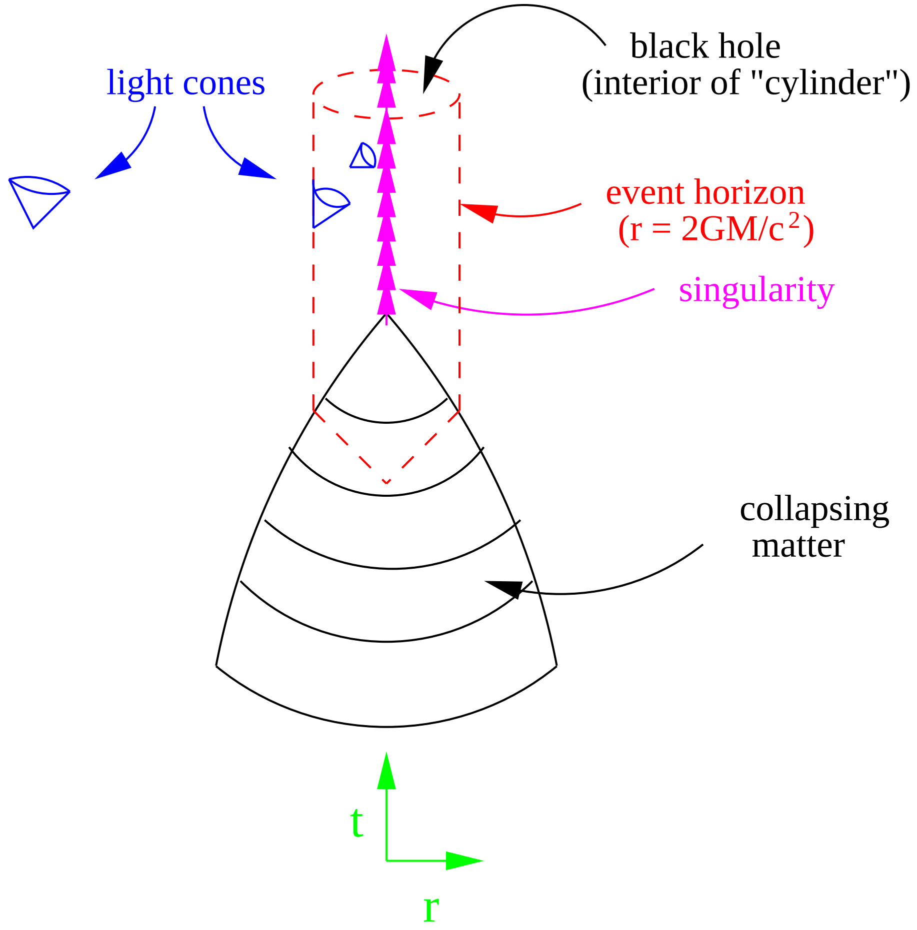
Information Loss

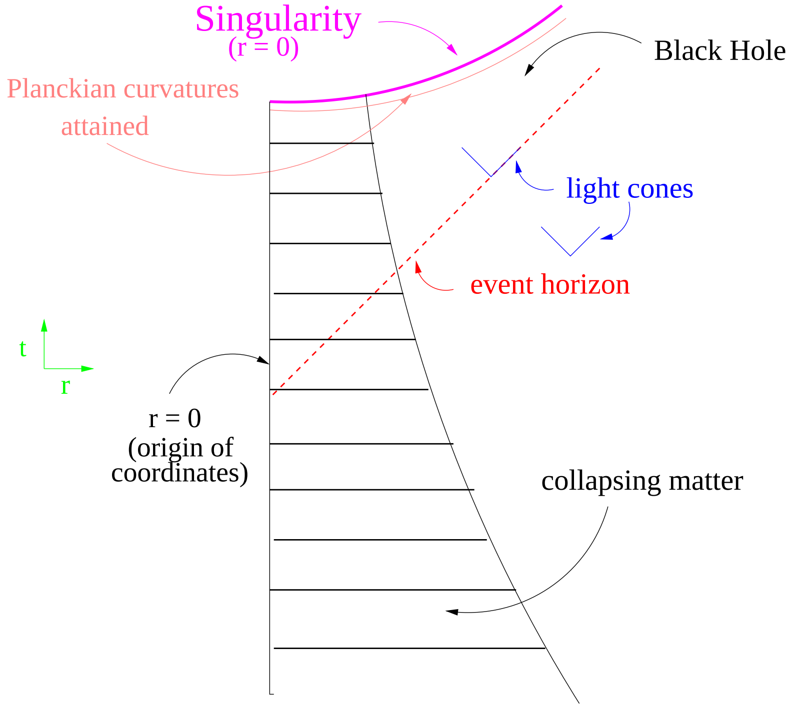
Robert M. Wald

[Reference: W.G. Unruh and R.M. Wald, arXiv:1703.02140, to appear in ROPP.]

Spacetime Diagram of Gravitational Collapse



Spacetime Diagram of Gravitational Collapse with Angular Directions Suppressed and Light Cones “Straightened Out”



Particle Creation by Black Holes

