Ghirardi-Rimini-Weber model with massive flashes

Antoine Tilloy[*]

Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, 85748 Garching, Germany
(Dated: January 29, 2018)

I introduce a modification of the Ghirardi-Rimini-Weber (GRW) model in which the flashes (or collapse space-time events) source a classical gravitational field. The resulting semi-classical theory of Newtonian gravity preserves the statistical interpretation of quantum states of matter in contrast with mean field approaches. It can be seen as a discrete version of recent proposals of consistent hybrid quantum classical theories. The model is in agreement with known experimental data and introduces new falsifiable predictions: (1) single particles do not self-interact, (2) the $1/r$ gravitational potential of Newtonian gravity is cut-off at short ($\lesssim 10^{-7}$m) distances, and (3) gravity makes spatial superpositions decohere at a rate inversely proportional to that coming from the vanilla GRW model. Together, the last two predictions make the model experimentally falsifiable for *all* values of its parameters.

Most attempts to unify gravity and quantum theory rely on the quantization (fundamental or effective) of the former. Yet this route has proved so hard that, in the absence of experimental evidence, it may be time to wonder if it really is unavoidable. Could a classical curved space-time and quantum matter not coexist at the most fundamental level? This latter option has been historically dismissed on the ground that its naive instantiations yield crippling conceptual difficulties [1], already in the Newtonian limit. However, these difficulties are not shared by all semiclassical approaches [2]. Recently, new proposals have been put forward to construct consistent (non-relativistic) hybrid theories of gravity [3–6]. At a formal level, these proposals implement gravity as a continuous measurement and feedback scheme on quantum matter. *Equivalently*, the proposals take the form of continuous collapse models in which a noisy modification of the mass density sources the gravitational field. The first point of view ensures that the models have a transparent empirical content and the second guarantees that they have a clear ontological interpretation and solve the measurement problem as a byproduct.

The objective of this short article is to introduce similar but conceptually more transparent and technically easier model. Starting from the simplest discrete collapse model, the Ghirardi-Rimini-Weber (GRW) model [7], instead of its continuous refinements [8–10], we will implement a gravitational interaction by making the flashes (collapse space-time events) "massive". The resulting model will yield a similar phenomenology while keeping technicalities to a minimum. The aim is not to propose a "final" model of semi-classical gravity but to demonstrate that a consistent approach, with clear physical and empirical contents, is possible. More modestly, even if fundamental semiclassical gravity is ultimately discarded, the present model may provide an effective description for a deeper quantum theory.

I. THE GRW MODEL AND ITS INTERPRETATION

We consider for simplicity an isolated system of N distinguishable particles of coordinates $\{\mathbf{x}_1, \cdots, \mathbf{x}_N\}$, masses $\{m_1, \cdots, m_N\}$, and without spin. The N-particle system is described by the wave-function $\psi_t(\mathbf{x}_1, \cdots, \mathbf{x}_n)$, with $\psi_t \in \mathrm{L}^2(\mathbb{R}^{3N}, \mathbb{C})$ which obeys a standard Schrödinger equation $\partial_t \psi_t = -\frac{i}{\hbar}\hat{H}_0\,\psi_t$, where $\hat{H}_0$ can be any many-body Hamiltonian. In the GRW model, this unitary evolution is interrupted by spontaneous jumps of the wave-function that can hit each particle independently:

$$\psi_t \longrightarrow \frac{\hat{L}_k(\mathbf{x}_f)\psi_t}{\|\hat{L}_k(\mathbf{x}_f)\psi_t\|}, \tag{1}$$

with:

$$\hat{L}_k(\mathbf{x}_f) = \frac{1}{(\pi r_C^2)^{3/4}}\mathrm{e}^{-(\hat{\mathbf{x}}_k - \mathbf{x}_f)^2/(2r_C^2)}, \tag{2}$$

where $\hat{\mathbf{x}}_k$ is the position operator associated to particle k. These N types of jumps occur independently of each other and are uniformly distributed in time with intensity λ. Finally, if a jump hits particle k at time t, the probability density for the center of the collapse $\mathbf{x}_f$ is given by $P(\mathbf{x}_f) = \|\hat{L}_k(\mathbf{x}_f)\psi_t\|^2$.

