. *Theoret. Comput. Sci.*, 419:73–91, 2012.

[7] Alex Meiburg. Quantum Information in Lean. https://github.com/Meiburg/Lean-QuantumInfo, 2024.

[8] Alex Meiburg, Leonardo A. Lessa, and Rodolfo R. Soldati. A formalization of the generalized quantum Stein's lemma in Lean, 2025.

[9] Abuzer Yakaryılmaz and A. C. Cem Say. Unbounded-error quantum computation with small space bounds. *Inform. and Comput.*, 209(6):873–892, 2011.