A megaelectronvolt emission line in the spectrum of a γ -ray burst

Maria Edvige Ravasio,^{1,2*} Om Sharan Salafia,^{2,5*} Gor Oganesyan,^{3,4*} Alessio Mei,^{3,4} Giancarlo Ghirlanda,^{2,5} Stefano Ascenzi,^{6,7,3,4} Biswajit Banerjee,^{3,4} Samanta Macera,^{3,4} Marica Branchesi,^{3,4} Peter G. Jonker,^{1,9} Andrew J. Levan,^{1,10} Daniele B. Malesani,^{1,11,12} Katharine B. Mulrey,^{1,13} Andrea Giuliani,⁸ Annalisa Celotti,^{14,2,15} Gabriele Ghisellini²

¹Department of Astrophysics/Institute for Mathematics, Astrophysics and Particle Physics Radboud University, Nijmegen 6525 AJ, The Netherlands ² Osservatorio Astronomico di Brera, Istituto Nazionale di Astrofisica, Merate 23807, Italy ³Gran Sasso Science Institute, L'Aquila I-67100, Italy ⁴Laboratori Nazionali del Gran Sasso, Istituto Nazionale di Fisica Nucleare, L'Aquila I-67100, Italy ⁵Sezione di Milano-Bicocca, Istituto Nazionale di Fisica Nucleare, Milano 20146, Italy ⁶Institute of Space Sciences - Consejo Superior de Investigaciones Científicas, Campus Universitat Autònoma de Barcelona, Barcelona E-08193, Spain ⁷Institut d'Estudis Espacials de Catalunya, Barcelona E-08034, Spain ⁸Istituto di Astrofisica Spaziale e Fisica Cosmica, Istituto Nazionale di Astrofisica, Milano I-20133, Italy ⁹Netherlands Institute for Space Research, Leiden 2333 CA, The Netherlands ¹⁰Department of Physics, University of Warwick, Coventry, CV4 7AL, UK ¹¹Cosmic Dawn Center, Denmark ¹²Niels Bohr Institute, University of Copenhagen, Copenhagen N 2200, Denmark ¹³Nationaal Instituut voor Kernfysica en Hoge-Energiefysica, Amsterdam 1098 XG, The Netherlands ¹⁴Scuola Internazionale Superiore di Studi Avanzati, Trieste I-34136, Italy ¹⁵Sezione di Trieste, Istituto Nazionale di Fisica Nucleare, Trieste I-34127, Italy *Corresponding authors; E-mail: mariaedvige.ravasio@ru.nl, om.salafia@inaf.it, gor.oganesyan@gssi.it

Long gamma-ray bursts (GRBs) are observed when the collapse of massive stars produces ultra-relativistic outflows pointed towards Earth. Their gamma-ray spectra are smooth, typically modelled by joint power-law segments describing a continuum, with no detected spectral lines. We report a significant (> 6σ) narrow emission feature at around 10 mega-electron-volts (MeV) in the spectrum of the bright GRB 221009A. Across 80 s, it evolves in energy (~ 12 to ~ 6 MeV), with a constant width of ~ 1 MeV, and in luminosity (~ 1.1 to < 0.43×10^{50} erg/s). We interpret it as a blue-shifted spectral line produced by the annihilation of electron-positron pairs, potentially in the same location responsible for emitting the brightest GRB pulses.

Gamma-ray bursts (GRBs) are transient phenomena appearing as brief (from a fraction of a second up to several hundred seconds) energetic flashes of gamma-rays at energies of kiloelectronvolt to megaelectronvolt (keV – MeV) distributed randomly in the sky. During the intense and highly variable γ -ray radiation phase, termed the prompt emission, an enormous amount of energy is released, typically ~ 10^{52} to 10^{53} erg, given the cosmological distances of GRB sources and assuming the energy is emitted isotropically. Observations and theoretical studies have shown that some GRBs are produced during the formation of a stellar-mass black hole during the collapse of a massive star. The extraction of rotational energy from the black hole powers a relativistic jet: the prompt emission of GRBs is then produced by the conversion of a small fraction of the jet kinetic or magnetic energy into electromagnetic radiation (1, 2).

The physics of the prompt emission is poorly understood: the dominant form of energy in the relativistic jet is unknown, as is the nature of the radiative process responsible for the observed radiation. The gamma-ray spectrum during the prompt emission phase is smooth, typically described using a model consisting of two power-laws with slopes α and β , smoothly connected at a peak photon energy E_{peak} where most of the power is emitted (hereafter referred to as SBPL). For some GRBs, detailed broad-band modelling of the spectral shape, when possible, shows fundamental deviations from that typical double power-law spectrum, such as spectral breaks at low energies (3, 4) or an exponential cutoff at high energies (5, 6), which can potentially provide information about the underlying physical processes.

Fermi/GBM observations of GRB 221009A

On 9 October 2022, the Gamma-Ray Burst Monitor on the Fermi spacecraft (Fermi/GBM) was triggered by GRB 221009A, an extremely bright GRB (with reported fluence $F \sim 0.2 \,\mathrm{erg}\,\mathrm{cm}^{-2}$ (7–10)). The redshift z of the host galaxy of GRB 221009A was measured as z = 0.151 (11). This distance and the flux at the brightest pulse of GRB 221009A result in extreme values for the isotropic equivalent gamma-ray energy $E_{iso} \sim 10^{55}$ erg and peak luminosity $L_{peak,iso} \sim 10^{54}$ erg/s (7–10). Given the extreme brightness of this GRB, most gamma-ray observations, including from Fermi/GBM, are affected by saturation effects, so analysis of Fermi/GBM data taken in time intervals affected by saturation (flagged as bad time interval, BTI) has been discouraged (12).

We investigate the relatively less bright portions of the prompt emission outside the period flagged as BTI over the full spectral range covered by GBM (8 keV - 40 MeV). We performed a time-resolved spectral analysis of Fermi/GBM data, extracting spectra from 0 to 460 s after the GBM trigger time, excluding the BTI from 219 to 277 s (*12, 13*). We find that the spectra at times 280 s to 320 s after the GBM trigger contain a narrow emission feature at around 10 MeV (Figure 1). We fit each of these spectra with a model consisting of a Gaussian, representing the emission feature, superimposed on a smoothly broken power-law (SBPL, (*13*)), representing the typical GRB prompt emission continuum. Inclusion of the Gaussian component substantially improves the fitting residuals, compared to an SBPL-only model, as shown, as an example, in

the two time bins 290-295 s (Figure 1A&C) and 300-320 s (Figure 1B&D). The spectra of the other time bins are shown in Fig S9.

Figure 2A shows the light curve of GRB 221009A, as recorded by one of the sodium iodide (NaI) detectors part of Fermi/GBM, overlain by 13 selected time intervals (listed in Table S1). We chose the time intervals based on the behaviour of the variable emission (ignoring times when the observed emission drops to background levels). We fit the spectra extracted in each time interval using a range of models, including models with and without the Gaussian emission feature (*13*). Figure 2B&C shows the models fitted to the 8 selected time intervals, spanning the first 360 s of emission (excluding the BTI). The model and the observed data are shown in Fig. S9 (*13*). Figure 2B shows the models of the four time intervals before the brightest part of the light curve; in all four of these intervals there is no evidence for the narrow emission feature. Figure 2C shows the models of the four time intervals after the BTI, which all include the narrow feature. The continuum component is similar in both panels (the specific functions used are described in (*13*)).

Statistical significance of the emission feature

We assess the evidence for the additional narrow emission feature using the Akaike Information Criterion (AIC) (14). Including the Gaussian component in the model improves the AIC by Δ AIC = 49 and 141 in the 280-300 s and the 300-320 s time intervals, respectively, favoring its inclusion. To assess the statistical significance of the feature, we performed Monte Carlo simulations to evaluate the probability that such improvement is due to random fluctuations (13). Using a hypothesis testing framework, accounting for the look-elsewhere effect [inherent in a blind search for a feature with a priori unknown properties (15, 16)] and accounting for the 13 time bins analyzed, we estimated the significance as 6.2 σ in the 280-300 s bin and 11 σ in the 300-320 s bin. The combined significance of the feature found in multiple time bins is 13

 σ (13).

In the time intervals 7 and 8, extending from 320 to 360 s, the prompt emission has substantially faded and modelling the spectrum requires an additional power law component, which we attribute to the rising GRB afterglow, as seen before in MeV observations of other bursts (10, 17–19). Although there is tentative evidence for the Gaussian emission feature in intervals 7 and 8, at energies $7.22^{+1.63}_{-1.72}$ MeV and $6.12^{+0.74}_{-0.59}$ MeV respectively (Figure 2C and Figure S9), the additional free parameters required to model the afterglow and the apparently weaker Gaussian imply the AIC test does not favour its inclusion (Δ AIC values of -2 and 0) (Table 1 and (13)). We set 2σ upper limits on the line's luminosity of $< 0.49 \times 10^{50} \,\mathrm{erg \, s^{-1}}$ and $< 0.43 \times 10^{50} \,\mathrm{erg \, s^{-1}}$ in those time intervals.

If we do include the Gaussian feature in the models of intervals 7 and 8, the best-fitting parameters are well constrained (see (13)) and consistent with the trend of shifting towards lower energies and lower fluxes over time (Fig S1 and S9). The peak energy of the Gaussian model component decreases over time, from an initial 12.56 ± 0.30 MeV in interval 5 to $6.12^{+0.74}_{-0.59}$ MeV in interval 8. In the 80 s between time intervals 5 and 8, the luminosity of the emission feature must have decreased by at least a factor of two. There is no change in the best-fitting width of the emission feature. Table 1 lists the best-fitting parameters for the Gaussian emission component in our models of the time intervals 5 to 8.

Analysis of sub-intervals

To investigate the evolution of the emission feature at higher time resolution, we further divided intervals 5 and 6 (280-300 s and 300-320 s, respectively) into several sub-intervals, then repeated our modelling. Interval 5 (280-300 s) has the higher signal-to-noise ratio, so we subdivided it into four bins of 5 s, while interval 6 (300-320 s) was subdivided into two 10 s bins. We detect the emission feature in each of the six finer time intervals, finding it shifts from $14.40^{+0.86}_{-0.87}$ MeV down to $9.77^{+0.42}_{-0.49}$ MeV (Table 1). With the exception of sub-intervals 5.1 and 5.2 (Table 1), inclusion of the Gaussian component in the model is favoured by the AIC test, indicating it is statistically significant in the sub-intervals 5.3, 5.4, 6.1, 6.2 (Δ AIC = 42, 5, 45, 36, respectively). Figure 1 shows the spectrum during sub-interval 5.3. The model and the spectral data in the other sub-intervals are shown in Fig. S10.

Comparison with other studies

The narrow feature is found in data outside the periods flagged as BTI (Fig. 2A) (12). An independent analysis of the same GBM spectral data in a similar time interval, from 277 s to 324 s, found no evidence of instrumental problems at those times (10). Their analysis does show an excess of flux above the fitted continuum around 10 MeV, represented in their case by Band function and an additional power-law [(10), their figure 5]. The longer integration timescale of 46 s adopted in (10) and the time evolution of the feature likely contribute to make this excess appear broader with respect to the width we found in our analysis in 5s, 10s and 20s-long time intervals. We investigated a possible instrumental origin of the emission feature (see (13)), finding no evidence of an instrumental cause for the emission feature in the spectrum of GRB 221009A, or its evolution over time. The feature is found also in the data of the other BGO detector of GBM, with consistent spectral parameters (see (13)). We searched for relevant data from other γ -ray instruments that observed GRB 221009A, but found no usable data covering the relevant time period and photon energies(see (13)).