This model introduces two new parameters: the width of the collapsed wave-function r_C and the jump rate λ. They need to have small enough values that collapse events go unnoticed for small N, making the standard unitary evolution approximately valid, but large enough that macroscopic superpositions (with $N \gtrsim 10^{20}$) are almost immediately collapsed in definite locations. The typical values proposed in the literature, $\lambda = 10^{-16}\mathrm{s}^{-1}$ and $r_C = 10^{-7}$m, can be shown to achieve this goal [7, 11] (although they might soon be experimentally falsified).

The GRW model is a quantum theory without observer in the sense that what happens in measurement situations is derived rather than postulated. As measurement results are no longer primitive in this approach, it raises

* antoine.tilloy@mpq.mpg.de

the question: what is the theory fundamentally about? One proposal, going back to John Bell [12], is to regard the localized collapse events or flashes $(\mathbf{x}_f, t_f)$ as the elementary building blocks of matter ("local beables" [13] or "primitive ontology" [14–16]), of which, in particular, all macroscopic objects are composed. Our proposed model takes this interpretation seriously: if the flashes are matter, they should gravitate. The role of the wave-function, on the other hand, is only to provide the probabilistic law for their distribution. It does not represent matter and hence does not source a gravitational field.

An important feature of the GRW model, and of all collapse models for that matter, is that it is consistent with the statistical interpretation of the wave-function. Once one understands how the measurement postulate is *derived* from the model, one can use the wave-function exactly as in orthodox quantum theory (with the Born rule and collapse in measurement situations) but with a small modification of the dynamics. This is crucial to make the theory usable in practice.

The probabilities one can compute in orthodox quantum theory are obtained as sesquilinear forms on the state $|\psi\rangle$, *i.e.* they are quantities of the form $a = \langle\psi|\hat{A}|\psi\rangle$, where $\hat{A}$ is a Hermitian operator. In collapse models, one does not have a direct empirical access to the random jumps. One can only know their macroscopic configuration a posteriori as they determine measurement outcomes. As a result, the empirical content of the model –what one can predict– is contained in quantities of the form $\bar{a} = \mathbb{E}\left[\langle\psi|\hat{A}|\psi\rangle\right] = \mathrm{tr}\left(\hat{A}\,\mathbb{E}[|\psi\rangle\langle\psi|]\right)$, where $\mathbb{E}[\,\cdot\,]$ denotes the statistical average over the random jumps. Thus, although the non-linear random state evolution is necessary to understand why the theory makes sense, all the empirical content lies ultimately in the density matrix $\rho = \mathbb{E}[|\psi\rangle\langle\psi|]$. A straightforward computation shows that it obeys the equation:

$$\partial_t \rho_t = -\frac{i}{\hbar}[\hat{H}_0, \rho_t] + \lambda \sum_{k=1}^{N} \int \mathrm{d}\mathbf{x}_f \hat{L}_k(\mathbf{x}_f)\rho_t \hat{L}_k(\mathbf{x}_f) - \rho_t. \quad (3)$$

Importantly, equation (3) is *linear*. If this were not the case, the standard statistical interpretation would necessarily break down [17–19]. Indeed, if (3) were non-linear, it would mean that proper mixtures (statistical ensembles of pure states) would no longer be empirically equivalent to improper mixtures (decohered quantum superpositions of pure states obtained by tracing out unknown degrees of freedom). This equivalence is what usually guarantees that Alice and Bob cannot communicate using an EPR pair: measuring her state, Alice cannot know what Bob did with his. With non-linearities, the statistics on Alice's side would depend on Bob's measurements. The standard objection is that it would allow faster than light signalling. The difficulty is actually even more crippling: with non-linearities, reduced states can simply no longer be trusted to make predictions. The only states one ever has access to in practice being reduced states,

there can be no consistent orthodox statistical interpretation. This does not mean that the model as such would be inconsistent, but since reduced states are the only ones to which we have empirical access in practice, the usual statistical formalism of quantum mechanics could no longer be applied[1].

I insist on this linearity at the master equation level because it is at the same time necessary and unexpected. After all, we started with a non-linear collapse evolution! Yet, the jump probabilities exactly simplify with the normalization factors of (1) when averaging over the noise, thus magically removing the non-linearity for the density matrix. Of course, there is no magic here: the GRW model has been constructed in such a way as to reproduce quantum statistics. However, an arbitrary modification of the stochastic evolution will most likely not preserve this feature. Unless there is a particular reason why linearity (of the master equation) should hold, it typically will not. Conversely, this means that the prerequisite of linearity puts tight constraints on possible extensions or modifications of the GRW model (e.g. when adding gravity).