Black holes are perfect black bodies! As a result of particle creation effects in quantum field theory, a distant observer will see an exactly thermal flux of all species of particles appearing to emanate from the black hole. The temperature of this radiation is

$$kT = \frac{\hbar\kappa}{2\pi} \, .$$

For a Schwarzshild black hole ($J = Q = 0$) we have $\kappa = c^3/4GM$, so

$$T \sim 10^{-7} \frac{M_\odot}{M} \, .$$

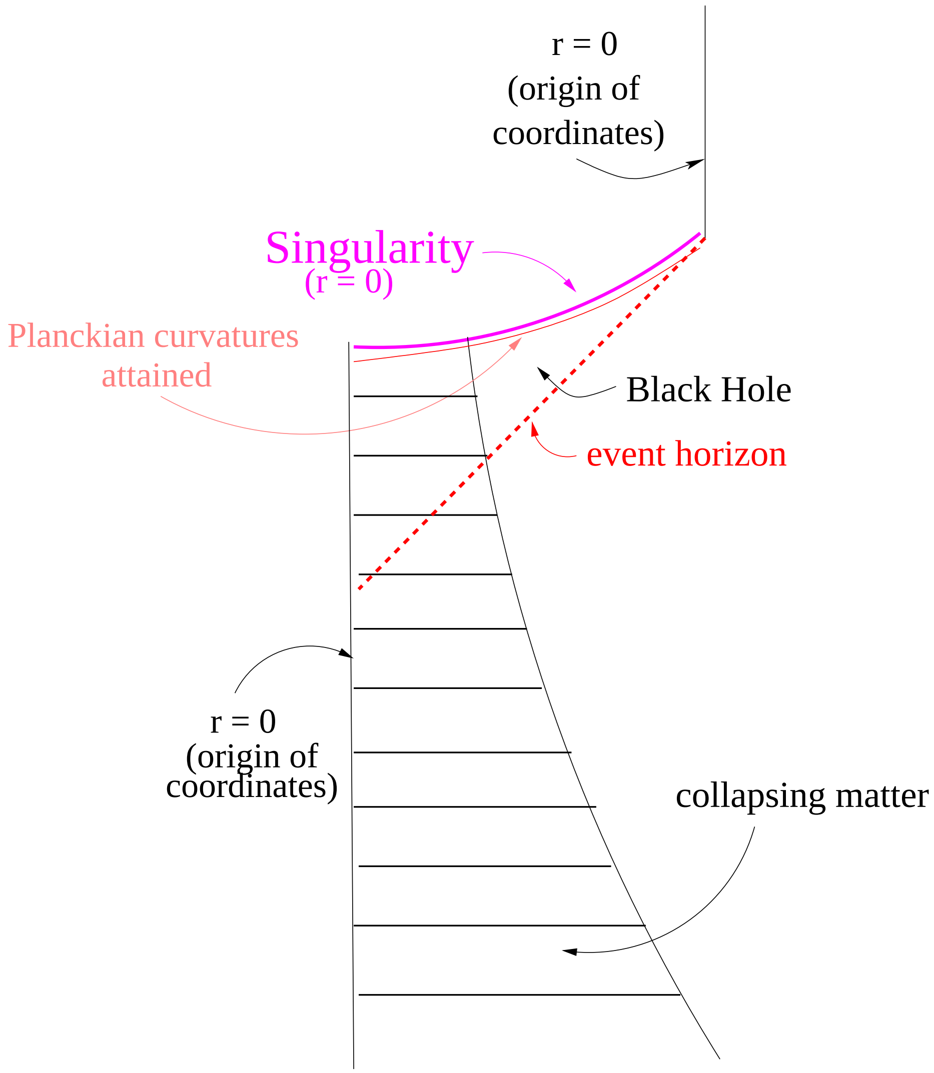
The mass loss of a black hole due to this process is

