Challenging spontaneous collapse models with ultracold layered force sensors

A. Vinante,[1,2,*] M. Carlesso,[3,4] A. Bassi,[3,4] A. Chiasera,[5] S. Varas,[5] P. Falferi,[2] B. Margesin,[6] R. Mezzena,[7] and H. Ulbricht[1]

[1]*Department of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, United Kingdom*
[2]*IFN-CNR and Fondazione Bruno Kessler, I-38123, Trento, Italy*
[3]*Department of Physics, University of Trieste, Strada Costiera 11, 34151 Trieste, Italy*
[4]*Istituto Nazionale di Fisica Nucleare, Trieste Section, Via Valerio 2, 34127 Trieste, Italy*
[5]*IFN-CNR CSMFO Lab and FBK Photonics Unit, I-38123 Trento, Italy*
[6]*Fondazione Bruno Kessler - CMM, I-38123, Trento, Italy*
[7]*Department of Physics, University of Trento, I-38123, Trento, Italy*

(Dated: February 25, 2020)

Despite the unquestionable empirical success of quantum theory, witnessed by the recent uprising of quantum technologies, the debate on how to reconcile the theory with the macroscopic classical world is still open. Spontaneous collapse models, based on stochastic and nonlinear modifications of the Schrödinger equation, are the only experimentally testable solution so far proposed. Here, we describe a new experiment based on monitoring a high quality factor microcantilever loaded by a layered test mass, specifically designed to test the Continuous Spontaneous Localization (CSL) model. The measurements are in good agreement with pure thermal motion for temperatures down to 100 mK. From the absence of excess noise we infer a new bound on the collapse rate at the characteristic length $r_c = 10^{-7}$ m, which improves over previous mechanical experiments by more than one order of magnitude. Our results are explicitly challenging a well-motivated region of the CSL parameter space proposed by Adler.

The question whether the quantum superposition principle remains valid all the way up to the macroscopic domain is still debated. While the widespread belief is that linearity is a fundamental property of nature [1, 2], over and over this assumption has been questioned[3–7]. Spontaneous collapse models [8–12] offer a clear and, under fairly general assumptions [13–15], unique phenomenology describing the break-down of quantum superpositions when moving towards the macroscopic scale, while preserving the quantum properties of microscopic systems. By construction they are empirically falsifiable [16], and are therefore attracting increasing experimental interest [17–38].

The general assumption of collapse models is that a universal classical noise drives the state of any material system towards a localized state, even in absence of any measurement process. An inbuilt amplification mechanism makes sure that the collapse scales with the size of the system, so that only sufficiently macroscopic objects are effectively localized [15, 16].

In this work, we present a new experimental test of the Continuous Spontaneous Localization (CSL) model [9, 10]. In CSL the noise is characterized by two phenomenological parameters: the collapse rate λ, measuring the strength of the collapse, and a characteristic length r_c, defining its spatial resolution. The conservative values $\lambda \simeq 10^{-17}\,\mathrm{s}^{-1}$ and $r_c = 10^{-7}\,\mathrm{m}$ [8, 9] were initially proposed by Ghirardi *et al.* [8, 9] by assuming that the collapse becomes effective at the transition between the mesoscopic and the macroscopic world. A larger value for λ has been suggested by Adler [10], under the assumption that the collapse is already effective at mesoscopic scale, resulting in $\lambda \sim 10^{9\pm2}$ times larger than at $r_c = 10^{-7}\,\mathrm{m}$, and $\sim 10^{11\pm2}$ times larger at $r_c = 10^{-6}\,\mathrm{m}$. Moreover, according to Adler, values much larger or smaller of r_c are physically less motivated [10].

The current strongest experimental bounds on the CSL parameters are shown in Fig. 4. They come from non-interferometric tests, which exploit an unavoidable indirect effect of collapse models, namely a tiny violation of the energy conservation [8]. Relevant examples are spontaneous X-ray emission from Germanium [19–21], spontaneous heating of massive bulk systems [22–26] or universal force noise on mechanical systems [27–38]. Bounds based on the first two effects are already ruling out Adler's parameters, but they can be easily evaded by reasonable assumptions on the spectrum of the CSL noise [22, 39]. Conversely, experiments based on mechanical resonators, with characteristic frequency in the mHz–kHz range, are more robust against changes in the noise properties.

Following Refs. [30–32], the one-sided spectral density of the CSL force noise on the x direction acting on a mass density distribution $\rho(\boldsymbol{r})$ can be written as:

$$S_{F_{\mathrm{CSL}}} = \frac{\hbar^2 \lambda r_c^3}{\pi^{3/2} m_0^2} \int \mathrm{d}\boldsymbol{k}\, k_x^2 e^{-k^2 r_c^2} |\tilde{\rho}(\boldsymbol{k})|^2, \qquad (1)$$

where $\tilde{\rho}(\boldsymbol{k})$ is the Fourier transform of $\rho(\boldsymbol{r})$ and m_0 is the nucleon mass. The effect described by Eq. (1) features a non-trivial dependence on the geometry, and can be enhanced around a given r_c by a properly designed multilayered test mass, as discussed in detail in Ref. [40]. In order to detect the smallest possible CSL effect, one needs to minimize the thermal force noise spectral density $S_{F_{\mathrm{th}}} = 4k_B T m\omega_0/Q$, which calls for mechanical res-

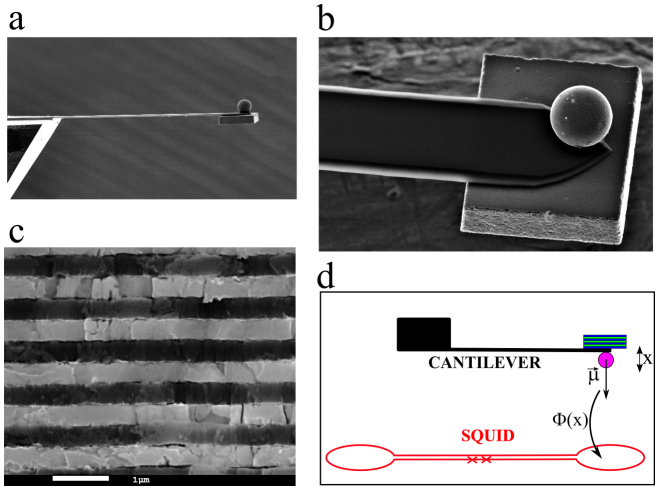

FIG. 1: Details of the experiment. (a) Low resolution SEM micrograph of the assembled cantilever, with the multilayer test mass and the magnetic microsphere. (b) and (c) SEM micrographs of the multilayer test mass, from top (b) and side (c) view respectively. Here, the alternate layers of WO_3 and SiO_2 are shown in bright and dark respectively. (d) Simplified scheme of the detection technique, with a gradiometric SQUID magnetometer which detects the variable magnetic field induced by the oscillating ferromagnetic microsphere.

onators with low temperature T, low frequency ω_0 and high Q.