Previous studies have reported evidence for absorption or emission features in other GRBs, but none were statistically significant (> 5σ). During the prompt emission phase, absorption lines at 30-70 keV and emission lines at 400-460 keV were reported for multiple bursts by the Konus experiment onboard the *Venera* 11 and 12 missions (20) and by the Ginga satellite (21) (which was interpreted as blueshifted atomic emission line (22)). The search for lines, mostly in absorption and < 100 keV (23) found no detection of any spectral line in a sample of 192 bursts (24). A possible transient Fe absorption feature has been reported during the prompt emission phases of GRB 990705 (25) and GRB 021211, with significance of 2.8 to 3.1 σ (26). Line searches in the afterglow emission phase revealed possible features in the soft X–ray data of BeppoSAX (27, 28), ASCA (29), XMM-Newton (30, 31) and Chandra (32, 33). However a reanalysis of those cases (34) has argued that the significances were over-estimated and no lines were detected. An extensive search for emission or absorption lines in X-ray spectra of GRB afterglows (35) also did not find any statistically significant feature.

GRB 221009A differs from those cases in that we find a bright and narrow, statistically significant emission line at several MeV energies and in the Fermi/GBM spectral data of a GRB. The extraordinary brightness of this GRB led to a high signal-to-noise ratio in the GBM detectors, allowing the emission feature to be detected. We tested this by simulating spectra similar to those observed in interval 6 (when the feature is found with the highest significance) but with progressively lower flux, finding that a line flux 20-40 times lower would not have been detected above the noise (see (*13*) and Figure S6).

We investigated whether a similar emission feature could have been detected in other bright GRBs. We used the GBM data for the three next brightest GRBs in the energy band 10-1000 keV, GRB 130427A, GRB 160625B and GRB 230307A (see (*13*) for details). For each of these GRBs, we extracted the spectrum corresponding to the peak of the lightcurve and between 2 and 8 spectra during the decaying phase of the pulse with the highest count rate. Although the fluxes of these spectra are 1.4 to 29 times higher than interval 6 for GRB 221009A (in the energy range 10 keV–40 MeV) (Figure S7), none of the spectra analyzed shows evidence for an emission feature with similar width and flux. We conclude that a similar emission feature would have been detectable in those three GRBs, but was not present around the time of peak brightness. We extended the search to other three GRBs following different selection criteria

(i.e. most favorable source inclination angle with the detectors, see (13) for the details), finding no statistically significant excess over the continuum model.

Interpretation of the emission feature

A transient, narrow emission feature at MeV energies is not predicted by the standard prompt emission models (1, 36–38). We explored several potential explanations for its presence, all assuming that the narrow spectral component is produced within the GRB jet. The large bulk velocity of GRB jets imply a very low baryon content, so we do not expect such baryons to participate in significant nucleosynthesis. Instead, baryons remain in the form of free protons, deuterium and, at most, α -particles (39). This prevents the production of observable narrow lines by, for instance, fluorescent recombination within the jet. Cold electrons within the jet could conceivably interact with nearly monochromatic photons from some sort of 'narrow line region' that surrounds the progenitor, up-scattering them by means of bulk Comptonization (40, 41), which would result in a blue-shifted and Doppler boosted emission line. We investigated this possibility but found theoretical difficulties with it (see Supplementary Text).

More naturally, a narrow spectral feature could arise in the form of a blue-shifted electronpositron pair annihilation line. The physical conditions required to efficiently produce electronpositron pairs within the jet are probably reached in regions where energy dissipation processes (internal shocks and/or magnetic reconnection events) take place (42). We estimate that during the brightest pulse in GRB 221009A a sufficient number of electron-positron pairs could have formed through two-photon annihilation within a region of the jet moving at a moderate bulk Lorentz factor $\Gamma \sim 20$ and located at $R \sim 10^{15}$ cm from the central engine (see (13)). The annihilation of electron-positron pairs then produces a spectral feature with duration, luminosity and spectrum consistent with that observed. The moderate Lorentz factor $\Gamma \sim 20$ is required to place the line at ~10 MeV as observed. This value is lower than the typical $\Gamma \gtrsim 100$ expected in powerful GRB jets, and previous estimates of $\Gamma \sim 200 - 1000$ for this particular burst (10). Regions with lower Γ could arise temporarily (see (13)) during the collision of a very fast portion of the jet with a slower one, as required for efficient energy dissipation in the leading GRB prompt emission mechanisms (1, 38).

A slight modification of such a scenario, which would allow for a larger Lorentz factor (possibly more in line with the expectations, given the large luminosity) and would naturally accommodate the relatively fast evolution of the narrow feature, is one where the pair annihilation line sweeps the GBM band during the steep decline of one of the brightest pulses of the GRB, due to the high-latitude emission effect (43-45) (see supplementary text).

References

- 1. M. J. Rees, P. Meszaros, Astrophys. J. 430, L93 (1994).
- 2. R. Sari, R. Narayan, T. Piran, Astrophys. J. 473, 204 (1996).
- G. Oganesyan, L. Nava, G. Ghirlanda, A. Celotti, *Astronomy and Astrophysics* 616, A138 (2018).
- M. E. Ravasio, G. Ghirlanda, L. Nava, G. Ghisellini, *Astronomy and Astrophysics* 625, A60 (2019).
- 5. Fermi Large Area Telescope Team, et al., Astrophys. J. 754, 121 (2012).
- 6. G. Vianello, et al., Astrophys. J. 864, 163 (2018).
- 7. D. Frederiks, et al., Astrophys. J. 949, L7 (2023).
- 8. Z.-H. An, et al., arXiv e-prints p. arXiv:2303.01203 (2023).

- 9. E. Burns, et al., Astrophys. J. 946, L31 (2023).
- 10. S. Lesage, et al., Astrophys. J. 952, L42 (2023).
- 11. D. B. Malesani, et al., arXiv e-prints p. arXiv:2302.07891 (2023).
- 12. Fermi Collaboration, https://fermi.gsfc.nasa.gov/ssc/data/analysis/grb221009a.html.
- 13. Materials and methods are available as supplementary materials .
- 14. H. Akaike, IEEE Transactions on Automatic Control 19, 716 (1974).
- 15. K. P. Burnham, D. R. Anderson, Sociological Methods & Research 33, 261 (2004).
- 16. E. Gross, O. Vitells, European Physical Journal C 70, 525 (2010).
- 17. P. Meszaros, M. J. Rees, Astrophys. J. 405, 278 (1993).
- 18. R. Sari, T. Piran, R. Narayan, Astrophys. J. 497, L17 (1998).
- 19. M. E. Ravasio, et al., Astronomy and Astrophysics 626, A12 (2019).
- E. P. Mazets, S. V. Golenetskii, R. L. Aptekar, I. A. Gurian, V. N. Ilinskii, *Nature* 290, 378 (1981).
- 21. T. Murakami, M. Fujii, K. Hayashida, M. Itoh, J. Nishimura, Nature 335, 234 (1988).
- 22. C. J. Hailey, F. A. Harrison, K. Mori, Astrophys. J. 520, L25 (1999).
- 23. D. L. Band, et al., Astrophys. J. 458, 746 (1996).
- 24. D. M. Palmer, et al., Astrophys. J. 433, L77 (1994).
- 25. L. Amati, et al., Science 290, 953 (2000).

- 26. F. Frontera, et al., Astrophys. J. 616, 1078 (2004).
- 27. L. Piro, et al., Astronomy and Astrophysics S. 138, 431 (1999).
- 28. L. A. Antonelli, et al., Astrophys. J. 545, L39 (2000).
- 29. A. Yoshida, et al., Astrophys. J. 557, L27 (2001).
- 30. J. N. Reeves, et al., Nature 416, 512 (2002).
- D. Watson, J. N. Reeves, J. Hjorth, P. Jakobsson, K. Pedersen, Astrophys. J. 595, L29 (2003).
- 32. L. Piro, et al., Science 290, 955 (2000).
- 33. N. R. Butler, et al., Astrophys. J. 597, 1010 (2003).
- 34. M. Sako, F. A. Harrison, R. E. Rutledge, Astrophys. J. 623, 973 (2005).
- 35. S. Campana, et al., Astronomy and Astrophysics 592, A85 (2016).
- 36. G. Drenkhahn, H. C. Spruit, Astronomy and Astrophysics 391, 1141 (2002).
- 37. D. Lazzati, B. J. Morsony, M. C. Begelman, Astrophys. J. 700, L47 (2009).
- 38. B. Zhang, H. Yan, Astrophys. J. 726, 90 (2011).
- 39. A. M. Beloborodov, Astrophys. J. 588, 931 (2003).
- 40. M. Sikora, M. C. Begelman, M. J. Rees, Astrophys. J. 421, 153 (1994).
- 41. M. Vietri, G. Ghisellini, D. Lazzati, F. Fiore, L. Stella, Astrophys. J. 550, L43 (2001).
- 42. A. Pe'er, E. Waxman, Astrophys. J. 613, 448 (2004).

- 43. P. Kumar, A. Panaitescu, Astrophys. J. 541, L51 (2000).
- 44. G. Oganesyan, et al., Astrophys. J. 893, 88 (2020).
- 45. S. Ascenzi, et al., Astronomy and Astrophysics 641, A61 (2020).
- 46. C. Meegan, et al., Astrophys. J. 702, 791 (2009).
- 47. https://heasarc.gsfc.nasa.gov/W3Browse/fermi/fermigbrst.html .
- 48. https://fermi.gsfc.nasa.gov/ssc/data/analysis/GBM_caveats.html .
- 49. https://fermi.gsfc.nasa.gov/ssc/data/analysis/scitools/gtburst.html .
- 50. https://heasarc.gsfc.nasa.gov/xanadu/xspec/.
- 51. D. Gruber, et al., ApJS 211, 12 (2014).
- 52. M. E. Ravasio, et al., Astronomy and Astrophysics 613, A16 (2018).
- 53. B. P. Gompertz, et al., Nature Astronomy 7, 67 (2023).
- 54. G. Ghisellini, A. Celotti, D. Lazzati, Mon. Not. R. Astron. Soc. 313, L1 (2000).
- 55. F. Daigne, Ž. Bošnjak, G. Dubus, Astronomy and Astrophysics 526, A110 (2011).
- 56. J. Yang, et al., Astrophys. J. 947, L11 (2023).
- 57. A. Ursi, et al., GRB Coordinates Network 32650 (2022).
- 58. G. Piano, et al., GRB Coordinates Network 32657 (2022).
- 59. M. Tavani, et al., Astrophys. J. 956, L23 (2023).

- G. Fitzpatrick, S. McBreen, V. Connaughton, M. Briggs, Space Telescopes and Instrumentation 2012: Ultraviolet to Gamma Ray, T. Takahashi, S. S. Murray, J.-W. A. den Herder, eds. (2012), vol. 8443 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, p. 84433B.
- G. Fitzpatrick, V. Connaughton, S. McBreen, D. Tierney, *arXiv e-prints* p. arXiv:1111.3779 (2011).
- 62. https://heasarc.gsfc.nasa.gov/W3Browse/fermi/fermigdays.html .
- 63. D. Foreman-Mackey, D. W. Hogg, D. Lang, J. Goodman, PASP 125, 306 (2013).
- 64. https://fermi.gsfc.nasa.gov/ssc/data/analysis/user/.
- 65. E. Bissaldi, et al., Experimental Astronomy 24, 47 (2009).
- 66. R. Preece, et al., Science 343, 51 (2014).
- 67. G. Breit, J. A. Wheeler, *Physical Review* 46, 1087 (1934).
- 68. R. J. Gould, G. P. Schréder, *Physical Review* 155, 1404 (1967).
- 69. J. M. Jauch, F. Rohrlich, Texts and Monographs in Physics (Springer, ed. 2, 1976).
- 70. M. J. Rees, P. Mészáros, Astrophys. J. 628, 847 (2005).
- 71. K. Ioka, K. Murase, K. Toma, S. Nagataki, T. Nakamura, Astrophys. J. 670, L77 (2007).
- 72. K. Murase, K. Ioka, Astrophys. J. 676, 1123 (2008).
- 73. S. M. Rahaman, J. Granot, P. Beniamini, Mon. Not. R. Astron. Soc. 528, 160 (2024).
- 74. R. Svensson, Astrophys. J. 258, 335 (1982).

- 75. A. M. Beloborodov, Mon. Not. R. Astron. Soc. 407, 1033 (2010).
- 76. R. Sari, T. Piran, Astrophys. J. 455, L143 (1995).
- 77. F. Daigne, R. Mochkovitch, Mon. Not. R. Astron. Soc. 336, 1271 (2002).
- 78. J. H. Shiode, E. Quataert, Astrophys. J. 780, 96 (2014).
- 79. A. Panaitescu, Astrophys. J. 895, 39 (2020).
- S. Salafia, G. Ghisellini, A. Pescalli, G. Ghirlanda, F. Nappo, *Mon. Not. R. Astron. Soc.* 450, 3549 (2015).