$$\frac{dM}{dt} \sim AT^4 \propto M^2 \frac{1}{M^4} = \frac{1}{M^2}\,.$$

Thus, an isolated black hole should "evaporate" completely in a time

$$\tau \sim 10^{73} (\frac{M}{M_\odot})^3 \mathrm{sec}\,.$$

Spacetime Diagram of Evaporating Black Hole



Quantum Entanglement

If a quantum system consists of two subsystems, described by Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, then the joint system is described by the Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$. In addition to simple product states $|\Psi_1\rangle \otimes |\Psi_2\rangle$, the Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$ contains linear combinations of such product states that cannot be re-expressed as a simple product. If the state of the joint system is not a simple product, the subsystems are said to be *entangled* and the state of each subsystem is said to be *mixed*. Interactions between subsystems generically result in entanglement.

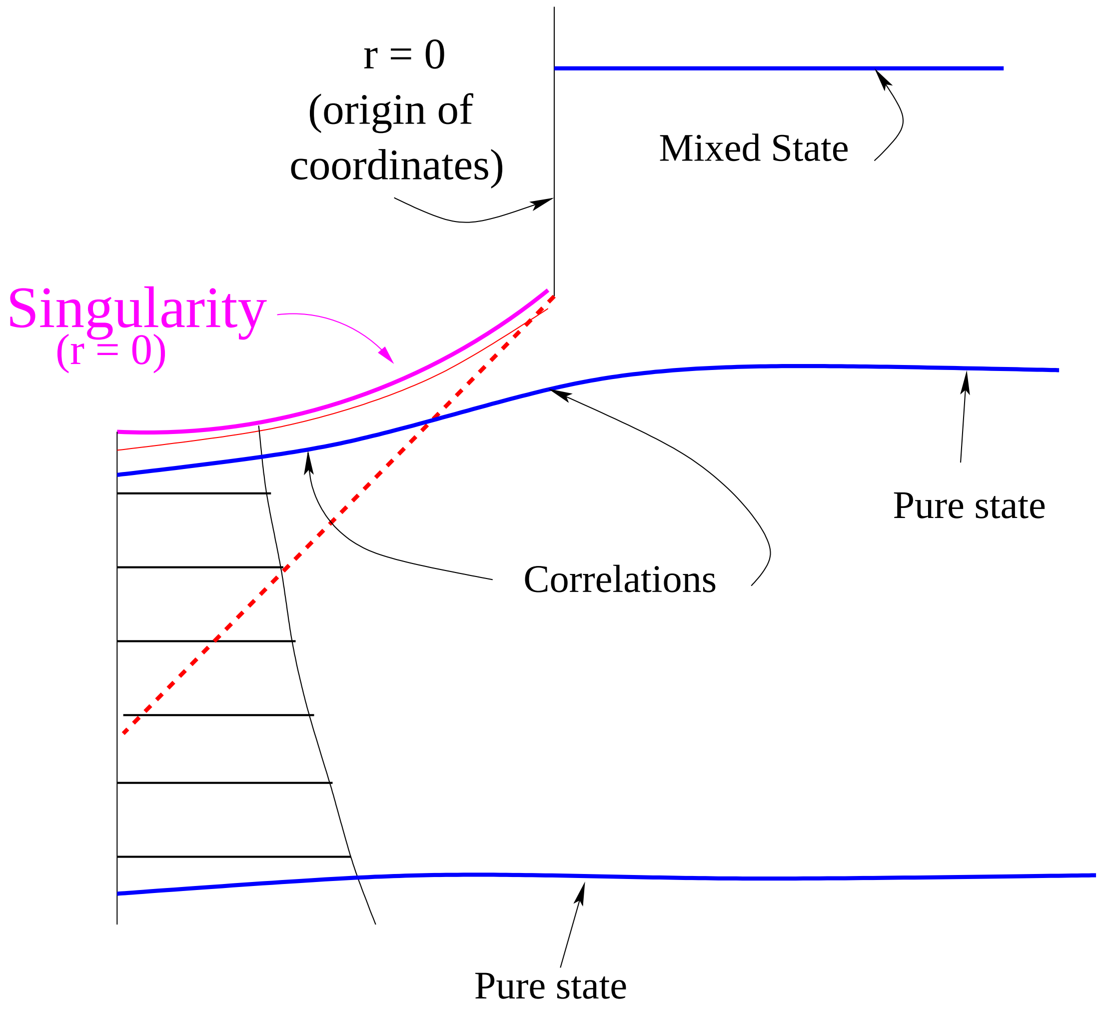
Entanglement is a ubiquitous feature of quantum field theory. At small spacelike separations, a quantum field is always strongly entangled with itself, as illustrated by the following formula for a massless KG field in Minkowski spacetime:

$$\langle 0|\phi(x)\phi(y)|0\rangle = \frac{1}{4\pi^2}\frac{1}{\sigma(x,y)}$$

If there were no entanglement, we would have $\langle 0|\phi(x)\phi(y)|0\rangle = \langle 0|\phi(x)|0\rangle\langle 0|\phi(y)|0\rangle = 0$.

Information Loss

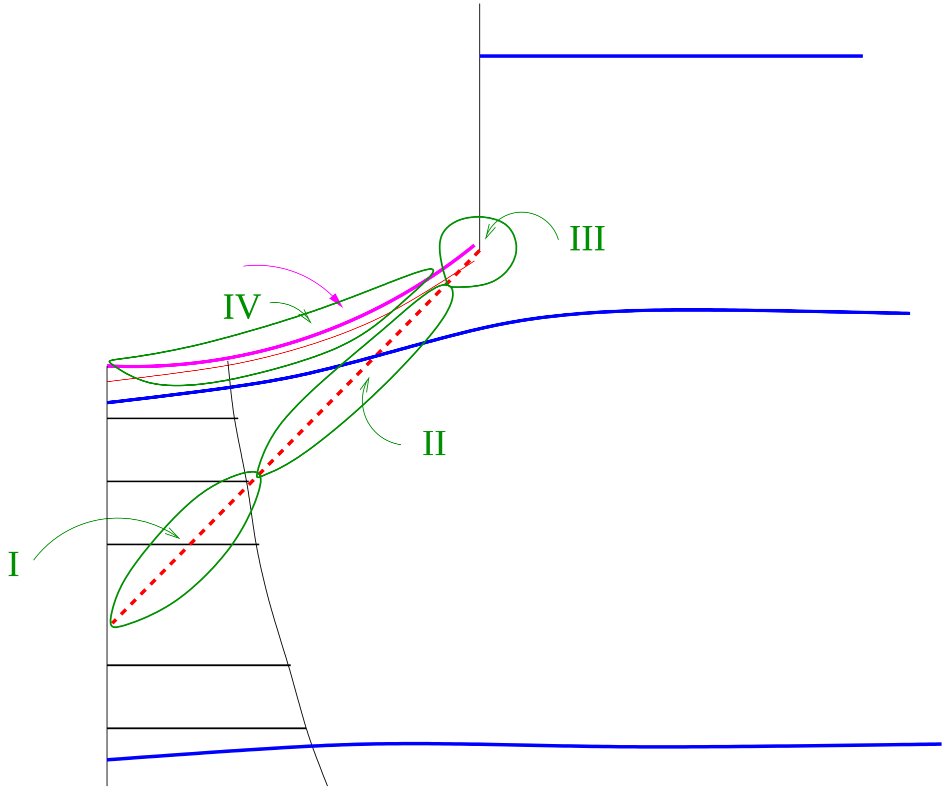
In a spacetime in which a black hole forms, there will be entanglement between the state of quantum field observables inside and outside of the back hole. This entanglement is intimately related to the Hawking radiation emitted by the black hole. In addition to the strong quantum field entanglement arising on small scales near the horizon associated with Hawking radiation, there may also be considerable additional entanglement because the matter that forms (or later falls into) the black hole may be highly entangled with matter that remains outside of the black hole.