In our experiment, the mechanical sensor is a silicon cantilever (Fig. 1a) of the type developed for atomic force microscopy. The same sensor was used in previous tests of CSL [33]. We load the cantilever with a test mass, engineered to optimize the effect of the CSL noise for $r_c \sim 10^{-7}\,\mathrm{m}$ (Figs. 1b and 1c). The test mass is a cuboidal multilayer structure formed by 47 alternate layers of SiO_2 and WO_3, fabricated by sputtering (see Methods for details on the fabrication). As described in detail in Ref. [40], the multilayer structure enhances the effect of the CSL noise for $r_c \lesssim 3d$ where d is the mean layer thickness. The enhancement scales as the density contrast $\Delta\rho = \rho_1 - \rho_2$. In this experiment, we have $\rho_1 = 7.17 \times 10^3\,\mathrm{kg/m^3}$ and $\rho_2 = 2.20 \times 10^3\,\mathrm{kg/m^3}$, which are respectively the densities of WO_3 and SiO_2. The mean layer thickness $d = (370 \pm 4)\,\mathrm{nm}$, is chosen to maximize the CSL noise enhancement at $r_c \approx 10^{-7}$ m. Based on the measured geometrical parameters, we estimate the value of the multilayer mass $m = (7.1 \pm 0.2) \times 10^{-10}\,\mathrm{kg}$.

We attach to the cantilever a second smaller mass, a ferromagnetic microsphere, whose motion is detected by a Superconducting Quantum Interference Device (SQUID) magnetic flux sensor (see Fig. 1d) placed at a distance of $\sim 50\,\mu\mathrm{m}$ [33]. This detection method is very convenient and, due to the low power dissipated by the SQUID, is compatible with the low temperature regime of the experiment. Moreover, the magnetomechanical coupling in our setup is low enough to make the back-action force noise from the SQUID negligible. We notice that the ferromagnetic sphere and the cantilever itself give additional, although smaller, contributions to the CSL force noise, which have been accounted for.

Cantilever and SQUID are enclosed in a mechanically isolated shielded copper box, thermally linked to the mixing chamber plate of a dry dilution refrigerator. The mixing chamber temperature is stabilized by a PID controller. Before performing any measurement we wait for at least two hours to ensure that the temperature is settled, although the thermalization time is expected to be much shorter. We measure resonance frequency and quality factor of the fundamental flexural mode of the cantilever by means of ringdown measurements. The resonance frequency is $f_0 = 3532.7\,\mathrm{Hz}$, while it was measured as $f_0' = 8174\,\mathrm{Hz}$ before attaching the multilayer test mass and with the magnetic sphere already in place [33]. Accordingly, we use the added mass method to estimate the effective stiffness k of the cantilever with respect to the effective position of the test mass on the cantilever. We obtain the value $k = (0.43 \pm 0.01)\,\mathrm{N/m}$. As observed in a previous experiment, the intrinsic quality factor depends slightly on temperature. This is likely due to the effect of two-level systems in the silicon cantilever [33]. The maximum measured quality factor is $Q = (2.83 \pm 0.03) \times 10^6$ at the lowest operation temperature $T = 30\,\mathrm{mK}$.

At a given temperature T, we estimate the power spectral density of the force noise by acquiring and averaging a large number of high resolution periodograms of the SQUID magnetic flux signal. The cantilever motion appears as a Lorentzian peak centered at f_0 on top of a white noise floor mainly due to the SQUID imprecision noise. The amplitude of the peak depends on T. Some representative averaged spectra are shown in Fig. 2. We perform a weighted fit of each spectrum with the theoretical curve expressed by:

$$S_\Phi = A + \frac{Bf_0^4 + C(f^2 - f_1^2)^2}{(f^2 - f_0^2)^2 + \left(\frac{ff_0}{Q_a}\right)^2}, \qquad (2)$$

.

Here, the apparent quality factor Q_a takes the place of the intrinsic (or true) Q, which is related to the thermal noise. Q_a is generally different from Q, due to a well understood cold damping effect induced by the SQUID feedback electronics [33]. The same mechanism is responsible for a small asymmetry term proportional to C. The parameter A is the white noise floor, mainly due to the SQUID noise. The Lorentzian term amplitude B contains the relevant information on the force noise, and can be expressed as:

$$B = \Phi_x^2 \left(\frac{S_{F0}}{k^2} + \frac{4k_B T}{k\omega_0 Q} \right), \qquad (3)$$

where $\Phi_x = \mathrm{d}\Phi/\mathrm{d}x$ is the magnetomechanical coupling factor which converts a cantilever displacement x into a

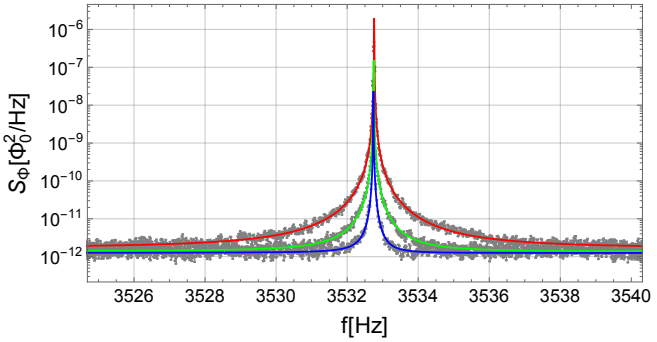

FIG. 2: Representative averaged power spectra of the SQUID flux noise around the cantilever fundamental resonance. The curves refer to the temperatures $T = 1000\,\mathrm{mK}$, $T = 200\,\mathrm{mK}$ and $T = 30\,\mathrm{mK}$ respectively from the top to the bottom. The solid lines are the best fits to the three dataset with Eq. (2).