Figure 1: **GBM spectra of GRB 221009A in the time intervals 290-295 s and 300-320 s. A and B:** The GBM spectra (data points) in the time interval 290-295 s (interval 5.3, panel A) and in the time interval 300-320 s (interval 6, panel B) are fitted with a standard GRB empirical continuum model (SBPL). The narrow feature appears as an excess around $\sim 12 - 10$ MeV. Data are from GBM three sodium iodide (NaI) detectors (light blue, yellow and purple crosses, see legend) and its Bismuth Germanate (BGO) detector (orange crosses). **C and D:** Same as panels A and B, but with a model consisting of the SBPL (black dotted line) and an additional Gaussian emission component (black dashed lines). In all panels, data points have been rebinned for graphical purposes. Error bars represents the 1σ uncertainty on data points, upper limits at 3σ are represented by arrows. Black lines indicate the best-fitting model and gray shading is its 1σ uncertainty intervals. Residuals between the data and model are shown below each panel.



Figure 2: **GBM light curve and time-resolved spectra of GRB 221009A.** A: Count rate light curve of GRB 221009A (blue solid line) in the energy band 8-900 keV. Labelled regions are the 13 time intervals we analyse (separated by dashed black vertical lines) and the BTI (12) excluded from the analysis due to detector saturation (gray shading, see text). Eight time intervals are colour-coded (colors match the other panels). B: Solid curves show the best-fitting models and shaded areas their 90% credible intervals, in νF_{ν} representation, for time intervals 1, 2, 3 and 4 (colors match panel A). Each model fits the data with only the typical continuum for GRB prompt emission (the specific functions used are described in (13), parameters listed in Table S1). C: Same as panel B, but for time intervals 5, 6, 7 and 8. These models include a Gaussian emission component, with dashed lines, in addition to the continuum (parameters listed in Tables 1 and S1). The SBPL is represented with dotted lines, and the power-law model (PL), representing the afterglow, with dash-dotted lines. In interval 7 and 8, the Gaussian is not statistically significant, and its 2σ upper limits are represented with dashed lines and downward arrows. See Figure SS9 for the version of this plot showing the models and data overlaid, with each interval in a separate panel.

Table 1: Spectral parameters of the Gaussian feature. The best-fitting parameters of the Gaussian emission feature in our models are listed for time intervals 5 to 8. For intervals 5 and 6 we also list results for shorter sub-intervals (see text). L_{gauss} is the feature's luminosity, E_{gauss} is its central photon energy, and σ_{gauss} is its width. The uncertainties are reported at 1σ level, and the upper limits on the luminosity when the feature is not statistically significant are quoted at the 2σ level. ΔAIC is the change in AIC when adding the Gaussian component to the model. The corresponding best-fitting parameters for the continua in each model are listed in Table S1.

Time interval [s]	Interval number	$\frac{L_{\rm gauss}}{[10^{50} \text{ erg s}^{-1}]}$	$E_{ m gauss}$ [MeV]	$\sigma_{ m gauss}$ [MeV]	ΔAIC
280 to 300	5	$1.12_{-0.19}^{+0.19}$	$12.56_{-0.31}^{+0.30}$	$1.31_{-0.30}^{+0.31}$	49
280 to 285 285 to 290 290 to 295	5.1 5.2 5.3	< 1.5 < 0.99 $1.84^{+0.36}_{-0.33}$	$14.4_{-0.87}^{+0.86}$ $13.2_{-1.5}^{+6.4}$ $12.2_{-0.3}^{+0.3}$	$\begin{array}{c} 0.99\substack{+0.66\\-0.57}\\ 1.14\substack{+0.59\\-0.62}\\ 1.08\substack{+0.34\\-0.30\\-0.30\end{array}$	2.4 - 1.2 42
295 to 300 300 to 320	<u> </u>	$\frac{0.63^{+0.28}_{-0.27}}{1.14^{+0.20}_{-0.18}}$	$\frac{12.55^{+0.47}_{-1.4}}{10.19^{+0.29}_{-0.28}}$	$\frac{0.79^{+0.81}_{-0.45}}{1.70^{+0.52}_{-0.42}}$	5 141
300 to 310 310 to 320	6.1 6.2*	$\begin{array}{c} 1.08\substack{+0.19\\-0.17}\\ 0.75\substack{+0.21\\-0.19} \end{array}$	$\begin{array}{c} 10.42\substack{+0.31\\-0.30}\\ 9.77\substack{+0.42\\-0.49}\end{array}$	$1.14_{-0.29}^{+0.36}\\1.24_{-0.21}^{+0.25}$	45 30
320 to 340 340 to 360	7* 8*	< 0.49 < 0.43	$7.2^{+1.6}_{-1.7} \\ 6.12^{+0.74}_{-0.59}$	$2.38^{+0.45}_{-0.83}\\1.35^{+1.1}_{-0.74}$	$-2 \\ 0$

^{*}These spectra require an extra power-law component, representing the afterglow (see text).

Acknowledgments

We thank the anonymous referees for providing constructive and valuable comments that helped improve the manuscript. We acknowledge Dr. Felicia Carla Tiziana Barbato for providing useful insights on the BGO detectors. This research has made use of data obtained through the High Energy Astrophysics Science Archive Research Center (HEASARC) Online Service provided by the NASA/Goddard Space Flight Center.

Funding MER acknowledges support from the research programme Athena with project number 184.034.002, which is financed by the Dutch Research Council (NWO). OS acknowledges funding by the European Union-Next Generation EU, PRIN 2022 RFF M4C21.1 (202298J7KT - PEACE). SA is supported by the H2020 ERC Consolidator Grant "MAGNESIA" under grant agreement No. 817661 (Principal Investigator: Rea) and National Spanish grant PGC2018-095512-BI00. AJL and DBM are supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 725246). DBM thanks the Cosmic Dawn Center, supported by the Danish National Research Foundation under grant DNRF140, for the kind hospitality. GO and MB acknowledge financial support from the European Union's Horizon 2020 Programme under the AHEAD2020 project (grant agreement n. 871158). BB and MB acknowledge financial support from MIUR (PRIN 2017 grant 20179ZF5KS).

Authors contributions MER extracted the Fermi data, modelled the background, performed the spectral analysis, produced most figures, and led the writing of the text. OSS provided the theoretical interpretations of the spectral feature, investigated the statistical significance derived from the simulations, performed the orbital background test and contributed parts of the text. GO provided the theoretical interpretations of the spectral feature and contributed parts of the text. AM extracted and analyzed the Fermi data for the three other GRBs and contributed parts of the text. GiG extracted and analyzed the Fermi data for GRB 221009A and some of the bright bursts, performed the spectral simulations and the orbital background test and contributed parts of the text. SA contributed to the theoretical interpretations. BB, SM, MB PGJ, AJL, DBM, KBM and AG discussed the results and contributed to manuscript preparation. AC and GaG contributed to the direction of the theoretical interpretations and manuscript preparation.

Competing interests The authors declare no conflicts of interest.

Data and materials availability The *Fermi* data are available from the HEASARC archive at https://heasarc.gsfc.nasa.gov/db-perl/W3Browse/w3table.pl?tablehead= name%3Dfermigbrst&Action=More+Options, and can be retrieved using the GRB name. The results of our model fitting are listed in Tables 1 and S1. The code used to derive the background using the orbital method is publicly available at https://github.com/ omsharansalafia/GBM_bkg_orbital_method.

Supplementary Materials

Materials and Methods Supplementary Text Figs. S1 to S11 Tables S1 References (45–80)

Supplementary materials for

A megaelectronvolt emission line in the spectrum of a γ -ray burst

Maria Edvige Ravasio, Om Sharan Salafia, Gor Oganesyan, Alessio Mei, Giancarlo Ghirlanda, Stefano Ascenzi, Biswajit Banerjee, Samanta Macera, Marica Branchesi, Peter G. Jonker, Andrew J. Levan, Daniele B. Malesani, Katharine B. Mulrey, Andrea Giuliani, Annalisa Celotti, Gabriele Ghisellini Corresponding authors: mariaedvige.ravasio@ru.nl, om.salafia@inaf.it, gor.oganesyan@gssi.it

The PDF file includes:

Materials and Methods Supplementary Text Figs. S1 to S11 Tables S1 References

Materials and Methods

Data analysis The GBM instrument is composed of 12 NaI detectors (sensitive to photons over the energy range of 8 keV to 900 keV) and two bismuth germanate (BGO, 300 keV to 40 MeV) scintillation detectors (*46*). Following the recommendations of the instrument team (*12*), we initially analysed the data from two NaI, namely NaI 8 and NaI 4, and one BGO detector, BGO 1. These NaI detectors registered the highest count rates and at the time of the trigger they observed the source at an angle of less than 60° (35° and 51°, respectively), for which the effective area is maximised. The BGO 1 detector has a smaller viewing angle (80°) than BGO 0. At a later stage, we also analysed data from one additional NaI detector, NaI 6 (source viewing angle of 46°), and the second BGO 0 (100°), to investigate the observed spectral shape (see below).

We retrieved spectral data files and the corresponding response matrix files (rsp2) from the HEASARC online archive (47). Following the standard GBM data analysis procedure, we selected energy channels in the range 10 to 900 keV for NaI detectors, and 0.3 to 40 MeV for BGO detectors, and excluded channels in the range 25 to 45 keV from our analysis due to the presence of the Iodine K-edge at 33.17 keV (48). We also excluded energy channels corresponding to the 45–90 keV energy range from the NaI4 detector data and to the 10–45 keV from the NaI8 detector, due to their systematically different behaviour with respect to the other NaI detectors. We used inter-calibration constant factors among NaI and BGO detectors, scaled to the most illuminated NaI and allowed to vary within 30%. We used the CSPEC data, which have 1024 ms time resolution. To model the background, we selected different time intervals (see also below) before and after the burst and fitted them with a polynomial function up to the fourth order. Spectra were extracted using the public software GTBURST (49) and analysed with XSPEC (50). We used PG-Stat, valid for Poisson data with a Gaussian background (50), as our goodness of fit statistic.

We extracted time-resolved spectra excluding the BTI (12). Before the start of the BTI at 219 s, we selected four time intervals with ~ 10 s width, according to the pulse trend in the light curve and excluding the quiescent time (Fig. 2A). To assess the spectral evolution after the BTI end at t = 277 s, we extracted a sequence of 20-second-wide time bins (also shown in Fig. 2A), up to 460 s after the trigger time. The time intervals selected are reported in Table S1.

Model comparison We fitted the extracted spectra with five different models. The simplest is the smoothly broken power-law (SBPL) model, which is composed of two power laws with low-energy photon index α and high-energy photon index β , smoothly connected at one break energy, representing the peak energy E_{peak} in the νF_{ν} representation. The SBPL is a standard function used to model the non-thermal spectral shape of the GRB prompt emission (51).

The second model is a double smoothly broken power-law (2SBPL), a modified version of the SBPL model which allows for the presence of an additional spectral break at low-energy. The 2SBPL is composed of three power-law segments (with photon indices α_1 , α_2 and β) smoothly connected at two break energies, E_{break} and E_{peak} . Previous studies have shown 2SBPL model is a better representation of the data than SBPL for bright GRBs (4, 52, 53).

In cases where the SBPL or 2SBPL models did not adequately fit the spectrum, as indicated by a poor goodness of fit statistic and/or by the presence of systematic trends in the residuals, we added a Gaussian component, an extra power law component (PL), or both. We also investigated using a blackbody function to fit the excess emission at MeV energies. However, a blackbody is too broad to fit the emission feature. We compared the different models through the AIC, and considered a more complex model as being statistically preferred over a simpler one whenever the difference in AIC was > 4 (15). To determine uncertainties, we explored the parameter space using a Markov Chain Monte Carlo (MCMC) approach, using the chain command in XSPEC. We adopt the median values of the marginalized posterior probability densities as the best-fitting parameters. The values and the 1σ uncertainties for each model parameter are reported in Table 1 and Table S1. Their evolution over time is shown in Figure S1.