A possible understanding of why collapse models are consistent with the statistical interpretation comes with noticing that they can be formally rewritten as purely orthodox models. The "jump" part of the GRW evolution (1) takes the form of a positive-operator valued measure (POVM), *i.e.* a generalized measurement perfectly authorized in "operational" quantum mechanics. Averaging over measurement results is operationally equivalent to tracing out the degrees of freedom of the measurement apparatus and naturally gives a perfectly legitimate linear open system evolution (3) of the Lindblad form. The interpretation of the GRW model is naturally different: no observer is fuzzily measuring the particle positions and the jumps are taken as fundamental. Yet, the fact that the formalism can be recast in orthodox terms can be seen as a guarantee that the operational toolbox of quantum theory will survive.

In the orthodox picture, the flashes correspond to measurement outcomes. Measurement outcomes have a nice status in quantum theory: they are the only "classical" variables one has access to. In particular, measurement outcomes can be fed back into the evolution of the quantum state without causing trouble. Some controller can apply an external field to a quantum system contingent on a measurement outcome without wreaking havoc on the statistical interpretation of states. If it were not so, the operational toolbox of quantum theory would be unusable. Obviously in the GRW model, no observer has access to individual flashes. However, because they are equivalent to measurement outcomes at a formal level,

---

[1] The linearity of the equation for the density matrix in collapse models plays the same role as "equivariance" in Bohmian mechanics [20]. Without it, the theory might still make sense but its empirical content becomes almost impossible to extract.

one can make the wave-function dynamics depend on them. If we are in the business of constructing new theories, it is legitimate to use the flashes as classical sources. This makes the flashes very different from other primitive ontologies proposed in foundations. Neither the quantum expectation value of the mass density operator nor the Bohmian particle trajectories can backreact on the wave-function without ruining the statistical interpretation. To my knowledge, the flashes (or their field equivalent for continuous collapse models) are the only candidates for a primitive ontology so far proposed that need not be purely *passive*[2] (*i.e.* determined by the wave-function without influencing it in return).

II.   SOURCING GRAVITY FROM THE FLASHES

Guided by the previous intuition, we can use the flashes to source a classical gravitational field Φ (instead of the expectation value of the mass density discussed in [21]). We start by considering that a flash just happened and forget the rest of the dynamics. For a flash $(\mathbf{x}_f, t_f)$ associated to the collapse operator $\hat{L}_k$ and thus to particle k, we source Φ using the Poisson equation:

$$\Delta\Phi(\mathbf{x},t) = 4\pi G m_k \lambda^{-1} f(t - t_f, \mathbf{x} - \mathbf{x}_f), \qquad (4)$$

where f is a space-time form factor that smears the flash in space-time. It should satisfy $\int \mathrm{d}^3\mathbf{x}\,\mathrm{d}t\, f(t,\mathbf{x}) = 1$, $f \geq 0$ and $\forall t < 0, f(t,\mathbf{x}) = 0$. As the particles flash at a rate λ, the factor λ^{-1} insures that the field generated by the flashes corresponds on average to the field created by a classical particle in the same position. We will mainly be interested in the singular limit where the gravitational potential is sourced by the sharp flash only and thus take $f(t - t_f, \mathbf{x} - \mathbf{x}_f) \to \delta(t - t_f)\delta^3(\mathbf{x} - \mathbf{x}_f)$. However, a general f may be useful[3] for extensions out of the Newtonian limit where point sources would yield divergent metrics.

We now add the corresponding gravitational potential in the evolution for the wave-function as a totally benign external potential:

$$\hat{V}_{\mathrm{G}} = \int \mathrm{d}\mathbf{x}\,\Phi(\mathbf{x})\,\hat{M}(\mathbf{x}) \qquad (5)$$

where $\hat{M}(\mathbf{x}) = \sum_{\ell=1}^{N} m_\ell\, \delta(\mathbf{x} - \hat{\mathbf{x}}_\ell)$ is the mass density operator. Inverting the Poisson equation (4) gives:

$$\hat{V}_G(t) = -G\lambda^{-1} \sum_{\ell=1}^{N} m_k m_\ell \int \mathrm{d}\mathbf{x} \frac{f(t - t_f, \mathbf{x} - \mathbf{x}_f)}{|\mathbf{x} - \hat{\mathbf{x}}_\ell|}. \qquad (6)$$

---

[2] I should insist that being passive is in no way a fundamental problem for a primitive ontology. Being able to consistently back react on the wave-function simply offers more flexibility for theory building.