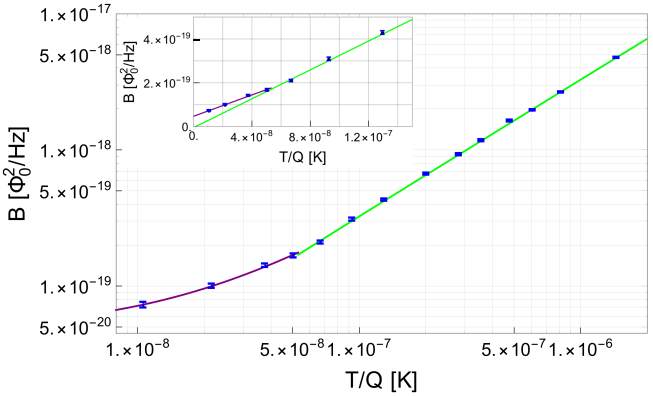

FIG. 3: Measured amplitude of the Lorentzian peak B, as function of T/Q. The main panel shows all data in Log-Log scale, for better visualization. The inset reports only the low temperature points in linear scale, to underline the crossover between the high temperature and low temperature regimes. The two solid curves represent the linear fits for the points below and above crossover temperature. The linear fit of the data at high temperature is used to bound the CSL noise.

SQUID magnetic flux Φ, S_{F0} is the spectral density of any nonthermal force noise, and the last term is the thermal noise, which according to the fluctuation-dissipation theorem is proportional to T/Q. The identification of the latter term in Eq. (3) allows us to determine Φ_x and thus to calibrate any non-thermal contribution to B.

Fig. 3 shows the measured B as a function of T/Q. The data follow the expected linear behaviour described in Eq. (3) down to $T/Q = 67\,\mathrm{nK}$, which corresponds to $T = 100\,\mathrm{mK}$. However, the data at lower temperatures (inset of Fig. 3) indicate a crossover to a different linear regime, characterized by a lower slope and positive intercept. A natural way to explain the crossover is by assuming that at least two dissipation channels are acting on the cantilever motion, one of which is not cooling further below the crossover temperature [32, 41]. Formally, we split the dissipation, expressed in terms of the inverse quality factor, as $1/Q = 1/Q_1 + 1/Q_2$, where Q_1 and Q_2 are associated to different thermal baths, respectively at the temperatures T_1 and T_2. In the high temperature limit the system is well thermalized, and $T_1 = T_2 = T$, where T is the temperature measured by a calibrated thermometer placed on the experimental stage. In the low temperature limit, one of the two temperatures, say T_1, saturates to a constant crossover temperature $T_1 \approx T_{\mathrm{co}}$, while the second bath is still well thermalized, $T_2 = T$. Thermal saturations in low temperature systems are typically described by the relation:

$$T_1 = \left(T_{\mathrm{co}}^n + T^n\right)^{1/n}. \tag{4}$$

Such a relation is obtained by assuming a residual heat load on the bath at T_1 combined with a finite thermal conductance towards the main bath at T varying as T^{n-1} [41, 42]. Considering that thermal conductivities through any material and interface tend to 0 as $T \to 0$ and residual heat loads are practically unavoidable, thermal saturation is an universal effect which is expected to arise in any system at sufficiently low temperature.

In our case, we suspect that the cantilever motion couples magnetically to dissipating elements located on the SQUID substrate surface. These could be either surface electron spins [43] or vortices in superconducting films [44]. The SQUID substrate temperature will be higher than T due to the constant power dissipation $\sim 1\,\mathrm{nW}$ in the SQUID shunt resistors. Consequently, we estimate that the saturation process takes places in the $50 - 100\,\mathrm{mK}$ temperature range. Such a process is well described by Eq. (4) with $n = 4$, which corresponds to a Kapitza boundary thermal resistance as limiting thermalization mechanism [42].

Indeed, the whole dataset of Fig. 3 can be fitted with the combined function $B = B_0 + B_a\left(x^4 + x_{\mathrm{co}}^4\right)^{1/4} + B_b x$, where $x = T/Q$. This analysis provides the determination of the crossover at $(T/Q)_{\mathrm{co}} \approx 53\,\mathrm{nK}$ which corresponds to $T_{\mathrm{co}} \approx 85\,\mathrm{mK}$. Note that the saturation is effectively very sharp, so that the related excess noise rapidly vanishes for $T \gtrsim T_{\mathrm{co}}$. However, the data are in principle compatible with other models with larger n, and do not allow to make conclusive claims on the actual saturation mechanism.

In the following, we will make the assumption that the observed crossover is indeed related to a thermal saturation, regardless of its precise physical origin. To estimate the magnitude of a possible CSL noise effect compatible with the experiment, we restrict our analysis to the range $T \geq 100\,\mathrm{mK}$, thus excluding the data below the crossover temperature. The restricted dataset follows a linear behaviour remarkably well. A weighted orthogonal linear fit with the function $B_0 + B_1 T/Q$ yields the values $B_0 = (-4.64 \pm 5.31) \times 10^{-21}\,\Phi_0^2/\mathrm{Hz}$ and the slope $B_1 = (3.29 \pm 0.03) \times 10^{-12}\,\Phi_0^2/(\mathrm{K}\cdot\mathrm{Hz})$. The fact that

the intercept is compatible with 0 is in full agreement with the fluctuation-dissipation theorem, thus indicating that the system is well thermalized in the restricted high temperature range. According to Eq. (3), the fitting parameters B_0 and B_1 can be used to estimate the residual non-thermal force noise, which reads:

$$S_{F0} = \frac{4k_B k}{\omega_0}\frac{B_0}{B_1}, \tag{5}$$

thus giving $S_{F0} = (-1.51 \pm 1.77) \times 10^{-36}\,\mathrm{N^2/Hz}$. We use the procedure described in Ref. [45] to determine the upper limit on a strictly positive CSL force noise $S_{F0,\mathrm{CSL}} \leq 2.07\times10^{-36}\,\mathrm{N^2/Hz}$ at the 95% confidence level. Note that, according to the form of Eq. (4), any residual effect of saturation in the high temperature data would increase the noise in such a way to increase the value of B_0. Therefore, our estimation of CSL noise should be regarded as conservative.