In the time intervals analyzed before the BTI, up to 216 s (intervals1, 2, 3 and 4), there is no evidence for the presence of narrow features. In the first time interval, corresponding to the first pulse of the light curve, the spectrum is well fitted using a simple SBPL. In this time bin, the low-energy photon index $\alpha = -1.68^{+0.01}_{-0.01}$ is close to the value expected for synchrotron emission in fast cooling regime ($\alpha_{syn} = -1.5$, (54)) and there is no evidence for an additional spectral break in the continuum at low energies. In the following three time intervals, which are more than an order of magnitude brighter, the 2SBPL model provides a statistically significant improvement over the simpler SBPL model. The break energy slowly evolves from ~ 230 keV to ~ 115 keV over 20 s. Overall, we find the prompt emission spectra at T< 216 s to be consistent with synchrotron from non-thermal relativistic electrons in a marginally fast cooling regime (55). Independent modelling of time-resolved spectra of GRB 221009A also found them to be consistent with synchrotron radiation from non-thermal electrons (56).

After the BTI ends at 277 s, GBM data are again usable for spectral analysis. The spectrum we extracted from 280 to 300 s (interval 5) shows a strong excess in the BGO data with respect to the SBPL model. We modelled this excess by adding a Gaussian to the SBPL model, finding $E_{gauss} = 12.56^{+0.30}_{-0.31}$ MeV with a width $\sigma_{gauss} = 1.31^{+0.31}_{-0.30}$ MeV. The addition of the Gaussian is strongly favoured by $\Delta AIC = 49$ over the simpler SBPL model. To check if this feature is also present on shorter timescales and study its evolution with time, we further split this 20 s time interval into four 5-second-long time bins (intervals 5.1, 5.2, 5.3 and 5.4, see Table 1). The addition of the Gaussian line component improves the goodness of fit statistic in all of the four shorter-integration spectra analysed, but its presence is statistically significant only in spectra 5.3 and 5.4, with $\Delta AIC = 42$ and 5, respectively. Figure 1A&C shows the spectrum during time interval 5.3, and the rest of the spectra of these sub-intervals are shown in Figure S10.

In the following time bin (300-320 s, interval 6), the best-fitting model is the SBPL+Gaussian, with the inclusion of the Gaussian line significantly required by $\Delta AIC = 141$. By the end of interval 6, the prompt emission spectrum has decreased by more than 75% in luminosity, leaving the line as the highest peak in the GBM energy range (the line luminosity is similar to what was found in the previous time bin). The central energy of the Gaussian has decreased to $10.19^{+0.29}_{-0.28}$ MeV, while its width is consistent with the value found in the previous time bin. As with interval 5, we performed a finer time-resolved analysis by splitting this bin into two 10-second-long finer bins (intervals 6.1 and 6.2 in Table 1). In both finer time intervals, the prompt spectrum shows higher flux between the peak energy and the Gaussian line (where previously the spectrum had a typical decaying trend with a photon index β between -3 and -2.5). The best-fitting SBPL+Guassian model has a flatter high-energy photon index $\beta \sim -2$. We therefore tested the addition of a PL to the SBPL+Gaussian model in both time intervals. The AIC indicates this PL is not required in the 6.1 time bin, which is instead well modelled by SBPL+Gaussian. However, the spectrum in the second time bin 6.2 required the SBPL+Gaussian+PL model (with a $\Delta AIC = 12$ with respect to the SBPL+Gaussian model). The addition of the PL allows the model to better fit the spectrum between the two peaks and yields a β parameter similar to the previous values ($\beta = -2.17 \pm 0.06$). The PL has a photon index $\Gamma_{PL} = 1.85^{+0.03}_{-0.02}$ and a luminosity in the 10 keV–40 MeV energy band $L_{\rm PL} = 0.42^{+0.09}_{-0.14} \times 10^{51}$ erg/s. We ascribe this to the rising afterglow component.

In the next two time intervals, 320-340 s and 340-360 s (intervals 7 and 8), there is still a small excess with respect to the progressively fainter prompt emission spectrum. The addition of a Gaussian to the model is still favored over the SBPL alone ($\Delta AIC = 53$ and 91, for intervals 7 and 8, respectively). However, given the similar high flux between peak energy and the Gaussian line and the presence of the PL in the previous time bin, we tested its inclusion also in these two

time intervals. We found that the SBPL+PL model better fits the flux between the two peaks ($\Delta AIC = 82$ and 172 with respect to the SBPL model, for intervals 7 and 8, respectively). The AIC values indicate the SBPL+PL is the best-fitting model for both spectra, which is favored over the SBPL+Gaussian model. In both time intervals 7 and 8, the PL has a photon index consistent with that found in interval 6.2 ($\Gamma_{PL} = -1.85$; -1.84, respectively) and a similar luminosity in the 10 keV-40 MeV energy band ($L_{PL} = (0.54 \text{ to } 0.46) \times 10^{51} \text{ erg/s}$). Although the Gaussian line is not required ($\Delta AIC = 0 - 2$) by the data over the SBPL+PL model in these two time intervals, the highly significant detection of the line in intervals 5 and 6 provides support that the weaker excess in intervals 7 and 8 might indicate the line presence. If the line is present, the fit with SBPL+PL+Gaussian yields well-constrained parameters, suggesting its central energy has decreased down to $E_{gauss} = 7.22^{+1.63}_{-1.72}$ MeV and $E_{gauss} = 6.12^{+0.74}_{-0.59}$ MeV, in the 7th and 8th intervals, respectively. We set $2-\sigma$ upper limits on the line luminosity during these two time intervals of $L_{gauss} < 0.49 \times 10^{50} \text{erg s}^{-1}$ and $L_{gauss} < 0.43 \times 10^{50} \text{erg s}^{-1}$, respectively, revealing a fading by at least 50% with respect to the previous time intervals.

From 360 s onwards (intervals 9, 10, 11, 12, and 13), there is no evidence for the presence of the line. The four spectra, covering the time interval 360-460 s, are each best fitted by the SBPL model. The luminosity of the prompt spectrum is steadily increasing during this period, although in interval 9 it is similar to interval 6, so we would expect to detect the emission line, if it were present and bright enough. The low-energy photon index α of the SBPL model is -1.42 to -1.64, as observed in the previous time intervals, thus not requiring the presence of an additional spectral break at low energy.

Statistical significance of the line In order to assess the statistical significance of the narrow feature, we tested the hypothesis \mathcal{H}_1 that the data is described by the background model plus an SBPL continuum and a Gaussian feature against the null hypothesis \mathcal{H}_0 that only the

background and the SBPL continuum are needed, using the spectra of intervals 5 (280 - 300 s) and 6 (300 - 320 s). This test determines the probability (p-value) that the test statistic λ exceeds the value that we obtained in our analysis under the assumption that \mathcal{H}_0 holds. The test statistic is defined as

$$\lambda = -2\ln\left(\frac{\max_{\theta_1} \mathcal{L}(\theta_1 \mid \mathcal{H}_1)}{\max_{\theta_0} \mathcal{L}(\theta_0 \mid \mathcal{H}_0)}\right),\tag{S1}$$

where θ_1 and θ_0 represent the parameter vectors for \mathcal{H}_1 and \mathcal{H}_0 , respectively, and $\mathcal{L}(\theta_i | \mathcal{H}_i)$ represents the data likelihood given the model implied by hypothesis \mathcal{H}_i . This hypothesis testing approach, which should not be confused with a likelihood ratio test, accounts for the lookelsewhere effect (the 'trials factor') inherent in a search for a feature whose properties are not known a priori in an observed spectrum, as long as the maximisation of $\mathcal{L}(\theta_1 | \mathcal{H}_1)$ over $\theta_1 =$ $(\alpha, \beta, E_{\text{peak}}, n_{\text{SBPL}}, E_{\text{gauss}}, \sigma_{\text{gauss}}, n_{\text{gauss}})$ (where n_{SBPL} and n_{gauss} are the normalisations of the SBPL and Gaussian spectral components, respectively) is performed over a wide parameter space that includes all values of E_{gauss} and σ_{gauss} that would have been considered acceptable. In our case, we have $-2 \ln \mathcal{L} = \text{PGstat}$, where the quantity on the right-hand side is the goodness of fit statistic for a Poisson source with a Gaussian background, output by XSPEC. The test statistic is the difference between the minimum goodness of fit statistic that can be attained under each of the hypotheses, $\lambda = \min_{\theta_0} \text{PGstat}(\theta_0 | \mathcal{H}_0) - \min_{\theta_1} \text{PGstat}(\theta_1 | \mathcal{H}_1) \equiv$ ΔPGstat . To estimate the statistical significance of the emission line, we need to determine the probability distribution $p(\lambda | \mathcal{H}_0)$ of λ under the hypothesis \mathcal{H}_0 , then determine the p-value

$$p(\lambda > \Delta PGstat | \mathcal{H}_0) = \int_{\Delta PGstat}^{\infty} p(\lambda | \mathcal{H}_0) d\lambda.$$
(S2)

The integral can be computed using a Monte-Carlo approach, by simulating a large number N_{sim} of realizations of the spectrum under hypothesis \mathcal{H}_0 , so that

$$p(\lambda > \Delta PGstat | \mathcal{H}_0) \sim \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} \Theta(\lambda_i - \Delta PGstat),$$
 (S3)

where

$$\Theta(x) = \begin{cases} 0 & x < 0\\ 1 & x \ge 0 \end{cases}$$
(S4)

is the Heaviside step function and λ_i is the test statistic in the *i*-th realization. For this approximation to be valid, the number of simulated spectra $N_{\rm sim}$ must be much larger than the reciprocal of the p-value that is to be tested; to demonstrate a 5 sigma significance, $N_{\rm sim} \gg 10^6$. However such a large number of simulations is not actually necessary (16), because the p-value is asymptotically equal to

$$p(\lambda > \Delta PGstat | \mathcal{H}_0) \approx p(\chi_s^2 > \Delta PGstat) + \mathcal{N}p(\chi_{s+1}^2 > \Delta PGstat)$$
 (S5)

for $\Delta PGstat \gg s$, where s is the difference in the number of free parameters between \mathcal{H}_1 and \mathcal{H}_0 , $p(\chi_d^2 > x)$ is the probability that a stochastic variable with a chi-square distribution with d degrees of freedom has a value larger than x, and \mathcal{N} is a constant that depends on the specific statistical model. In our case, s = 3 is the number of parameters of the Gaussian spectral component, so that the asymptotic equality S5 is valid for $\Delta PGstat \gg 3$. To determine the value of the \mathcal{N} constant, we use the Monte Carlo method described below.

Starting from the best-fitting SBPL model of the spectrum (either interval 5 or 6), we simulated $N_{\rm sim} = 8 \times 10^4$ spectra (using the fakeit command of XSPEC) of the source and the corresponding background spectra. We considered the same detectors as used for the data analysis, namely NaI 4, 6, 8 and BGO 1, and the corresponding detector response matrices used for the analysis of the real spectra. To determine $\max_{\theta_1} \mathcal{L}(\theta_1 | \mathcal{H}_1)$, we then refit each simulated spectrum with the SBPL and SBPL+Gaussian models. All the parameters of the fitting models were left free to vary over wide ranges: the Gaussian line energy $E_{\rm gauss} \in [1, 100]$ MeV, width $\sigma_{\rm gauss} \in [0.01, 10]$ MeV and normalization $n_{\rm gauss} \in [0, 10^{24}]$ cm⁻² s⁻¹. To ensure that the true maximum Δ PGstat was found, we repeated the SBPL + Gauss fit 11 times for each spectrum realization, initializing each time the fitting procedure from different initial Gaussian line energy values. We used the resulting distribution of test statistics ΔPG stat to estimate a an upper limit on the p-value at a reference test statistic value ΔPG stat = 5 × s = 15, that is $p(\lambda > 15 | \mathcal{H}_0) < C_{sim}$, where

$$C_{\rm sim} = \frac{N_{\rm sim}(\Delta PG \text{stat} > 15) + 3\sqrt{N_{\rm sim}(\Delta PG \text{stat} > 15)}}{N_{\rm sim}}$$
(S6)

is the Monte Carlo p-value estimator plus three times the Poisson standard deviation. We then set N so that the asymptotic form of the p-value equals C_{sim} at $\Delta PGstat = 15$:

$$\mathcal{N} = \frac{C_{\rm sim} - p(\chi_3^2 > 15)}{p(\chi_4^2 > 15)}.$$
(S7)

Figure S2 shows the p-value as a function of ΔPG stat as estimated through our Monte Carlo simulations along with the extrapolation defined in Eq. S5 for each of the two spectra and the ΔPG stat attained in the analysis of the real data. For interval 5, this is ΔPG stat = 55, which gives an extrapolated p-value of 4.7×10^{-11} , corresponding to 6.6σ significance; for interval 6, ΔPG stat = 147, which gives a p-value of 2×10^{-30} , corresponding to 11.5σ significance.