[3] An example of a more regular sourcing would be to take the same Gaussian space smearing as the one used in the collapse operators $\hat{L}_k(\hat{\mathbf{x}})$, while keeping a Dirac in time to preserve Markovianity. This minor extension can be treated without difficulty.

This is the potential created by a single flash. By linearity, the total potential is simply the sum of the potentials associated with all the past flashes. Taking the limit $f(t - t_f, \mathbf{x} - \mathbf{x}_f) \to \delta(t - t_f)\delta^3(\mathbf{x} - \mathbf{x}_f)$ will make the potential (6) created by each flash divergent but the corresponding unitary $\hat{U}_k(\mathbf{x}_f, k) = \exp\left(-i\int_{t_f}^{+\infty} \mathrm{d}t\, \hat{V}_G(t)/\hbar\right)$ applied on the wave-function –instantaneous in the limit– will remain finite. Indeed, in that limit we get:

$$\hat{U}_k(\mathbf{x}_f) = \exp\left(i\,\frac{G}{\lambda\hbar} \sum_{\ell=1}^{N} \frac{m_k m_\ell}{|\mathbf{x}_f - \hat{\mathbf{x}}_\ell|}\right). \qquad (7)$$

Hence in the absence of smearing, the physical picture becomes particularly simple. Immediately after a flash, an instantaneous unitary transformation is applied to the wave-function corresponding to the singular gravitational pull of the flash. The evolution is thus very close to that of the original GRW model but with a modification of the jump operators. Indeed, after particle k flashes, the wave-function experiences the transformation:

$$\psi_t \longrightarrow \hat{U}_k(\mathbf{x}_f)\frac{\hat{L}_k(\mathbf{x}_f)\psi_t}{\|\hat{L}_k(\mathbf{x}_f)\psi_t\|} \equiv \frac{\hat{B}_k(\mathbf{x}_f)\psi_t}{\|\hat{B}_k(\mathbf{x}_f)\psi_t\|}. \qquad (8)$$

As before we obtain the master equation for the density matrix:

$$\partial_t \rho_t = -\frac{i}{\hbar}[\hat{H}_0, \rho_t] + \lambda\sum_{k=1}^{N} \int \mathrm{d}\mathbf{x}_f\, \hat{B}_k(\mathbf{x}_f)\rho_t\hat{B}_k^\dagger(\mathbf{x}_f) - \rho_t. \qquad (9)$$

This is the central result of this short article.

III.   PHENOMENOLOGY

We now study the evolution of a single isolated particle. Ignoring the free Hamiltonian $\hat{H}_0$ to focus on the gravitational and collapse effects, we see that the master equation (3) is diagonal in position:

$$\partial_t \rho_t(\mathbf{x}, \mathbf{y}) = \lambda(\Gamma(\mathbf{x}, \mathbf{y}) - 1)\rho_t(\mathbf{x}, \mathbf{y}), \qquad (10)$$

where $\Gamma(\mathbf{x}, \mathbf{y})$ is an *a priori* complex kernel:

$$\Gamma(\mathbf{x},\mathbf{y}) = \int \frac{\mathrm{d}\mathbf{x}_f}{(\pi r_C^2)^{3/2}} \exp\left(i\frac{Gm^2}{\lambda\hbar}\left[\frac{1}{|\mathbf{x} - \mathbf{x}_f|} - \frac{1}{|\mathbf{y} - \mathbf{x}_f|}\right]\right)$$
$$\times \exp\left(-\frac{(\mathbf{x} - \mathbf{x}_f)^2 + (\mathbf{y} - \mathbf{x}_f)^2}{2r_C^2}\right). \qquad (11)$$

Changing of variable with $\tilde{\mathbf{x}}_f = \mathbf{x}_f + (\mathbf{x} + \mathbf{y})/2$, one sees that the integral on the upper half space is equal to the conjugate of the integral on the lower half space, hence Γ is *real*. It means that the effect of collapse and gravity on a single particle is simply a decay of the phases in the position basis. A single particle does not attract itself in

this model. This absence of self-attraction is a counterintuitive prediction markedly distinguishing this approach from other semi-classical theories relying on mean field.