The corresponding upper bound on λ is derived taking into account the actual geometry and the materials of the whole mechanical resonator, which is composed by the multilayer mass, the magnetic sphere and the cantilever. The contribution from the multilayer mass is largely dominant at $r_c < 10^{-7}\,\mathrm{m}$, and is responsible for a second minimum of the upper bound at $r_c \approx 10^{-7}\,\mathrm{m}$ [40]. The resulting exclusion plot is shown in Fig. 4. The fluctuations in the upper bound due to the uncertainties in the geometry and density of the different subsystems are of the order of the thickness of the curve in Fig. 4 and cannot be appreciated due to the very compressed logarithmic scale.

The current experiment improves by almost two orders of magnitude the previous upper bounds from cantilever experiments at the correlation length $r_c = 10^{-7}\,\mathrm{m}$ [33], and by more than one order of magnitude the final bound from LISA Pathfinder [36]. We are thus substantially challenging the parameter region proposed by Adler [10]. Moreover, the data reveal that the excess noise observed in a previous related experiment [33] is incompatible with a CSL effect for $r_c = 10^{-7}\,\mathrm{m}$. On the other hand, they do not provide substantial new insight into the origin of that excess noise. Indeed, the absolute value of the excess force noise in the previous experiment, featuring the same cantilever, magnet, and SQUID, is compatible with the error bar of the new experiment. The improved bound on the CSL parameters arises entirely from the largest mass load and the specific multilayer structure.

We underline that the strong improvement of the bound at $r_c = 10^{-7}\,\mathrm{m}$ depends on the peculiar features of the CSL model, which make the force noise sensitive to spatial variations of the test mass internal density [40, 46]. Different localization models may lead to different behaviour. For instance, in the Diosi-Penrose model, the force noise is essentially insensitive to the shape and the spatial distribution of the mass [30, 31]. In principle,

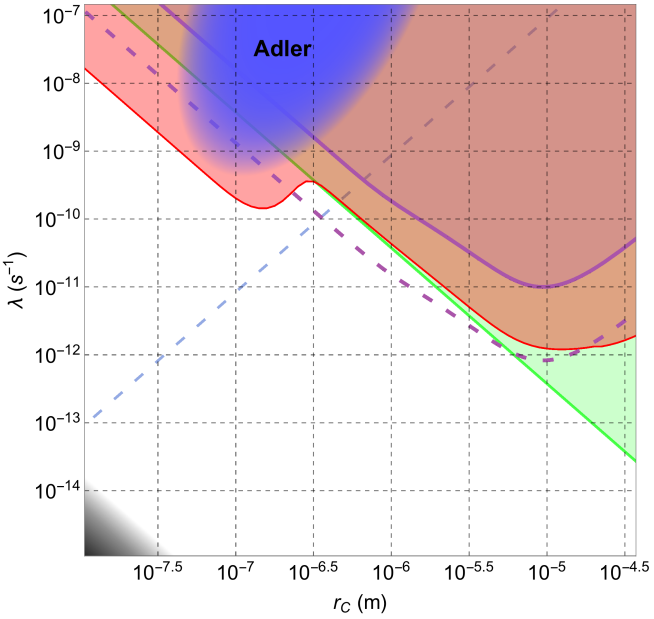

FIG. 4: Exclusion plot for the CSL collapse parameters. Red line and corresponding shaded area: upper bound and excluded region from the present experiment at the 95% confidence level. Green line and corresponding shaded area: upper bound and excluded area from LISA Pathfinder [36]. Violet solid line and corresponding shaded area: upper bound and excluded area from a previous cantilever experiment [33]. Purple dashed line: lower limit of a possible CSL effect from the excess noise observed in the latter experiment [33]. Blue dashed line: upper bound from X-ray emission from a Germanium sample [21]. Since this experiment probes CSL at much higher energies $\sim 10^{19}\,\mathrm{Hz}$, this upper bound is easily evaded by assuming a spectral cutoff of the CSL noise [39]. The blue region represents crude estimations of CSL parameters from Adler, assuming CSL is effective at mesoscopic scale [10]. The grey region at the bottom is a lower bound to the CSL parameters based on theoretical arguments consistent with the initial CSL proposal by Ghirardi *et al.* [9, 51].

specific ad-hoc modifications of the CSL model may lead to a different behaviour as well.

Another point to consider here is that the original estimation of the value of λ by Adler was based on very crude assumptions and analysis. Thus, the proposed parameter space represented by the blue region in Fig. 4 should be taken as indicative [47]. In this sense, a further improvement by at least one order of magnitude, possibly with different experimental techniques, may be needed to provide a strong falsification of CSL under Adler's assumptions.

Despite these caveats, our measurements are clearly reducing the probability that CSL effects will be found at $\lambda \gtrsim 10^{-10}\,\mathrm{Hz}$. Eventually, one should explore the more conservative framework initially proposed by Ghirardi *et al.* [9]. In this case, the CSL effects, if existing, could feature a much lower collapse rate λ. One would need a dramatic improvement of sensitivity and novel experi-

mental techniques to fully probe the entire CSL parameter space. Recent advances in levitation technologies at low frequencies [37, 48, 49] combined with low temperatures and ultrahigh vacuum may be, together with more direct interferometric techniques on earth [18] and in space [50], the most promising route towards this ambitious goal.

We gratefully thank S.L. Adler for many stimulating discussions, and N. Bazzanella for technical help. AB acknowledges hospitality from the Institute for Advanced Study, Princeton where part of this work was done. We acknowledge financial support from the EU H2020 FET project TEQ (Grant No. 766900), the Leverhulme Trust (RPG-2016-046), the COST Action QTSpace (CA15220), INFN and the Foundational Questions Institute (FQXi).