This methodology accounts for the lack of a priori expectations for the energy and width of the line, but it does not address the increased probability of a false positive when multiple time bins are tested. We focus on the N = 13 twenty-second-long time bins over which we carried out our spectral analyses. The post-trials p-value in the *i*-th time bin after a search carried out on N time bins is p_iN , where $p_i = p(\lambda_i > \Delta PGstat_i | \mathcal{H}_0)$ is the pre-trials p-value. This reduces the post-trials significance of the Gaussian line to around 6.2σ in interval 5, and to around 11.2σ in interval 6. Because we found a large improvement in the goodness of fit for two neighbouring time bins, we calculate the probability of finding such improvement by chance in two bins. The combined post-trials p-value for an improvement ($\Delta PGstat_i, \Delta PGstat_j$) in any two time bins (i, j) was computed by re-interpreting the search as if it were carried out on time bin couples rather than on single time bins. The number of distinct couples of time bins is M = N(N - 1)/2, and the pre-trials p-value for a deviation in two time bins is $p_{i,j} = p_i p_j$. The post-trials p-value is therefore $p_{i,j}M = N(N-1)p_ip_j/2$. Applying this correction to the two time bins 5 and 6 where we find the largest improvement, we obtain a post-trials p-value $p_{5,6}M = p_5p_6N(N-1)/2 \approx 7.3 \times 10^{-39}$, which corresponds to 13σ significance. This analysis shows that the Guassian emission component is statistically significant in intervals 5 and 6, even after accounting for trials factors.

Search for potential instrumental effects GRB 221009A has been observed by several γ -ray instruments and in almost all of them the brightness of the burst induced severe instrumental effects due to saturation and deadtime.

Of those instruments, only Konus-Wind (7) and AGILE (57, 58) observe the γ -ray sky extending to the energy range where we identify the Gaussian feature in Fermi/GBM spectral data. We therefore investigate whether those datasets contain evidence of the emission line. The detector S2 on the Konus-Wind satellite was affected by saturation, and the time bin 249-257 s was the latest time interval analyzed in previous work (7). Unfortunately, no further spectral information is publicly available for the emission beyond 257 s. Prompted by a pre-print report of our findings, the Konus-Wind team analyzed their data of the brightest pulses of the burst (i.e. prior to when we observed the line in Fermi/GBM), which have been corrected for saturation and pileups. They found (7) that in the time interval 225-233 s there is a slight but not statistically significant ($\leq 2\sigma$) excess in the residuals over the fitted continuum around 15 MeV. Fitting the excess with a Gaussian (with fixed 1 MeV width), they constrained the luminosity of a putative feature at that position to be below $L_{gauss} < 1.7 \times 10^{52}$ erg/s at 2σ .

The AGILE/MCAL instrument observes in the energy range 0.4-100 MeV. A previous analysis of those data (*59*) found that in the time interval relevant for the feature presence (the first part of their interval c, from 273.01 to 393.61 s, see their figure 1 and table 1) there are no data available from MCAL detector (due to saturation), precluding any spectral analysis. The team responsible for the GBM instrument has performed checks on the reliability of data collected by GBM and has released caveats about their use, flagging specific time intervals as BTI (bad time interval) and discouraging their analysis (*12*). In performing our spectral analysis, we selected time intervals either before or after the BTI and analyzed them with the standard procedures and software. We also performed the following tests to investigate whether an instrumental artefact could mimic a Gaussian line at MeV energies.

We first visually inspected the background spectra extracted during the time intervals showing the emergence of the line. Figure S4 A&C show the spectra accumulated by the BGO1 detector during the time interval 280-300 s and 300-320 s, respectively. The background counts spectra (black crosses) do not show the presence of the same excess that is visible in the total observed data (source+background, orange crosses) at \sim 10 MeV. We also visually inspected the response matrices used in the same time intervals, and they also did not show any unusual behaviour at MeV energies. We therefore exclude the possibility that the line is produced by any artefact in either the background spectrum or the response matrix.

We also visually inspected the observed count rate spectra, which are model-independent, namely they are independent from the modelling with an empirical model (e.g. SBPL+Gaussian) and can be thought of as raw data. Figure S4 C&D show the count rate spectra of the source with the background subtracted. The excess at $\sim 10-12$ MeV is evident in the counts rate spectra, in both time intervals considered.

The selection of time intervals for the background subtraction can also have an impact on the source spectra analyzed. The background time intervals selected for each detector were as follows. For NaI4: -126 to -19 s, 1724 to 2056 s, 2101 to 2334 s; for NaI6: -309 to -35 s, 1015 to 1433 s; for NaI8: -285 to -23 s, 50 to 107, 1490 to 1671; for BGO 1: -428 to -10 s, 1670 to 2104 s; for BGO 0: -423 to -19, 1592 to 1707 s. Although the background time interval selec-

tion for NaI6 includes the final phase of fading source activity observable in the other two more on-axis detectors (NaI8 and NaI4), we checked that the average count rate for this more off-axis detector in this time interval is consistent with the pre-burst selected time interval, indicating it is background-dominated. To test the presence of the line, we extracted the spectra with a different selection of the time windows for the background spectrum computation. For this test, we estimated the background for BGO 1 in the following time intervals: -125.571 to -19.072, 1723.904 to 2055.687, 2100.744 to 2334.221 s. After checking these revised background spectra and response matrices for the presence of possible features in the MeV energy range, we found that the source spectra still show the excess at MeV energies, confirming the presence of the Gaussian line in the BGO data.

We also performed a previously-proposed test (60, 61), by using the count rates recorded during the 30th orbit preceding and following the observation, as a proxy of the source-free background count rates. The choice of these orbits is because the spacecraft is almost identically oriented, with respect to the sources that produce the background (60). We retrieved the Fermi/GBM count rate data for the dates 2023-10-07, 2023-10-08, 2023-10-09, 2023-10-10 and 2023-10-11 from the daily catalog (62). We determined the average orbit duration as follows. We define $\rho(t)$ as the total count rate (sum of the count rates in all channels) in the BGO 1 detector at time t, as recorded in the daily data files. The auto-correlation of ρ with a displacement Δt is

$$\mathcal{A}(\Delta t) = \frac{\int \left(\rho(t) - \langle \rho \rangle\right) \left(\rho(t + \Delta t) - \langle \rho \rangle\right) \,\mathrm{d}t}{\int \left(\rho(t) - \langle \rho \rangle\right)^2 \,\mathrm{d}t},\tag{S8}$$

where the integrals extend over the interval $(\min(t), \max(t) - \Delta t)$. We found that \mathcal{A} is maximised for $\Delta t = 171090.0$ s, which corresponds to an average orbital duration of 95.05 minutes over the considered period, assuming the auto-correlation peak corresponds to a shift of 30 orbits. Using this average orbital time, we estimated the average background count rate at a time offset by ± 30 orbits from that of our intervals 5 and 6, that is, 280-300 s and 300-320 s post-trigger, where the line is found with the highest significance. We fitted the ansatz empirical model

$$b(t, \rho_0, \vec{a}) = \rho_0 \exp\left(a_1(t - t_0) + a_2(t - t_0)^2 + a_3(t - t_0)^3 + a_4(t - t_0)^4\right), \quad (S9)$$

where $a_1, ..., a_4$ are free parameters, to the count rate light curve in each of the j = 1, ..., 128energy channels of the BGO 1, with $t_0 = 100 \text{ s} \pm \Delta t$ as reference time, assuming Poissonian statistics for the binned counts, which translates into a log-likelihood (defined up to an additive constant)

$$\ln p(\{c_{i,j}\}_{i=1}^{N} \mid \rho_{0}, \vec{a}) = \sum_{i=1}^{N} c_{i,j} \ln (b(t_{i}, \rho_{0}, \vec{a})\delta t_{i}) - b(t_{i}, \rho_{0}, \vec{a})\delta t_{i},$$
(S10)

where $c_{i,j}$ is the number of counts in energy channel j and in time bin i (which corresponds to the observing time t_i), and δt_i is the width (exposure time) of the *i*-th time bin, for a total of N time bins extending up to $t_1 = 500 \,\mathrm{s} \pm \Delta t$. We performed Bayesian model fitting, by defining the posterior probability $p(\rho_0, \vec{a} \mid \{c_{i,j}\}_{i=1}^N) \propto p(\{c_{i,j}\}_{i=1}^N \mid \rho_0, \vec{a}) \pi(\rho_0, \vec{a})$, adopting a flat prior $\pi(\rho_0, \vec{a}) \propto 1$, and sampling the posterior probability density for each channel j = 1, ..., 128using the emcee software (63). For each channel, the posterior probability density of the estimated background rate in the relevant time bin is the posterior predictive distribution of the mean of the average background rates at ± 30 orbits. To compute this, we took M = 1000samples from the posterior probability density of the +30 orbits background model parameters, and an equal number of samples from the posterior probability density of the -30 orbits model parameters. From these, we computed the corresponding samples $\{b_{\pm 30,j,k}\}_{k=1}^{M}$ of the modelled average background rate in channel j at ± 30 orbits, using Equation S9, and used them to compute M samples of the estimated background rate in $counts s^{-1} keV^{-1}$ in our actual time bin as $\hat{b}_{j,k}/\Delta E_j = (b_{+30,j,k} + b_{-30,j,k})/2\Delta E_j$, for k = 1, ..., M, where ΔE_j is width of channel j in terms of photon energy. We then estimated the background rate in the channel from the median of the resulting samples, and the uncertainty from the $16^{\rm th}$ and $84^{\rm th}$ percentiles of the samples. The result is shown in Fig. S4A&C (red error bars), which is consistent with the background model used in our analysis above (black error bars). As a further test, we also used the software osv1.3 (64), with consistent results (blue crosses in Figure S4 A& C). We therefore conclude that our results are unaffected by the method we used to estimate the background.

In general, if multiple detectors observed the burst and a peculiar spectral feature is found in one of them, then it is worth testing the presence of a such feature by comparing the results of each detector. In fact, if the feature were a statistical fluctuation, having the same fluctuation in a different detector is improbable. The other BGO detector on Fermi, BGO 0, is mounted on the opposite side of the spacecraft with respect to BGO 1 (source viewing angle = 80°). Nevertheless, the GRB brightness was sufficient for BGO 0 to detect the burst (with a viewing angle of 100°) and register a large count rate. We analysed the spectra of BGO 0 extracted over the same time intervals as we used for BGO 1. As for BGO 1 detector, we estimated the orbital background counts rate spectra using the 30th orbit method (*60, 61*), finding consistent results with our estimate of the background during the time when the line appears (Figure S5A).