In a semiclassical treatment, if the black hole evaporates completely, the final state will be mixed, i.e., one will have dynamical evolution from a pure state to a mixed state. In this sense, there will be irreversible "information loss" into black holes.

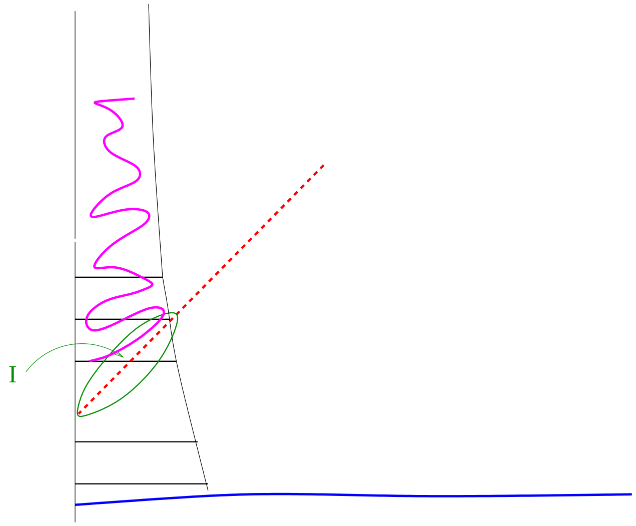
What’s Wrong With This Picture?

If the semiclassical picture is wrong, there are basically 4 places where it could be wrong in such a way as to modify the conclusion of information loss:



Possibility I: No Black Hole Ever Forms (Fuzzballs)

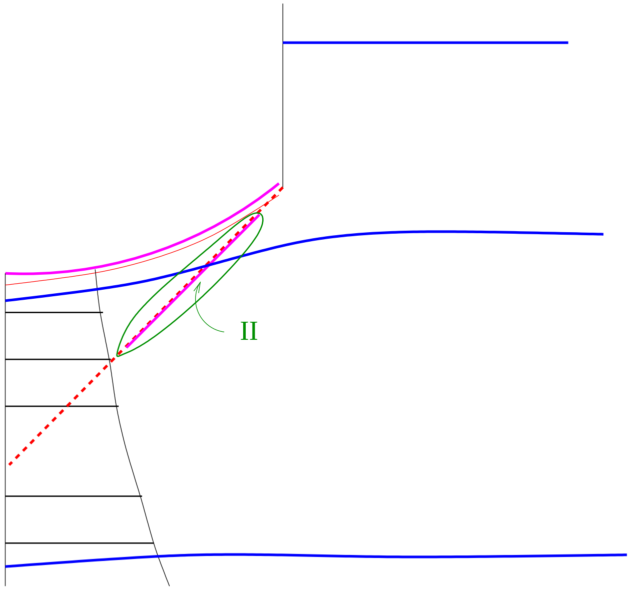
In my view, this is the most radical alternative. Both (semi-)classical general relativity and quantum field theory would have to break down in an arbitrarily low curvature/low energy regime.



Note that if the fuzzball or other structure doesn't form at just the right moment, it will be “too late” to do anything without a major violation of causality/locality in a low curvature regime as well.