* Electronic address: a.vinante@soton.ac.uk

[1] H.D. Zeh, *On the interpretation of measurement in quantum theory*, Foundations of Physics **1**, 69 (1970).
[2] W.H. Zurek, *Decoherence and the transition from quantum to classical*, Physics Today **44**, 36 (2003).
[3] E. Schrödinger, *Die gegenwärtige Situation in der Quantenmechanik*, Die Naturwissenschaften **23**, 807 (1935).
[4] J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics: Collected papers on quantum philosophy* (Cambridge University Press, Cambridge, 1987).
[5] S. Weinberg, *Collapse of the state vector*, Phys. Rev. A **85**, 062116 (2012).
[6] A.J. Leggett, *Testing the limits of quantum mechanics: motivation, state of play, prospects*, J. Phys.: Condens. Matter **14**, R415 (2002).
[7] M. Arndt and K. Hornberger, *Testing the limits of quantum mechanical superpositions*, Nature Physics **10**, 271 (2014).
[8] G.C. Ghirardi, A. Rimini, and T. Weber, *Unified dynamics for microscopic and macroscopic systems*, Phys. Rev. D **34**, 470 (1986).
[9] G.C. Ghirardi, P. Pearle, and A. Rimini, *Markov processes in Hilbert space and continuous spontaneous localization of systems of identical particles*, Phys. Rev. A **42**, 78 (1990).
[10] S.L. Adler, *Lower and upper bounds on CSL parameters from latent image formation and IGM heating*, J. Phys. A **40**, 2935 (2007).
[11] L. Diosi, *Models for universal reduction of macroscopic quantum fluctuations*, Phys. Rev. A **40**, 1165 (1989).
[12] R. Penrose, *On Gravity's role in Quantum State Reduction*, Gen. Relativ. Gravit. **28**, 581 (1996).
[13] N. Gisin, *Stochastic quantum dynamics and relativity*, Helv. Phys. Acta **62**, 363 (1989).
[14] N. Gisin, *Weinberg's non-linear quantum mechanics and supraluminal communications*, Phys. Lett. A **143**, 1 (1990).
[15] A. Bassi, and G. C. Ghirardi,*Dynamical reduction models*, Phys. Rep. **379**, 257 (2003).
[16] A. Bassi, K. Lochan, S. Satin, T. P. Singh, and H. Ulbricht, *Models of wave-function collapse, underlying theories, and experimental tests*, Rev. Mod. Phys. **85**, 471 (2013).
[17] K. Hornberger, S. Gerlich, P. Haslinger, S. Nimmrichter and M. Arndt, *Quantum interference of clusters and molecules*, Rev. Mod. Phys, **84**, 157 (2012).
[18] Y.Y. Fein, P. Geyer, P. Zwick, F. Kiaka, S. Pedalino, M. Mayor, S. Gerlich, and M. Arndt, *Quantum superposition of molecules beyond 25 kDa*, Nature Physics (2019).
[19] C. Curceanu, B.C. Hiesmayr, and K. Piscicchia, *X-rays Help to Unfuzzy the Concept of Measurement*, J. Adv. Phys. **4**, 263 (2015).
[20] C. Curceanu *et al.*, *Spontaneously emitted X-rays: an experimental signature of the dynamical reduction models*, Found. Phys. **46**, 263 (2016).
[21] K. Piscicchia *et al.*, *CSL Collapse Model Mapped with the Spontaneous Radiation*, Entropy **19**, 319 (2017).
[22] S.L. Adler and A. Vinante, *Bulk heating effects as tests for collapse models*, Phys. Rev. A **97**, 052119 (2018).
[23] M. Bahrami, *Testing collapse models by a thermometer*, Phys. Rev. A **97**, 052118 (2018).
[24] R. Mishra, A. Vinante, T.P. Singh, *Testing spontaneous collapse through bulk heating experiments: An estimate of the background noise*, Phys. Rev. D **98** 052121 (2018).
[25] S.L. Adler, A. Bassi, M. Carlesso, and A. Vinante, *Testing continuous spontaneous localization with Fermi liquids*, Phys. Rev. A **99**, 103001 (2019).
[26] A. Tilloy and T.M. Stace, *Neutron Star Heating Constraints on Wave-Function Collapse Models*, Phys. Rev. Lett. **123**, 080402 (2019).
[27] B. Collett and P. Pearle, *Wavefunction Collapse and Random Walk*, Found. Phys. **33**, 1495 (2003).
[28] S.L. Adler, *Stochastic collapse and decoherence of a nondissipative forced harmonic oscillator*, J. Phys. A **38**, 2729 (2005).
[29] M. Bahrami, M. Paternostro, A. Bassi, and H. Ulbricht, *Proposal for a Noninterferometric Test of Collapse Models in Optomechanical Systems*, Phys. Rev. Lett. **112**, 210404 (2014).
[30] S. Nimmrichter, K. Hornberger, and K. Hammerer, *Optomechanical Sensing of Spontaneous Wave-Function Collapse*, Phys. Rev. Lett. **113**, 020405 (2014).
[31] L. Diosi, *Testing Spontaneous Wave-Function Collapse Models on Classical Mechanical Oscillators*, Phys. Rev. Lett. **114**, 050403 (2015).
[32] A. Vinante, M. Bahrami, A. Bassi, O. Usenko, G. Wijts, and T.H. Oosterkamp, *Upper Bounds on Spontaneous Wave-Function Collapse Models Using Millikelvin-Cooled Nanocantilevers*, Phys. Rev. Lett. **116**, 090402 (2016).
[33] A. Vinante, R. Mezzena, P. Falferi, M. Carlesso, and A. Bassi, *Improved Noninterferometric Test of Collapse Models Using Ultracold Cantilevers*, Phys. Rev. Lett. **119**, 110401 (2017).
[34] M. Carlesso, A. Bassi, P. Falferi, and A. Vinante, *Experimental bounds on collapse models from gravitational wave detectors*, Phys. Rev. D **94**, 124036 (2016).
[35] B. Helou, B. Slagmolen, D.E. McClelland, and Y. Chen, *LISA pathfinder appreciably constrains collapse models*, Phys. Rev. D **95**, 084054 (2017).
[36] M. Carlesso, M. Paternostro, H. Ulbricht, A. Vinante. A. Bassi, *Non-interferometric test of the continuous spontaneous localization model based on rotational optomechanics*, New J. Phys. **20** 083022 (2018).
[37] D. Zheng *et al.*, *Room temperature test of wave-function collapse using a levitated micro-oscillator*, Phys. Rev. Re-