Figure S5 shows the spectral data in counts rate (Panel A and B) and in the νF_{ν} representation (Panel C) for both BGO detectors, together with the best-fitting models, corresponding to the time interval 280-300 s. The BGO 0 data are consistent with those of BGO 1 and confirm the presence of the line. The line, present in both BGO detectors and required by $\Delta AIC = 146$, has a fitted position of $E_{gauss} = 13.05^{+0.26}_{-0.24}$ MeV with $\sigma_{gauss} = 1.78^{+0.27}_{-0.25}$ MeV, which is consistent with the results we found for BGO 1 alone. We also fitted the spectrum only considering BGO 0, finding that the line is required by $\Delta AIC = 85.8$. The background fitting procedure was performed independently for each of the two BGO detectors (different time intervals and different order of the fitted polynomial). From this analysis, we consider that an un-modelled background effect or the same spurious spectral feature with consistent energy, width and luminosity, that also evolves in the same way, in two BGO detectors very unlikely. We observe the line becoming softer in energy and dimmer in luminosity by a factor ~ 2 and ~ 5 respectively, over ~ 80 s. While this evolution could be produced in a rapidly evolving astrophysical system as a GRB, it is less likely to occur for an instrumental artifact. As a comparison, the iodine K-edge instrumental feature present in GBM data at 33.17 keV (65), which we excluded from the spectral analysis, does not evolve in energy, nor in normalization, over time.

In conclusion, given that time intervals showing the presence of the line are safely placed outside those affected by pileup and saturation effects, and given that our tests described above returned no evidence for any potential instrumental or statistical issue producing the line, we are confident in the astrophysical origin of this feature.

Comparison with other bright GRBs The brightness of GRB 221009A provided high signal-to-noise data at MeV energies, enabling our detection of the statistically significant emission line and asses its temporal evolution. Similarly high quality BGO data is only available for bright GRBs. The typically lower BGO data quality for other GRBs may have prevented previous detections of this line, even if it were a common feature in GRBs. We investigated this by performing simulations of the spectrum from interval 6 (300-320 s), where the line is detected with highest significance, to determine the minimum flux it would have required to be detected (setting a detection threshold of $\Delta AIC > 4$, roughly corresponding to one-sigma significance). These simulations were performed using the fakeit routine in XSPEC and adopting the best-fitting model and parameters of interval spectrum 6, namely SBPL+Gaussian, but reducing the normalization by a factor of 2, 10, 20, 50, 80, and 100. For each value of the normalization, we simulated 100 realizations of the spectrum then fitted each simulated spectrum with both the SBPL and SBPL+Gaussian models to compute the AIC. Figure S6 shows the 16th to 84th percentiles of the resulting ΔAIC as a function of the flux reduction factor. We find that the sig-

nificance of the narrow feature decreases rapidly with the flux and that a reduction in brightness by a factor between 20 and 40 would make the feature undetectable.

This shows that the best candidate GRBs to search for a similar narrow feature are extremely bright bursts. Therefore, we searched for evidence of it in the three next-brightest (in terms of fluence) GRBs observed by Fermi/GBM (fluence of $F = 0.04 \,\mathrm{erg cm^{-2}}$, in the energy range 10-1000 keV). This value is the one reported in the Catalog (47), and it is not affected by the revised fluence of GRB 221009A published in (9), which is much larger (F = $0.2 \,\mathrm{erg/cm^2}$). These GRBs are GRB 230307A ($F = 3.15 \times 10^{-3} \,\mathrm{erg/cm^2}$), GRB 130427A ($F = 2.46 \times 10^{-3} \,\mathrm{erg/cm^2}$) and GRB 160625B ($F = 0.64 \times 10^{-3} \,\mathrm{erg/cm^2}$). A posteriori, we find their spectra (see below) have a flux in the energy range 10 keV–40 MeV from 1.4 to 29 brighter than interval 6 of GRB 221009A used for the simulations.

Figure S6 compares the count rate light curves of the selected GRBs. Fig. S6A shows the light curve observed by the most illuminated NaI detector for each burst (between 8 and 900 keV), while Fig. S6B shows the data for the most illuminated BGO (between 300 keV and 40 MeV). Without knowing the underlying mechanism of the line, it is difficult to predict where to expect it, but given the observed position of the line in our burst (~ 10 MeV), the search is most promising in bursts with counts rates in the BGO data as high as in GRB 221009A. For each burst we extracted the spectrum at the peak of the BGO light curve, and a few spectra (at least 2, depending on the light curve shape) during the steep decay following the peak.

The brightness of GRB 130427A also caused pile-up effects in the detectors after 2.4 s from the trigger time (*66*). Although this time interval does not include the main peak of the light curve in the BGO data, nor the decaying part of it, we restricted the extraction of spectra to the first 2.4 s. We found no evidence for a line-like excess at high energies in these data: the spectrum is well fitted by the 2SBPL model, with spectral indices consistent with synchrotron

radiation in marginally fast cooling regime. For both GRB 230307A and GRB 160625B, we did not find any evidence of a line at high energies both in the peak spectrum and in the spectra analysed during the decaying phase. In those two GRBs, the spectra are fitted by the 2SBPL model. In one time interval of GRB 160625B, 200.74 to 204.83 s, there is a hint of an excess around 16 MeV, but the addition of a Gaussian to the continuum is not statistically significant ($\Delta AIC = 0.47$).

We extended the search to three additional bright GRBs with a viewing angle from the axis of the BGO detectors $\theta_{BGO} < 40^{\circ}$, selecting the 3 brightest GRBs of this sub-sample, GRB 170409A, GRB 171227A and GRB 130606B. We considered time-bins around the main pulse peak and the following decaying phase in their light curves. We extracted between 3 and 6 spectra for each GRB, within temporal bins of 2 to 13 s. We fitted these spectra with the 2SBPL model in the same way as for GRB 221009A. We do not find any statistically significant excess over the continuum model at high energies or a systematic trend in the residuals. In one time interval including the peak of GRB 170409A (32-38 s), there is a hint of an excess around ~ 20 MeV (Figure S11), but the addition of a Gaussian component is not statistically significant ($\Delta AIC = 3.85$).

We stress that our search was focused on a few bright GRBs, likely representing the best candidates to find a line (if present) in the spectrum, rather than being a systematic search throughout the Fermi Catalog.

Theoretical interpretation Electron-positron pairs can be formed within a GRB jet: in the co-moving frame of the plasma, a fraction of the photons produced as a consequence of energy dissipation (internal shocks and/or magnetic reconnection, (1, 38)) is above the photonphoton annihilation (67–69) threshold $h\nu > m_ec^2$ (where h, ν, m_e and c are Planck's constant, the photon frequency, the electron rest mass and the speed of light, respectively) and can therefore form electron-positron pairs (70–72). This process has been discussed in detail (42) within an internal shock scenario under the assumption that dissipation takes place at radii where the jet is optically thin to Thomson scattering off baryon-associated electrons. Following (42), we define the compactness parameter (primed quantities are measured in the jet comoving frame)

$$\ell' = \frac{\sigma_{\rm T} \epsilon_{\pm} L \Delta'}{4\pi R^2 \Gamma^2 m_{\rm e} c^3},\tag{S11}$$

where $\sigma_{\rm T}$ is the Thomson cross section, L is the observed GRB luminosity, R is the radius (radial distance from the jet central engine) of the dissipation region within the jet, Γ is its bulk Lorentz factor, $\Delta' = \xi R/\Gamma$ is its co-moving width (ξ is a dimensionless parameter, which is typically $\xi \sim 0.1 - 1$ if the region of interest is downstream of an internal shock (73)) and ϵ_{\pm} is the fraction of the GRB luminosity in photons that are above the pair-production threshold in the co-moving frame. Adopting the notation $Q_x \equiv Q_{\rm cgs}/10^x$, where Q is any quantity and $Q_{\rm cgs}$ is its value in cgs units, Eq. S11 can be rewritten as

$$\ell' = 8.5 \times 10^3 \xi_{-0.5} \epsilon_{\pm,-1} L_{54} R_{15}^{-1} \Gamma_{1.3}^{-3}, \tag{S12}$$

where $\Gamma_{1.3} \approx \Gamma/20$. This implies that a region moving with $\Gamma \sim 20$ within the jet, at a radius $R \sim 10^{15}$ cm, that was illuminated by photons produced during the very bright peak of the emission of GRB 221009A, had a very large compactness parameter ℓ' . In such a condition, pairs are copiously and continuously created so their number density n'_{\pm} is set by the balance between the creation and annihilation rate (42, 70):

$$n'_{\pm} \sim \ell'^{1/2} \Gamma / \sigma_{\rm T} R \approx 2.8 \times 10^{12} \xi_{-0.5}^{1/2} \epsilon_{\pm,-1}^{1/2} L_{54}^{1/2} R_{15}^{-3/2} \Gamma_{1.3}^{-1/2} \,{\rm cm}^{-3}.$$
 (S13)

Such a pair number density causes the shell to become optically thick to Thomson scattering: indeed, the optical depth due to Thomson scattering off pairs within the shell,

$$\tau_{\mathrm{T},\pm} \sim \sigma_{\mathrm{T}} n'_{\pm} \xi R / \Gamma \approx 29 \xi_{-0.5}^{3/2} \epsilon_{\pm,-1}^{1/2} L_{54}^{1/2} R_{15}^{-1/2} \Gamma_{1.3}^{-3/2}, \qquad (S14)$$

is larger than unity. Pair-annihilation photons, on the other hand, can escape the shell if they are produced within a small external layer of thickness equal to the photon mean free path. For photons of energy $h\nu' = m_ec^2$, the mean free path is $\lambda' \sim (\sigma_{\gamma\gamma}n'_{\gamma} + \sigma_{T}n'_{\pm})^{-1} \sim \Delta'/\eta_{\gamma\gamma}\ell'$, where n'_{γ} is the number density of target photons for two-photon annihilation, $h\nu'_{\text{target}} \sim m_ec^2$, the quantity $\eta_{\gamma\gamma} = \sigma_{\gamma\gamma}/\sigma_{T} \sim 0.1$ (74) is the ratio of the (angle-averaged) Breit-Wheeler cross section $\sigma_{\gamma\gamma}$ to the Thomson cross section, and the approximation is valid when the mean free path for Breit-Wheeler annihilation is shorter than that for Thomson scattering, which is the case for our reference parameters. In this regime, the escaping luminosity L_{\pm} in pair-annihilation photons is a fraction λ'/Δ' of the luminosity $\epsilon_{\pm}L$ that is converted into pairs,

$$L_{\pm} \sim (\lambda'/\Delta') \epsilon_{\pm} L \sim 1.2 \times 10^{50} \eta_{\gamma\gamma,-1}^{-1} \xi_{-0.5}^{-1} R_{15} \Gamma_{1.3} \,\mathrm{erg \, s^{-1}}, \tag{S15}$$

and it does not depend on $\epsilon_{\pm}L$: the more pairs are produced, the more photons are generated when they annihilate, but this also produces a larger optical depth, cancelling out the dependence on $\epsilon_{\pm}L$. After the end of the bright pulse, when the compactness parameter drops, the already present electron-positron pairs continue annihilating over an observer-frame annihilation time scale

$$t_{\rm ann} \sim n_{\pm}' / \Gamma \dot{n}_{\pm,\rm ann}' \sim 1 / \Gamma \sigma_{\rm T} c n_{\pm}' \approx 0.9 \, \xi_{-0.5}^{-1/2} \epsilon_{\pm,-1}^{-1/2} L_{54}^{-1/2} R_{15}^{3/2} \Gamma_{1.3}^{-1/2} \, {\rm s} \tag{S16}$$

(where $\dot{n}'_{\pm,\text{ann}} \sim \sigma_{\text{T}} c {n'}^2_{\pm}$ is the pair annihilation rate per unit volume in the co-moving frame). Because the 'angular' time scale (related to the spread in the arrival times of photons produced within an angle $\theta \leq \Gamma^{-1}$ of the line of sight) $t_{\text{ang}} \sim R/\Gamma^2 c \approx 83 R_{15}\Gamma_{1.3}^{-2}$ s is much longer, $t_{\text{ang}} \gg t_{\text{ann}}$, an observer would see the line fade away over the latter time scale, which is similar to the delay of about 100 s between the brightest emission pulse (which peaks around the start of the BTI, 220 s after the GBM trigger) and the latest time of appearence of the line in our observations (the end of interval 6, 320 s after the GBM trigger). Before annihilating, pairs cool due to inverse Compton interaction with the abundant photons present, over an observer-frame time scale

$$t_{\rm cool} \sim \frac{3\pi m_{\rm e} c^2 \Gamma R^2}{\sigma_{\rm T} L} = 2.3 \times 10^{-4} R_{15}^2 \Gamma_{1.3} L_{54}^{-1} \,\mathrm{s},$$
 (S17)

which is much shorter than the annihilation time scale. This causes the annihilation line to be narrow, with an intrinsic relative width $\Delta \nu / \nu$ much smaller than unity (?). The observed width $\sigma_{\text{gauss}}/E_{\text{gauss}} \lesssim 0.1$ (see Table 1) agrees with this expectation, despite the possible broadening due to the evolution in the frequency of the line over the exposure time. The differs, for example, from the case of pairs formed due to collisional heating within the jet in the acceleration phase described in (75), which would produce a much broader line that might not be distinguishable from the continuum.