We further notice that the master equation (9) involves a new length scale $r_G = Gm^2/(\hbar\lambda)$. For $\lambda = 10^{-16}\text{s}^{-1}$, we get $r_G \simeq 1.8 \times 10^{-14}$m for protons and $r_G \simeq 5.3 \times 10^{-21}$m for electrons. This length scale appears notably in the short distance behavior of decoherence. Indeed, from (11) we get:

$$\lambda(1 - \Gamma(\mathbf{x}, \mathbf{y})) \underset{|\mathbf{x}-\mathbf{y}|\to 0}{\sim} \frac{\lambda}{\sqrt{\pi}} \frac{r_G^2}{r_C^3} |\mathbf{x} - \mathbf{y}| \propto \lambda^{-1} |\mathbf{x} - \mathbf{y}|. \quad (12)$$

Importantly, this means that the coupling with gravity creates decoherence effects of strength inversely proportional to the collapse rate λ: the range of values allowed for the latter is experimentally falsifiable from below.

Finally, we need to make sure that the model is at least empirically adequate in the Newtonian limit. For that matter, we consider a test particle of coordinate $\mathbf{x}_0$ interacting with a lump of N particles of coordinates $\mathbf{x}_1, \cdots, \mathbf{x}_N$. The only thing we assume is that the test particle is separated from the lump by a distance much larger than r_C. Ignoring again $\hat{H}_0$, we can expand (9) to first order in $\varepsilon = r_G/r_C < 10^{-7}m$ to write the approximate master equation for this set of $N + 1$ particles:

$$\partial_t \rho_t \simeq \lambda \sum_{k=0}^N \left(\int d\mathbf{x}_f \hat{L}_k(\mathbf{x}_f)\rho_t \hat{L}_k(\mathbf{x}_f) - \rho_t \right)$$
$$+ i \sum_{k,\ell=0}^N \frac{Gm_k m_\ell}{\hbar} \int d\mathbf{x}_f \left[\frac{1}{|\hat{\mathbf{x}}_\ell - \mathbf{x}_f|} - \frac{1}{|\hat{\mathbf{y}}_\ell - \mathbf{x}_f|} \right] \quad (13)$$
$$\times \hat{L}_k(\mathbf{x}_f)\rho_t \hat{L}_k(\mathbf{x}_f).$$

We shall be interested in the reduced density matrix of the test particle $\rho_t^\circ(\mathbf{x}_0, \mathbf{y}_0) = \text{tr}_N(\rho_t)(\mathbf{x}_0, \mathbf{y}_0)$. Taking the trace in (13) cancels most of the terms and one gets:

$$\partial_t \rho_t^\circ(\mathbf{x}_0, \mathbf{y}_0) = \int d\mathbf{x}_f \hat{L}_0(\mathbf{x}_f)\rho_t^\circ \hat{L}_0(\mathbf{x}_f) - \rho_t^\circ$$
$$+ i \sum_{k=1}^N \frac{Gm_0 m_k}{\hbar} \int d\mathbf{x}_1 \cdots d\mathbf{x}_N \, d\mathbf{x}_f \left[\frac{1}{|\mathbf{x}_0 - \mathbf{x}_f|} - \frac{1}{|\mathbf{y}_0 - \mathbf{x}_f|} \right]$$
$$\times \frac{1}{(\pi r_C^2)^{3/2}} e^{-\frac{|\mathbf{x}_k - \mathbf{x}_f|^2}{r_C^2}} \rho_t(\mathbf{x}_0, \mathbf{x}_1, \cdots, \mathbf{x}_N, \mathbf{y}_0, \mathbf{x}_1, \cdots, \mathbf{x}_N). \quad (14)$$

We now use the fact that the test particle is far from the lump to neglect the Gaussian convolution in (14):

$$\partial_t \rho_t^\circ \simeq \lambda(\Gamma_0(\rho_t^\circ) - \rho_t) - i \, \text{tr}_N \left\{ \left[-\sum_{k=1}^N \frac{Gm_0 m_k}{\hbar|\hat{\mathbf{x}}_0 - \hat{\mathbf{x}}_k|}, \rho_t \right] \right\}, \quad (15)$$

where $\Gamma_0(\rho_t^\circ)(\mathbf{x}_0, \mathbf{y}_0) = \Gamma_0(\mathbf{x}_0, \mathbf{y}_0)\rho_t^\circ(\mathbf{x}_0, \mathbf{y}_0)$ contains the single particle decoherence coming from the original GRW model. Equation (14) shows that the statistics

on the test particle are approximately *as if* it were interacting with the N particles of the lump through a (quantum) Newtonian pair potential. The crucial (and only) hypothesis to obtain this result is that the distance between the test particle and the lump is much larger than r_C.