search **2**, 013057 (2020).
[38] M. Bilardello, S. Donadi, A. Vinante and A. Bassi, *Bounds on collapse models from cold-atom experiments* Phys. A **462**, 764 (2016).
[39] M. Carlesso, L. Ferialdi and A. Bassi, *Colored collapse models from the non-interferometric perspective*, Eur. Phys. J. D **72**, 159 (2018).
[40] M. Carlesso, A. Vinante, and A. Bassi, *Multilayer test masses to enhance the collapse noise*, Phys. Rev. A **98**, 022122 (2018).
[41] O. Usenko, A. Vinante, G. Wijts, and T.H. Oosterkamp, *A superconducting quantum interference device based read-out of a subattonewton force sensor operating at millikelvin temperatures*, Appl. Phys. Lett. **98**, 133105 (2011).
[42] F. Pobell, *Matter and Methods at Low Temperatures*, 3rd ed. (Springer, Berlin, 2002).
[43] A. Vinante, G. Wijts, O. Usenko, L. Schinkelshoek, and T.H. Oosterkamp, *Magnetic resonance force microscopy of paramagnetic electron spins at millikelvin temperatures*, Nature Commun. **2**, 572 (2011).
[44] G. Stan, S.B. Field, and J.M. Martinis, *Critical Field for Complete Vortex Expulsion from Narrow Superconducting Strips*, Phys. Rev. Lett. **92** 097003 (2004).
[45] G.J. Feldman and R.D. Cousins, *Unified approach to the classical statistical analysis of small signals*, Phys. Rev. D **57**, 3873 (1998).
[46] L. Diosi, *Two invariant surface-tensors determine CSL of massive body wave function*, arXiv:1908.02195.
[47] S.L. Adler, private communication.
[48] A. Vinante, A. Pontin, M. Rashid, M. Toroš, P.R. Barker, H. Ulbricht, *Testing collapse models with levitated nanoparticles: Detection challenge*, Phys. Rev. A **100**, 012119(2019).
[49] B.R. Slezak, C.W. Lewandowski, J-F. Hsu and B. D'Urso, *Cooling the motion of a silica microsphere in a magneto-gravitational trap in ultra-high vacuum*, New J. Phys. **20**, 063028 (2018).
[50] R. Kaltenbaek *et al*, *Macroscopic Quantum Resonators (MAQRO): 2015 update*, EPJ Quantum Technology, **3**, 5 (2016).
[51] M. Toroš, A. Bassi, *Bounds on quantum collapse models from matter-wave interferometry: calculational details*, J. Phys. A **51**, 115302 (2018).
[52] A. Chiasera, C. Meroni, F. Scotognella, Y.G. Boucher, G. Galzerano, A. Lukowiak, D. Ristic, G. Speranza, S. Valligatla, S. Varas, L. Zur, M. Ivanda, G.C. Righini, S. Taccheo, R. Ramponi, M. Ferrari, *Coherent emission from fully Er3+ doped monolithic 1-D dielectric microcavity fabricated by rf-sputtering*, Optical Materials **87**, 107 (2019).
[53] M. Ketchen *et al.*, *Octagonal washer DC SQUIDs and integrated susceptometers fabricated in a planarized sub-μm Nb-AlO_x-Nb technology*, IEEE Trans. Appl. Supercond. **3**, 1795(1993).
[54] D. Freedman, P.Z. Diaconis, *On the histogram as a density estimator:L2 theory*, Wahrscheinlichkeitstheorie verw. Gebiete **57**, 453 (1981).

Supplemental Material

Sample fabrication

The WO_3/SiO_2 multilayer structure was fabricated by RF sputtering technique. The films were deposited on a silicon substrate. Before sputtering, the substrate was spin-coated with a standard $2.1\,\mu$m thick positive photoresist, in order to enable releasing of samples at the end of the process. The sputtering deposition was performed by alternating two targets of tungsten oxide (WO_3) and silica (SiO_2) both with size $15 \times 5\,\mathrm{cm}^2$. The deposition time necessary to reach the appropriate thickness, are about 2 h for silica layer and 1 h and 5 min for tungsten oxide layers respectively. The residual pressure before the deposition was 2.5×10^{-7} mbar. During the deposition procedure, the substrates were not heated and the temperature of the sample holder during the deposition was 30° C. The sputtering occurred with an Ar gas pressure of 5.4×10^{-3} mbar, the applied RF power was 130 W and 110 W for Silica and Tungsten targets respectively. To monitor the thickness of the layers during the deposition, two quartz microbalances Inficon instruments thickness monitor model SQM-160 faced on the two targets were employed. Thickness monitor was calibrated for the two materials by a long (24 h) deposition process and by directly measuring the thickness of the deposited layer by an m-line apparatus and SEM measurements [52]. The final resolution on the average effective thickness obtained by this quartz microbalance is about 1 Å.

After deposition the multilayer is cut into small rectangular chips with approximate dimensions $80 \times 110\,\mu$m using a dicing saw, and individual chips are finally released from the substrate using acetone. A selected multilayer chip is manually glued to the commercial cantilever (Nanosensors, type TL-CONT-10) using a small amount (< 1 picoliter) of Stycast 2850 epoxy. The ferromagnetic microsphere is glued on the other side of the cantilever (see Fig. 1). The microsphere, with radius $15.5\,\mu$m, is picked from a commercial ferromagnetic powder (Magnequench, type MQP-S-11-9). The dimensions of the multilayer chip, the single layer thickness and the microparticle have been estimated by SEM inspection. The dimensions of cantilever and microsphere have been cross-checked by optical microscopy inspection.

The SQUID is a commercial gradiometric microsusceptometer composed of two spatially separated Nb loops with radius $10\,\mu$m [53]. One loop is used for the cantilever detection, the other is used to apply feedback. The cantilever chip is manually placed above the SQUID (see Fig. 1) with the help of a Macor spacer and firmly held in place by a brass spring. The effective position of the center of the magnetic sphere during the measurements was about $50\,\mu$m above the SQUID loop center. The SQUID is read out by a commercial electron-

ics from Magnicon (model XXF-1) normally operated in flux-locked-loop mode, i.e. with negative feedback.

Experimental procedures

The temperature of the mixing chamber plate is measured by a RuO_2 thermometer calibrated against a superconducting reference point device. The temperature was separately cross-checked against a SQUID-based noise thermometer. The overall accuracy of the temperature measurement is estimated as better than 0.5%.

The mechanical quality factor was measured by ringdown measurements. We distinguish the apparent quality factor Q_a from the intrinsic quality factor Q. During the noise measurements the SQUID is operated in conventional flux-locked-loop mode to provide stable working point and high dynamics. Under these conditions, there is a dynamical change of the quality factor caused by the feedback electronics [33]. We measure Q_a under these conditions. We point out that Q_a appears only in the denominator of the Lorentzian term in Eq. (2) and due to the fitting procedure its precise value does not affect the estimation of the noise parameter B. In contrast, it is very important to determine with good accuracy the intrinsic quality factor Q, which does not include the effects from the feedback electronics, as it determines the strength of thermal noise through Eq. (3). We measure Q by ringdown measurement by operating the SQUID in open loop, with resolution better than 1%. The signal during this measurement is low enough to avoid SQUID nonlinearities. A different procedure involving the measurement of Q under variable feedback gain and extrapolating to infinite gain [33] was found to provide results consistent with the first procedure.