These analytical estimates lead us to interpret the feature we observe as due to pairs formed during the brightest emission phase of GRB 221009A, within a shell of material moving with a bulk Lorentz factor $\Gamma \sim 20$, that was located at a radius $R \sim 10^{15}$ cm from the central engine and participated in the production of the prompt emission photons. Such a slow shell is consistent with the internal shock scenario: we consider a jet shell with kinetic luminosity $L_{kin,4}$ and Lorentz factor Γ_4 that collides with another shell with luminosity $L_{kin,1} \ll L_{kin,4}$ and Lorentz factor $\Gamma_1 \ll \Gamma_4$ (in this paragraph, numerical subscripts indicate different shells as labelled in Figure S3). The collision produces a forward shock (FS) that propagates from the contact discontinuity (CD) into shell 1, and a reverse shock (RS) that propagates backwards (as seen by an observer co-moving with the CD) into shell 4. We identify the region between the CD and the FS as region 2, and the region between the RS and the CD as region 3. Regions 2 and 3 move at approximately the same speed, with a Lorentz factor Γ such that $\Gamma_4 \gg \Gamma \gg \Gamma_1$. We estimate the value of Γ by imposing pressure balance between regions 2 and 3 at the CD (76), finding $\Gamma \sim (L_{kin,4}/L_{kin,1})^{1/4}\Gamma_1$. Because a high efficiency of conversion of kinetic energy requires a large Lorentz factor contrast Γ_4/Γ_1 , a low $\Gamma_1 < 20$ would be a favourable condition to produce the high luminosity we observed. Hence, the Lorentz factor $\Gamma \sim 20$ provides a possible explanation for the observed luminosity. Given the observer-frame dynamical time $t_{\rm dyn} \sim R/\Gamma^2 c \approx 83 R_{15}\Gamma_{1.3}^{-2}$ s, such a region does not need to be long-lived, but it can disappear as soon as the forward shock crosses shell 1 and still produce the line over a time similar to that we observe.

Within this scenario, the observed time evolution of the line energy could be ascribed to a variation in the ratio $L_{\text{kin},4}/L_{\text{kin},1}$ during the propagation of the FS and RS, because the observed line photon energy is $E_{\pm} \sim \Gamma m_{\text{e}}c^2$. The luminosity evolution could additionally be due to the decay in the number density of pairs as they annihilate while the shell expands.

The two alternative scenarios we explored are described in Supplementary Text.

Supplementary Text

Alternative interpretations Above we detailed our preferred interpretation of the line as due to annihilation of electron-positron pairs. Here we consider alternative physical mechanisms. The first is an intrinsically low-energy spectral line (for instance 6.4 keV fluorescent K- α iron line) emitted by a narrow line region (which could be possibly part of the supernova ejecta) which is up-scattered to MeV energies by the relativistic jet. For the line not to be broadened more than we observe, the electrons that scatter the photons must be nonrelativistic in the jet comoving frame: this kind of Comptonisation has been proposed to operate in the jets of blazars, boosting the continuum and the broad-line photons (40). The most efficient configuration requires the seed photons to be produced or isotropized by Compton scattering near the relativistic jet Thomson photosphere. The boosted photon energy of the iron K- α line (which we focus on, but the outcome is similar for nickel or cobalt lines) is $E_{\text{line}} \approx \Gamma^2 E_{\text{Fe}}$, where Γ is the jet bulk Lorentz factor and $E_{\text{Fe}} \sim 6.4$ keV. Given the observed line energy is $E_{\text{line}} \sim 10 \text{ MeV}$, this would require a jet bulk Lorentz factor $\Gamma \sim 40$ for the (un-shocked) cold plasma. The observed luminosity of the boosted iron line would then be $L_{\text{line}} \sim \tau_{\text{T}} \Gamma^4 L_{\text{Fe}}$, where L_{Fe} is the luminosity (in fluorescent iron) of the narrow line region and $\tau_{\text{T}} \sim 0.5 \sigma_{\text{T}} L_{\text{jet}} / 8 \pi R \Gamma^3 m_{\text{p}} c^3 \approx 0.46 L_{\text{jet},52} R_{14}^{-1} \Gamma_{1.6}^{-3}$ (77) is the Thomson optical depth of the portion of the jet where most of the Comptonization occurs, and m_{p} is the proton rest mass. This would require a narrow line region luminosity $L_{\text{Fe}} \sim 8.5 \times 10^{43} L_{\text{jet},52}^{-1} R_{14} \Gamma_{1.6}^{-1} \text{ erg/s}$, which is equivalent to a recombination rate $\dot{N}_{\text{Fe}} = L_{\text{Fe}}/E_{\text{Fe}} \sim 8.7 \times 10^{51} L_{\text{jet},52}^{-1} R_{14} \Gamma_{1.6}^{-1} \text{ s}^{-1}$. We compare this to the expected iron recombination rate in a simple reflection model ((41), their equation 4), which indicates a required iron mass (assuming the narrow line region to be a spherical shell of thickness ΔR_{Fe} and temperature T) of $M_{\text{Fe}} \sim 2.8 \times 10^{-3} L_{\text{jet},52}^{-1} R_{14} \Gamma_{1.6}^{-1} T_7^{3/4} \Delta R_{\text{Fe},14} M_{\odot}$, where M_{\odot} is the mass of the Sun.

This scenario requires a large required radius for the narrow line region, $R \gtrsim \Delta R_{\rm Fe}$, which must be equal or larger than the jet photospheric radius (where $\tau_{\rm T} = 1$), $R_{\rm ph} = 4.6 \times 10^{13} L_{\rm jet,52} \Gamma_{1.6}^{-3}$ cm, and a low Lorentz factor of the jet $\Gamma \sim 40$ at the photosphere. These are difficult to reconcile because it requires the supernova (SN) ejecta to reach $R_{\rm ph}$ in a very short time (assuming the supernova exploded at the time marked by the GRB precursor), $t_{\rm SN} \sim t_{\rm line}/(1 + z) \sim 300(1 + z)^{-1}$ s, which would require relativistic expansion of the SN ejecta with a Lorentz factor $\Gamma_{\rm SN} \gtrsim \sqrt{R_{\rm ph}/ct_{\rm SN}} \sim 5(1 + z)^{1/2} L_{\rm jet,52}^{1/2} \Gamma_{1.6}^{-3/2} t_{\rm line,2.5}^{-1/2}$, and therefore an explosion energy $E_{\rm SN} \gtrsim \Gamma_{\rm SN} M_{\rm Fe} c^2 \approx 2.5 \times 10^{53} (1 + z)^{1/2} L_{\rm jet,52}^{-1/2} R_{14} \Gamma_{1.6}^{-5/2} T_7^{-3/4} t_{\rm line,2.5}^{-1/2}$ erg, inconsistent with even the most energetic supernovae. This problem could be alleviated if the line photons were emitted at lower radii and then isotropized by scattering by stellar wind material. Alternatively, the iron mass could have been ejected years before collapse in an ejection event during the final phases of the stellar evolution (78). The low jet bulk Lorentz factor $\Gamma \sim 40$ is also problematic: while the Lorentz factor of the jet could be variable, sometimes reaching low values, the average value is probably substantially larger, and the variability timescale is likely much faster than the observed duration of the spectral feature (the Lorentz factor variability would also substantially broaden the up-scattered line). This problem could be reduced if the seed photons were at lower energies (such as fluorescent lines of other elements).

A third possibility is high-latitude emission (HLE) from the shell that produced the most luminous pulse in the GRB light curve. Photons produced at latitudes (angle between the jet expansion direction and the line of sight) $\theta > 1/\Gamma$ reach the observer over a time longer than the dynamical timescale, producing a tail of progressively less Doppler-boosted emission that can be observed if the emission drops rapidly (43–45, 79). In this scenario, electron-positron pairs are produced within the dissipation region that produced the peak of the GRB 221009A luminosity, for the same arguments as in our preferred scenario, but the shell has a larger bulk Lorentz factor, $\Gamma \sim 1000$. When the dissipation ends, the luminosity drops (as observed around $t \sim 260$ s after the trigger), and the HLE tail could become visible. In the tail, photons emitted over a dynamical time and within a solid angle $d\Omega = 2\pi \sin \theta d\theta$ reach the observer over a time interval $dt_{obs} \sim (1 + z)R \sin \theta d\theta/c = (1 + z)R d\Omega/2\pi c$ (44, 45), where R is the emission (turn-off) radius. The energy in these photons, as measured by the observer, is related to the emitted energy (80) through $dE_{obs} = (\delta^3/\Gamma)(\partial E_{em}/\partial\Omega)d\Omega$, where $\delta = \Gamma^{-1}(1 - \beta \cos \theta)$ is the Doppler factor. This leads to the HLE luminosity evolution

$$L_{\rm HLE} = \frac{dE_{\rm obs}}{dt_{\rm obs}} = \frac{1}{2}(1+z)^2 \frac{R^2}{c^2 \beta^3 \Gamma^4 (t_{\rm obs} - t_{\rm obs, peak})^3} E_{\rm em},$$
 (S18)

where $E_{\rm em}$ is the isotropic-equivalent energy emitted during a dynamical time (assumed equal at all latitudes), and $t_{\rm obs,peak}$ is the emission turn-off time (which would correspond to the peak time). The advantages of this scenario are that (i) it can accommodate a large Lorentz factor, because the line observed photon energy is given by $E_{\rm line} = \delta m_{\rm e}c^2$ with $\delta \lesssim \Gamma$, and (ii) it predicts a decrease in both the line luminosity and energy, with $L_{\rm line} \propto (t_{\rm obs} - t_{\rm obs,peak})^{-3}$ and $E_{\rm line} \propto \delta \propto (t_{\rm obs} - t_{\rm obs,peak})^{-1}$. Taking $t_{\rm obs,peak}$ to be in the range 220 - 240 s, this gives an evolution of L_{line} and E_{line} that roughly matches the observed one (Figure S8).

The pairs in this scenario would be produced by photon-photon annihilation within the shell, similarly as in our preferred scenario. A difficulty is that the compactness parameter $\ell' \propto \Gamma^{-3}$ could be much lower (if the shell were to have a much larger Lorentz factor), which would prevent efficient pair creation, despite the very large observed luminosity.

Another potential difficulty is that the entire prompt emission spectrum would be affected by the same HLE mechanism, so the line might not be sufficiently bright to be detectable against the combined HLE broadband emission from the bright pulse, without the HLE broadband component being dominant over the on-axis emission (otherwise we would see the peak of the broadband component tracking the evolution of the line, which is not the case).

Although we cannot completely exclude these alternative scenarios, we prefer the model with the electron-positron annihilation line production in the slow-moving shell, as it does not present evident difficulties.



Figure S1: Light curve and temporal evolution of the spectral parameters for each of the 13 time intervals analyzed. Spectral parameters from the best-fit models are plotted as a function of time. The error bars represent the 16th to 84th percentiles of the samples of each parameter. In the 7th and the 8th time intervals, the Gaussian is not statistically significant, and the corresponding 2- σ upper limits on its luminosity are displayed. A: Light curve of GRB 221009A in the 8-900 keV, together with the 13 time intervals analyzed (dashed vertical black lines) as in Fig. 2A. In the other panels, the parameters of the Gaussian model are shown as red points, those related to the power law are shown in orange, while different colours have been used to represent the parameters of the continuum prompt models (SBPL or 2SBPL). B: The characteristic energies E_{break} (turquoise crosses), E_{peak} (blue crosses) and the central energy of the Gaussian E_{gauss} ; C: the photon indices α_1 (light blue crosses), α or α_2 (of the SBPL or 2SBPL function, respectively, yellow crosses), β (purple points) and Γ_{PL} . D: the luminosity of the typical prompt spectral function (either 2SBPL or SBPL, blue points), of the Gaussian and of the power-law functions. E: shows the evolution of the width σ of the Gaussian component.