Possibility II: Major Departures from Semiclassical Theory Occur During Evaporation (Firewalls)

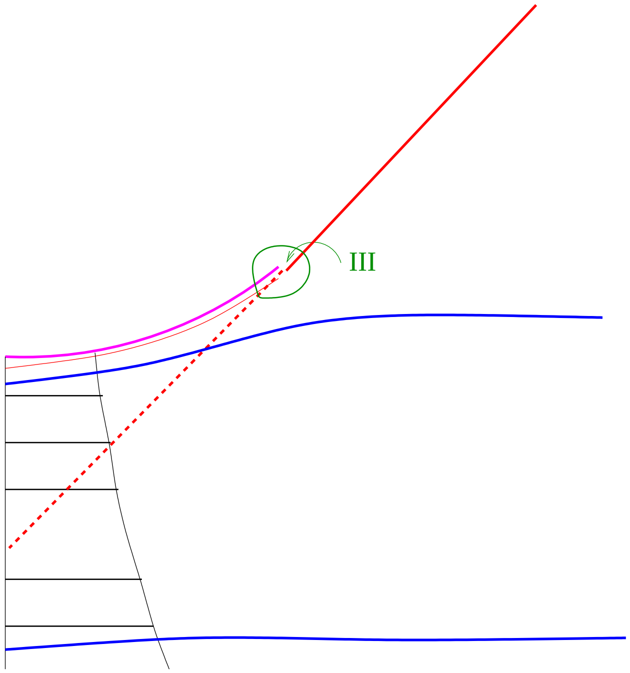
This is also a radical alternative, since the destruction of entanglement between the inside and outside of the black hole during evaporation requires a breakdown of quantum field theory in an arbitrarily low curvature regime.



"Firewalls" would need to come into existence at (or very near) the horizon in order to destroy entanglement. There is no theory of firewalls, but they would not only require a major breakdown of local laws of physics near the horizon but also require major violations of causality/locality in order to bring the entanglement from deep inside the black hole to outside the horizon.

Possibility III: Remnants

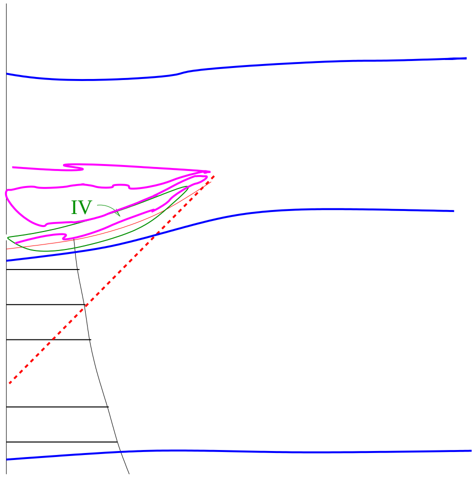
This is not a radical alternative, since the breakdown of the semi-classical picture occurs only near the Planck scale.



However, there are severe problems with invoking remnants to maintain a pure state. If the remnants cannot interact with the external world, it is not clear what "good" they do (since the "information," although still present, is inaccessible). If they can interact with the external world, then there are serious thermodynamic problems with them, since they must contain arbitrarily many states at tiny (Planck scale) energy and thus should be thermodynamically favored over all other forms of matter.

Possibility IV: A Final Burst

This alternative requires an arbitrarily large amount of "information" to be released from an object of Planck mass and size.



This is not necessarily as crazy as it might initially sound: Very recently, Hotta, Schutzhold, and Unruh have considered the model of an accelerating mirror that emits

Hawking-like radiation. The “partner particles” to the Hawking radiation are indistinguishable from vacuum fluctuations, and thus the information is “carried off” by vacuum fluctuations that are correlated with the emitted particles—at no energy cost!

This is a potentially viable means to restore information. However, it does not seem possible to do this for black holes in $3 + 1$ dimensions in a manner compatible with conservation of energy.