The noise measurements were performed by switching off the pulse tube compressor of the dilution refrigerator, in order to minimize vibrational noise. For each temperature we average a number of periodograms ranging from 40 to 80, corresponding to a total averaging time from 28 to 56 minutes.

Data Analysis

The Density Noise Spectrum (DNS) $\mathcal{S}(\omega)$ is experimentally obtained by averaging a number $n_{\rm av}$ of FFT periodograms of the SQUID signal. The sampling frequency is 100 kHz with a 2^{22} samples for each dataset. The latter are weighted with a Blackman window and a FFT is performed. The average of these blocks gives the DNS with a corresponding error which is set to be equal to $\mathcal{S}(\omega)/\sqrt{n_{\rm av}}$. Finally, the only post-processing of the data is the selection of a window around the peak corresponding to the mechanical motion. Such a window $[3515, 3550]$ Hz is chosen to be the same independently from the temperature at which the experiment is performed.

Although the accuracy of the FFT is high, it is worst by a factor $2 \sim 3$ than the width of the peak. This causes a spectral leakage, whose main effect is to distort the DNS around the peak. An effective way to account for this distortion is to remove from the fit all the data points (actually just 6 over more that thousand point used for the fit) that coincide with twice the width of the main lobe of the spectral leakage power. Outside this region, the signal to noise (due to the spectral leakage) ratio improves by -85 dB. Fig. S1 shows the comparison between the experimental DNS and the power spectrum of the Blackman window for the analysis for $T = 1000$ mK.

Each averaged DNS at given temperature is then fitted with the function [33]

$$S_\Phi = A + \frac{B f_0^4 + C(f^2 - f_1^2)^2}{(f^2 - f_0^2)^2 + \left(\frac{f f_0}{Q_a}\right)^2}, \tag{S1}$$

through ROOT. The fit was implemented by applying a recursive fit. In detail, the initial error $\mathcal{S}(\omega)/\sqrt{n_{\rm av}}$ is substituted at every step with the value of the fitted DNS divided by the square root of $n_{\rm av}$: $\mathcal{S}(\omega)/\sqrt{n_{\rm av}} \to \mathcal{S}_{\rm fit}(\omega)/\sqrt{n_{\rm av}}$. Then the recursive fit is performed until the value of χ^2/d.o.f. approaches an asymptotic value within 0.01%. For every fit, the fixed parameters are $n_{\rm av}$ and Q_a, while A, B, C, f_0 and f_1 are the free parameters of the fit. An example of the fit is shown in the top panel of Fig. S2, where the data for $T = 580$ mK are analyzed.

Once the recursive fit at the asymptotic value of χ^2/d.o.f. near to 1 is found, we implement an additional

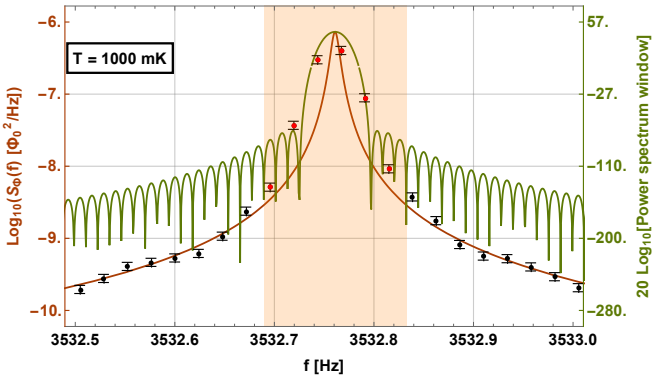

FIG. S1: Comparison of the experimental data (red and black dots with corresponding error bars) of the DNS at $T = 1000$ mK with the power spectrum of the Blackman window (green line), which leads to the spectral leakage. The error bars along f are equal to the distance between two subsequent points, here they are not shown to improve the figure readability. The orange shaded region identifies the frequencies for which the spectrum of the windowing is stronger than -85 dB. All the points in this region, here highlighted in red, are neglected in the subsequent analysis. Thus, only the black points are used for the fit (brown line).

quality check. This is the fit of the normalized residuals distribution with that of a χ^2-distribution with $2n_{\rm av}$ degrees of freedom. This is what one expects from the error distribution of a power spectrum. Here, we applied the Freedman-Diaconis rule [54] to compute the width of the histogram bins.

The resulting values of B are fitted against T/Q, which is the temperature dependent quantity appearing in a thermal force noise $S_{F_{\rm th}} = 4k_B T m\omega_0/Q$. Fig. S3 reports the corresponding experimental data and the functions used for their fit. Although the overall trend of B is proportional to T/Q, the data suffer of a saturation behaviour at low temperatures. This can be fitted with the following function

$$B = B_0 + B_1\frac{T}{Q} + B_2\left[\left(\frac{T}{Q}\right)^n + \left(\frac{T_{co}}{Q}\right)^n\right]^{1/n}, \qquad \text{(S2)}$$

where the term proportional to B_2 describes the deviation from the linear behaviour through the saturation at temperatures below T_{co}. In particular, the choice of $n = 4$ corresponds to a Kapitza contact thermal resistance as dominating thermalization path, which is the saturation process we expect. In the following we will consider $n = 4$. In the high temperature limit, Eq. (S2) reduces to

$$B = B_0 + (B_1 + B_2)\frac{T}{Q} + \mathcal{O}(\tfrac{1}{T^3}), \qquad \text{(S3)}$$

which corresponds to the linear high-temperature behaviour described by the classical fluctuation dissipation theorem. Here, B_0 will be the non-thermal contribution, that could be possibly produced by the CSL mechanism.