Figure S2: Results of the spectral simulations to assess the significance of the line. These plots show the probability that the improvement in the best goodness of fit statistic, after adding a Gaussian component to a background plus SBPL model, exceeds a given value ΔPG stat if the true model is background plus SBPL. A: The red line shows the p-value associated to a given ΔPG stat estimated from 8×10^4 Monte Carlo simulations produced using the best fitting parameters of our background plus SBPL model of interval 5 (corresponding to the 280 - 300 s time bin), assuming the background counts are normally distributed and the source counts are Poisson distributed. The pink shaded area comprises p-values within one standard deviation of the Monte Carlo estimate. The blue line shows the extrapolation of the p-value as defined in Eq. S5. The vertical dashed line shows the ΔPG stat from the real data. B: Same results for interval 6 (300 - 320 s). C: Zoom in to part of the plot in Panel B.



Figure S3: **Conceptual diagram of our preferred interpretation**. *Panel A*: the unsteady relativistic outflow, or 'jet', is composed of separate shells with different bulk Lorentz factors and kinetic luminosities. The blue plot shows the bulk Lorentz factor of the jet as a function of the radial distance R from the central engine, while the red plot shows the corresponding kinetic luminosity. The horizontal thin grey line in the blue plot represents the average bulk Lorentz factor of the jet $\langle \Gamma \rangle$. Different shades of blue in the sketch in the lower part of the panel represent different densities. We focus on a pair of consecutive shells, enclosed in the pink dashed rectangle in the sketch, whose kinetic luminosities $L_{kin,4} \gg L_{kin,1}$ and $\Gamma_4 \gg \Gamma_1$, where the subscript 1 refers to the outer, slower shell and 4 to the inner, faster one. *Panel B*: as the two shells collide, two regions form, separated by a contact discontinuity (CD). Region 2 is separated by the forward shock (FS) from region 1 and contains the shocked material from that region; region 3 is separated by the reverse shock (RS) from region 4 and contains the shocked material from that region. The bulk Lorentz factor across these two regions is approximately uniform and equal to $\Gamma \sim (L_{kin,4}/L_{kin,1})^{1/4}\Gamma_1$. The line is produced within region 2 or 3, or both.



Figure S4: Count rate spectra of BGO1 detector from GRB 221000A in intervals 5 and 6. A and C: The background (black crosses) and the total source plus background (orange crosses) counts rate spectra are compared with the orbital background estimates (in red from our analysis and in blue calculated with osv1.3) obtained during the 30th orbits before and after the investigated time intervals (see text). No emission line appears in any of the background spectra. **B and D:** Background-subtracted count rate spectrum, obtained by subtracting the black from the orange spectra shown in panels A and C. Raw data are shown without binning.



Figure S5: Spectrum of GRB 221000A in interval 5 from both BGO detectors, in counts rate and νF_{ν} representations. A: Count rate spectra of BGO0 detector related to the total (source plus background, in blue), the background (in black), and the orbital background estimates (in red). B: Background-subtracted count rate spectrum from both BGO detectors, along with the SBPL (dotted line) and SBPL+Gaussian line (solid line) models. C: νF_{ν} representation of the same spectrum with the data points corresponding to different detectors (including NaI ones, see legend) of Fermi/GBM, along with the best-fitting model SBPL+Gaussian line (solid line), and the SBPL (dotted line) and Gaussian (dashed line) models. The emission line is present in both BGO detectors (BGO 0 and BGO 1) data.



Figure S6: ΔAIC as a function of simulated flux ratio. Results of the simulations of a spectrum with the same parameters as the ones observed in interval 6, but with a normalization reduced by factors from 2 to 100 (see text). The dotted black line on the top indicates the ΔAIC found from observations, while the dashed red line represents the threshold $\Delta AIC = 4$ that roughly corresponds to a one-sigma significance. The blue error bars represent the 16th to 84th percentile ranges of each resulting ΔAIC distribution derived from our simulations.



Figure S7: Comparison of the lightcurves of the 4 brightest GRBs ever detected by Fermi in 15 years of activity. A: The lightcurve detected in the energy range 8-900 keV by the most illuminated NaI detector. B: The lightcurves of the same GRBs as detected in the high-energy band 300 keV - 40 MeV by the most illuminated BGO detector. Lightcurves are not background-subtracted.



Figure S8: **High-Latitude Emission scenario.** Time evolution of the observed line luminosity L_{line} and central photon energy E_{line} (black data points with error bars), compared with the evolution predicted in the HLE scenario. The blue, purple and orange lines correspond to different reference time choices for the peak time $t_{\text{obs,peak}}$ of the prompt emission pulse where the line originated, as reported in the legend. The normalizations are arbitrary and serve to the only purpose of demonstrating that the L_{line} and E_{line} evolution is compatible with the HLE scenario.



Figure S9: Evolution of the GBM spectrum up to 360 s. The GBM spectra (data points) in the same time intervals shown in Fig. 1, overlaid with the best-fitting model (reported in the legend) represented by grey solid lines. Gray shading represents the 16th to 84th percentile model uncertainty interval. The SBPL component is represented by the dotted gray lines, Gaussian component is represented by the dashed gray lines (with downward arrows, in case of $2-\sigma$ upper limit in panel G and H), and the power law component is shown with dotted-dashed gray lines.



Figure S10: **Evolution of the spectrum in the sub-intervals from 280 s to 320 s.** Same plot as Fig. S9, but showing the spectral data in the 6 sub-intervals analyzed, from 5.1 (280 to 285 s, panel A) to 6.2 (310 to 320 s, panel F).



Figure S11: **GBM spectrum of GRB 170409A in the time intervals from 32 s to 38 s.** The GBM spectra (data points, with downward arrows for 3- σ upper limits) overlaid with the best-fitting model (2SBPL) represented by grey solid lines. Gray shading represents the 16th to 84th percentile model uncertainty interval. The 2- σ upper limit on the Gaussian component is represented by the dashed gray lines (with downward arrows).

Table S1: **Continuum spectral parameters.** Spectral parameters of the best-fitting models describing the continuum spectral shapes observed in GRB 221009A (either the SBPL, 2SBPL or PL models), for each time interval. Parameters of the Gaussian emission feature are listed in Table 1.

Time interval [s]	Interval number	Model	L_{iso} (10 keV - 40 MeV) $[10^{51} erg s^{-1}]$	$lpha_1$ or $\Gamma_{ m PL}$	E_{break} [keV]	α or α_2	$E_{ m peak}$ [keV]	β
0 to 9	1	SBPL	$0.145\substack{+0.018 \\ -0.015}$	-	-	$-1.677\substack{+0.014\\-0.011}$	1257^{+420}_{-300}	$-3.14^{+0.36}_{-0.25}$
184 to 196	2	2SBPL	$3.612\substack{+0.035\\-0.035}$	$-0.927\substack{+0.019\\-0.020}$	231^{+32}_{-31}	$-1.656\substack{+0.050\\-0.047}$	1189^{+41}_{-40}	$-3.031^{+0.068}_{-0.056}$
196 to 206	3	2SBPL	$1.110\substack{+0.024\\-0.022}$	$-1.046\substack{+0.020\\-0.023}$	$104.7\substack{+9.3 \\ -8.8}$	$-1.877\substack{+0.017\\-0.040}$	279^{+36}_{-36}	$-2.762^{+0.081}_{-0.077}$
206 to 216	4	2SBPL	$1.496\substack{+0.028\\-0.026}$	$-1.084\substack{+0.018\\-0.021}$	$114.5^{+9.1}_{-10}$	$-1.911\substack{+0.025\\-0.030}$	947^{+140}_{-120}	$-3.27^{+0.20}_{-0.16}$
280 to 300	5	SBPL	$3.887^{+0.034}_{-0.034}$	-	-	$-1.509\substack{+0.0027\\-0.0028}$	683^{+13}_{-12}	$-2.417^{+0.014}_{-0.015}$
280 to 285 285 to 290 290 to 295 295 to 300	5.1 5.2 5.3 5.4	SBPL SBPL SBPL SBPL	$\begin{array}{r} 6.936 \substack{+0.077 \\ -0.081} \\ 4.359 \substack{+0.060 \\ -0.061} \\ 2.492 \substack{+0.061 \\ -0.057} \\ 1.876 \substack{+0.054 \\ -0.054} \end{array}$	-	-	$\begin{array}{r} -1.4271\substack{+0.0034\\-0.0035}\\-1.4646\substack{+0.0045\\-0.0045}\\-1.5796\substack{+0.0063\\-0.0058}\\-1.630\substack{+0.014\\-0.014}\end{array}$	$717^{+16}_{-17} \\ 554^{+16}_{-14} \\ 618^{+33}_{-29} \\ 880^{+150}_{-110}$	$\begin{array}{r} -2.377\substack{+0.016\\-0.017}\\-2.38\substack{+0.02\\-0.02}\\-2.449\substack{+0.042\\-0.044}\\-2.166\substack{+0.036\\-0.042}\end{array}$
300 to 320	6	SBPL	$0.942\substack{+0.019\\-0.049}$	-	-	$-1.681\substack{+0.022\\-0.013}$	543^{+170}_{-210}	$-2.0608\substack{+0.0084\\-0.035}$
300 to 310 310 to 320	6.1 6.2	SBPL SBPL PL	${}^{1.201^{+0.029}_{-0.033}}_{0.36^{+0.100}_{-0.068}}_{0.415^{+0.087}_{-0.140}}$	$-1.852^{+0.030}_{-0.019}$	- -	$-1.6913^{+0.012}_{-0.0065}\\-1.536^{+0.077}_{-0.055}$	1036^{+400}_{-280} $64.4^{+19.0}_{-7.1}$	$\begin{array}{c} -2.078 \substack{+0.019 \\ -0.026} \\ -2.173 \substack{+0.064 \\ -0.063} \end{array}$
320 to 340	7	SBPL PL	$0.195^{+0.025}_{-0.018}\\0.540^{+0.041}_{-0.041}$	$-1.846^{+0.015}_{-0.021}$	-	$-1.683^{+0.036}_{-0.034}$	$66.4^{+3.0}_{-2.8}$	$-2.96^{+0.14}_{-0.15}$
340 to 360	8	SBPL PL	$\begin{array}{c} 0.165 \substack{+0.017 \\ -0.015 \\ 0.456 \substack{+0.025 \\ -0.026 \end{array}} \end{array}$	$-1.837^{+0.029}_{-0.032}$	-	$-1.745^{+0.027}_{-0.026}$	$64.8^{+2.8}_{-2.6}$	$-3.57^{+0.17}_{-0.2}$
360 to 380	9	SBPL	$0.3531\substack{+0.0069\\-0.0071}$	-	-	$-1.420\substack{+0.052\\-0.053}$	$43.9^{+1.5}_{-1.6}$	$-2.251^{+0.017}_{-0.020}$
380 to 400	10	SBPL	$0.905\substack{+0.020\\-0.019}$	-	-	$-1.6356\substack{+0.0072\\-0.0068}$	$436.0^{+19.0}_{-17.0}$	$-2.48^{+0.041}_{-0.04}$
400 to 420	11	SBPL	$0.979\substack{+0.020\\-0.018}$	-	-	$-1.5859^{+0.0090}_{-0.0079}$	$314.7^{+10.0}_{-9.3}$	$-2.430\substack{+0.032\\-0.032}$
420 to 440	12	SBPL	$0.826\substack{+0.019 \\ -0.016}$	-	-	$-1.610\substack{+0.012\\-0.011}$	$246.1^{+8.2}_{-7.7}$	$-2.376^{+0.032}_{-0.032}$
440 to 460	13	SBPL	$1.60\substack{+0.02\\-0.02}$	-	-	$-1.5506\substack{+0.0058\\-0.0056}$	$332.8^{+7.1}_{-7.1}$	$-2.41\substack{+0.02\\-0.02}$