If we further assume that the particles in the lump are well localized (typically because of the collapse mechanism itself) around positions $\mathbf{r}_1, \cdots, \mathbf{r}_N$ on a length scale far smaller than the lump-test particle distance, we can remove the trace and obtain the classical limit:

$$\partial_t \rho_t^\circ \simeq \lambda(\Gamma_0(\rho_t^\circ) - \rho_t) - i \left[-\sum_{k=1}^N \frac{Gm_0 m_k}{\hbar|\hat{\mathbf{x}}_0 - \mathbf{r}_k|}, \rho_t^\circ \right]. \quad (16)$$

Apart again from the decoherence coming from the collapse, this is exactly what one would have wished for. A classical piece of matter approximately creates the classical gravitational field one would naively expect, at least at a distance $d \gg r_C \sim 10^{-7}$m away from it. At shorter distances, the effective Newtonian pair potential is smoothed by the collapse operators and the $1/r$ law of Newtonian gravity breaks down. However, this behavior is still compatible with all the experiments carried out so far where the gravitational field probed is typically sourced by large objects at least 10^{-4}m away from the test mass [22].

Including gravity in the GRW model in the way we have suggested, *i.e.* by making the flashes gravitate, has an interesting consequence on the parameter diagram of the model. The problem of vanilla collapse models is that parameters leading to very unsharp or very infrequent collapse can only ever be philosophically discarded. With a λ too small or r_C too large, the GRW model does not reduce macroscopic superpositions before they include different observer states. This leads, in effect, to a Many-Worlds theory that cannot be empirically discarded but undermines the main motivation for considering the collapse mechanism in the first place. However, there is no way to experimentally falsify this fuzzily defined zone of metaphysical discomfort.

The present model is free of this issue. The leading decoherence term is $\propto \lambda$ (the "instrinsic" GRW decoherence) but the coupling with gravity adds decoherence $\propto \lambda^{-1}$. This means that for a fixed value of r_C, all the values of λ can *in principle* be experimentally falsified. Intuitively, an infrequent collapse yields infrequent but supermassive flashes and thus a noisy gravitational interaction yielding strong gravitational decoherence. Reciprocally, reducing gravitational decoherence requires frequent flashes and thus a strong intrinsic decoherence.

Further, the effective potential one obtains is smoothed by the collapse operators and the $1/r$ law of gravity breaks down for distances $d \lesssim r_C$. Hence, small values of r_C can be falsified by decoherence and large values of r_C can be falsified by probing the gravitational force at short distances. The model is –in principle– experimentally falsifiable for all values of λ and r_C.

IV. DISCUSSION

The dynamical equations of collapse models like the GRW model are the same as the one that describe repeated unsharp measurements of particle positions (or mass density for continuous collapse models). This formal equivalence is ultimately what guarantees that the statistical interpretation of quantum states survives in such models. Further, it shows that the space-time events (or flashes) labeling collapse events are *formally* measurement outcomes. This means that, even though they are not known to observers, one can let them backreact explicitly on the quantum state as one would in a feedback scheme. Such a coupling does not create the usual difficulties of mean field couplings which yield non-linearities at the master equation level [23]. This is a crucial insight for theory building allowing to consistently couple matter with a classical gravitational field. Following this idea, we made the flashes themselves source the gravitational field, as infinitely "massive" space-time events. The model we obtained gives the Newtonian gravitational force one would expect between particles for distances larger than the collapse length scale r_C and thus is in agreement with known experiments. However, particles do not attract themselves, contrary to what one would have guessed from naive semi-classical theories. The model is falsifiable for all possible values of its two parameters λ and r_C in contrast with standard collapse models where only large values of λ and small values of r_C can be experimentally eliminated.