Since the precise thermal saturation mechanism at work is not fully understood, one is forced to neglect the low temperature data and restrict the analysis to high temperatures. The distinction between low and high temperatures is made by considering the fit of the experimental data with Eq. (S2), which gives a value for $T_{co}/Q = (5.30 \pm 2.14) \times 10^{-8}\,\mathrm{K}$ corresponding to a temperature just above $T = 85\,\mathrm{mK}$. Thus, all the data corresponding to $T = 100\,\mathrm{mK}$ and above are considered as high-temperature data, which will be used for a linear fit using

$$B = B_0 + B_1 T/Q. \qquad \text{(S4)}$$

This results to $B_0 = (-4.64 \pm 5.31) \times 10^{-21}\,\Phi_0^2/\mathrm{Hz}$ and $B_1 = (3.29 \pm 0.03) \times 10^{-12}\,\Phi_0^2/(\mathrm{Hz\,K})$. The high-temperature linear fit and that with the saturation at low temperatures are reported in Fig. S3.

The upper bound on the residual non-thermal force noise can be computed through Eq. (5), which gives $S_{F0} = (-1.23 \pm 1.44) \times 10^{-36}\,\mathrm{N^2/Hz}$. Such a value is compatible with zero, and has an upper bound of $S_{F0}^{\rm upper} = 1.69 \times 10^{-36}\,\mathrm{N^2/Hz}$ at the 95% of confidence level, which was computed through the methods in [45].

The expression for the CSL induced density force noise spectrum, expressed by Eq. (1), can be computed numerically as in [40]. The mass distribution that has been taken into account includes the cantilever, the multilayer mass and the sphere. The cantilever has density $\rho_C = 2.33 \times 10^3\,\mathrm{kg/m^3}$ and dimensions $450 \times 57 \times 2.5\,\mu\mathrm{m}^3$. The ferromagnetic sphere has density $\rho_S = 7.43 \times 10^3\,\mathrm{kg/m^3}$ and radius $R = 15.5\,\mu\mathrm{m}$. Finally, the multilayer mass has a basis of dimensions $113 \times 82\,\mu\mathrm{m}^2$ and it is made of 47 alternate layers of thickness $274 \pm 8\,\mathrm{nm}$. 24 of them have density $\rho_A = 7.17 \times 10^3\,\mathrm{kg/m^3}$, while the remaining 23 have $\rho_B = 2.2 \times 10^3\,\mathrm{kg/m^3}$. By comparing the numerically computed CSL force noise $S_{F_{CSL}}$ with the experimental upper value of the residual force noise $S_{F0}^{\rm upper}$, one obtains the experimental upper bound on the CSL parameters, which is shown in Fig. 4.

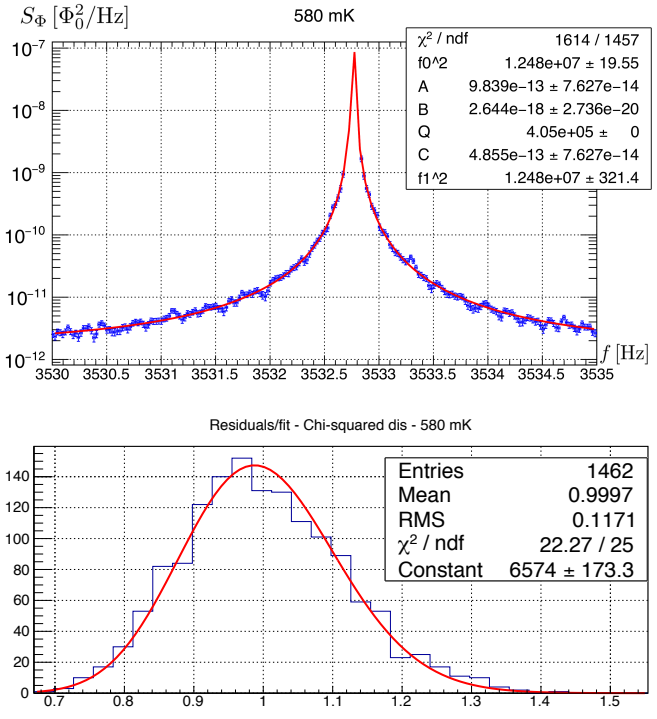

FIG. S2: (**Top panel**) Example of the experimental data (blue point with corresponding error bars) for $T = 580\,\mathrm{mK}$, which are fitted with Eq. (2) (red line). The value of the corresponding χ^2 and of the parameters obtained from the fit are reported in the inset. (**Bottom panel**) The residual distribution (blue histogram) for $T = 580\,\mathrm{mK}$, normalized over the fit, is fitted with a χ^2 distribution (red line).

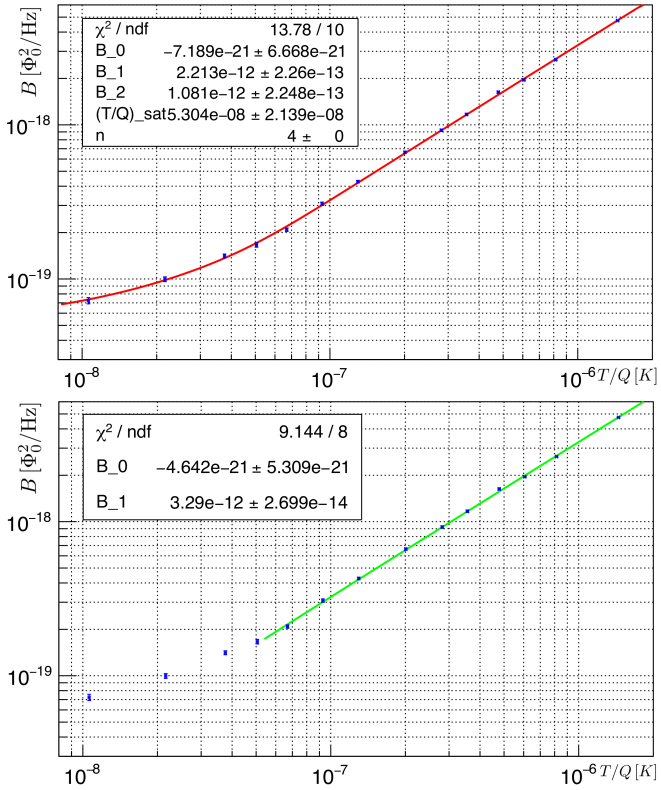

FIG. S3: Experimental data (blue point with corresponding error bars) for the values of B at different values of T/Q compared to two different fits. (**Top panel**) Fit with Eq. (S2) (red line), which corresponds to a thermal saturation at low temperatures. The value of the corresponding χ^2 and of the parameters obtained from the fit are reported in the inset. (**Bottom panel**) High temperature fit with Eq. (S4) (green line). The value of the corresponding χ^2 and of the parameters obtained from the fit are reported in the inset.