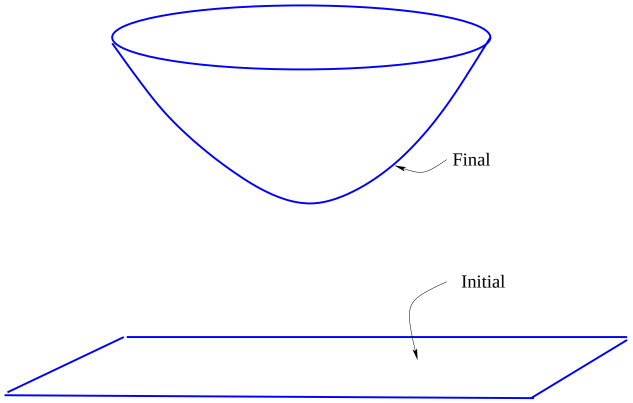
Arguments Against Information Loss:

Violation of Unitarity

In scattering theory, the word "unitarity" has 2 completely different meanings: (1) Conservation of probability; (2) Evolution from pure states to pure states.

Failure of (1) would represent a serious breakdown of quantum theory (and, indeed, of elementary logic). However, that is not what is being proposed by the semiclassical picture.

Failure of (2) would be expected to occur in any situation where the final "time" is not a Cauchy surface, and it is entirely innocuous.



For example, we get "pure → mixed" for the evolution of a massless Klein-Gordon field in Minkowski spacetime if the final "time" is chosen to be a hyperboloid. This is a *prediction* of quantum theory, not a *violation* of quantum theory.

The "pure → mixed" evolution predicted by the semiclassical analysis of black hole evaporation is of an entirely similar character.

I find it ironic that some of the same people who consider "pure → mixed" to be a violation of quantum theory then endorse truly drastic alternatives that *really are* violations of quantum (field) theory in a regime where it should be valid. I have a deep and firm belief in the validity of the known laws of quantum theory (on length and time scales larger than the Planck scale), and I will continue to vigorously defend quantum theory against those who may have initially set out to try to save it but who somehow got diverted into trying to destroy it.

Arguments Against Information Loss:
Failure of Energy and Momentum Conservation

Banks, Peskin, and Susskind argued that evolution laws taking "pure $\rightarrow$ mixed" would lead to violations of energy and momentum conservation. However, they considered only a "Markovian" type of evolution law (namely, the Lindblad equation). This would not be an appropriate model for black hole evaporation, as the black hole clearly should retain a "memory" of what energy it previously emitted.

There appears to be a widespread belief that any quantum mechanical decoherence process requires energy exchange and therefore a failure of conservation of energy for the system under consideration. This is true if the "environment system" is taken to be a thermal bath of oscillators. However, it is not true in the case where the "environment system" is a spin bath. In any case, Unruh has provided a very nice example of a quantum mechanical system that interacts with a "hidden spin system" in such a way that "pure → mixed" for the quantum system but exact energy conservation holds.

Bottom line: There is no problem with maintaining exact energy and momentum conservation in quantum mechanics with an evolution wherein "pure → mixed".

Arguments Against Information Loss: AdS/CFT

The one sentence version of AdS/CFT argument against the semiclassical picture is simply that if gravity in asymptotically AdS spacetimes is dual to a conformal field theory, then since the conformal field theory does not admit "pure → mixed" evolution, such evolution must also not be possible in quantum gravity.

AdS/CFT is a conjecture. The problem with using AdS/CFT in an argument against information loss is not that this conjecture has not been *proven*, but rather that it has not been *formulated* with the degree of precision needed to use it reliably in such an argument.

Implicit in all AdS/CFT arguments against information loss are assumptions such as (1) the correspondence is sufficiently "local" that the late time bulk observables near infinity are in 1-1 correspondence with the late time CFT observables, and (2) the CFT observables at one time comprise all of the observables of the CFT system (i.e., there is deterministic evolution of the CFT system). However, these assumptions would also suggest that a solution to Einstein's equation should be uniquely determined by its behavior near infinity at one moment of time—in blatant contradiction of the "gluing theorems" of general relativity.

I hope that the AdS/CFT ideas can be developed further so as to make a solid argument against (or for!)

information loss. A properly developed argument should provide some explanation of *how* information is regained—not just that it must happen somehow or other. Until then, I'm sticking with information loss!