In the end, if one pays the price of collapse models to solve the measurement problem, the possibility to construct consistent hybrid dynamics comes for free. The implementation of such dynamics then appreciably constrains the parameter diagram of collapse models. The present approach thus has appealing features. That said, how much should we *believe* in the model introduced here? Theoretical physics is already drowning in wild speculations, sometimes prematurely elevated to the status of truth, hence a bit of soberness is required. The present model is a proof of principle rather than a proposal for a realistic fundamental theory. Nonetheless, important lessons survive its limited scope:

1. There is no obstacle in principle to construct consistent semi-classical theories of gravity.

2. Collapse models can be empirically constrained by a natural coupling with gravity.

3. A primitive ontology can have a central dynamical role and need not be merely passive.

If semi-classical theories of the type presented here can be extended to general relativity[4] in a convincing way (which is arguably far from trivial) and if robust criteria can be found to make them less ad hoc (see *e.g.* [6]), then further optimism will be warranted. Unless, of course, experiments provide a hint in the meantime [25].

[1] K. Eppley and E. Hannah, Found. Phys. **7**, 51 (1977).
[2] M. Albers, C. Kiefer, and M. Reginatto, Phys. Rev. D **78**, 064051 (2008).
[3] D. Kafri, J. M. Taylor, and G. J. Milburn, New J. Phys. **16**, 065020 (2014).
[4] D. Kafri, G. J. Milburn, and J. M. Taylor, New J. Phys. **17**, 015006 (2015).
[5] A. Tilloy and L. Diósi, Phys. Rev. D **93**, 024026 (2016).
[6] A. Tilloy and L. Diósi, arXiv:1706.01856 (2017).
[7] G. C. Ghirardi, A. Rimini, and T. Weber, Phys. Rev. D **34**, 470 (1986).
[8] L. Diósi, Phys. Rev. A **40**, 1165 (1989).
[9] G. C. Ghirardi, P. Pearle, and A. Rimini, Phys. Rev. A **42**, 78 (1990).
[10] A. Bassi, K. Lochan, S. Satin, T. P. Singh, and H. Ulbricht, Rev. Mod. Phys. **85**, 471 (2013).
[11] W. Feldmann and R. Tumulka, J. Phys. A: Math. Theor. **45**, 065304 (2012).
[12] J. Bell, "Are there quantum jumps?" in *Schrödinger: Centenary Celebration of a Polymath*, edited by C. W. Kilmister (Cambridge University Press, 1987) pp. 41–52.
[13] J. S. Bell, Epistemological Letters **9** (1976).
[14] V. Allori, S. Goldstein, R. Tumulka, and N. Zanghì, Brit. J. Philos. Sci. **65**, 323 (2014).
[15] V. Allori, International Journal of Quantum Foundations **1**, 107 (201
[16] R. Tumulka, arXiv:1102.5767 (2011).
[17] N. Gisin, Phys. Lett. A **143**, 1 (1990).
[18] J. Polchinski, Phys. Rev. Lett. **66**, 397 (1991).
[19] A. Bassi and K. Hejazi, Eur. J. Phys. **36**, 055027 (2015).
[20] D. Dürr and S. Teufel, *Bohmian Mechanics* (Springer Berlin Heidelberg, 2009).

---

[4] On top of the difficulty to construct relativistic collapse models, one may worry that no divergence-free quantity $T_{\mu\nu}$ (needed to source gravity in general) can be constructed from a flash-like object. This latter problem might be bypassed by going to unimodular gravity. In this formulation, energy non conservation is tolerated and can source an effective cosmological constant [24].

[21] M. Derakhshani, Phys. Lett. A **378**, 990 (2014).

[22] C. A. Hagedorn, *A Sub-Millimeter Parallel-Plate Test of Gravity*, Ph.D. thesis, University of Washington (2015).

[23] M. Bahrami, A. Großardt, S. Donadi, and A. Bassi, New J. Phys. **16**, 115007 (2014).

[24] T. Josset, A. Perez, and D. Sudarsky, Phys. Rev. Lett. **118**, 021102 (2017).

[25] S. Bose, A. Mazumdar, G. W. Morley, H. Ulbricht, M. Toros, M. Paternostro, A. Geraci, P. Barker, M. S. Kim, and G. Milburn, arXiv:1707.06050 (2017).