A High Stellar Velocity Dispersion for a Compact Massive Galaxy at $z = 2.2$

Pieter G. van Dokkum[1], Mariska Kriek[2], and Marijn Franx[3]

[1] Astronomy Department, Yale University, 260 Whitney Ave, New Haven, CT 06511, USA

[2] Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA

[3] Leiden Observatory, Leiden University, NL-2300 RA Leiden, The Netherlands

---

Recent studies have found that the oldest and most luminous galaxies in the early Universe are surprisingly compact,[1–7] having stellar masses similar to present-day elliptical galaxies but much smaller sizes. This finding has attracted considerable attention[8–13] as it suggests that massive galaxies have grown by a factor of $\sim$ five in size over the past ten billion years. A key test of these results is a determination of the stellar kinematics of one of the compact galaxies: if the sizes of these objects are as extreme as has been claimed, their stars are expected to have much higher velocities than those in present-day galaxies of the same mass. Here we report a measurement of the stellar velocity dispersion of a massive compact galaxy at redshift $z = 2.186$, corresponding to a look-back time of 10.7 billion years. The velocity dispersion is very high at 510^{+165}_{-95} km s^{-1}, consistent with the mass and compactness of the galaxy inferred from photometric data and indicating significant recent structural and dynamical evolution of massive galaxies. The uncertainty in the dispersion was determined from simulations which include the effects of noise and template mismatch. However, we caution that some subtle systematic effect may influence the analysis given the low signal-to-noise ratio of our spectrum.

We observed the galaxy, dubbed 1255–0, with the Gemini Near-Infrared Spectrograph (GNIRS) on the Gemini South telescope for a total of 29 hrs. The de-redshifted spectrum is shown in Fig. 1a. A detailed description of the observations and reduction, as well as an analysis of the continuum emission and detected (weak) emission lines, is presented in a companion paper.[14] In the companion paper we derive a stellar mass of $\approx 2.0 \times 10^{11}$ M$_\odot$ for a Kroupa initial mass function, by fitting stellar population synthesis models to the broad band photometry and the GNIRS spectrum. The effective radius of 1255–0 $r_e = 0.78 \pm 0.17$ kpc, as previously measured[3] from deep Hubble Space Telescope (HST) NICMOS2 observations. The galaxy was selected from a well-studied[15,16,3] sample of nine spectroscopically-confirmed galaxies with evolved stellar populations at $z \sim 2.3$, and its properties are similar to those of other galaxies in this sample. The median stellar mass of the nine objects is 1.7×10^{11} M$_\odot$ and their median effective radius $r_e = 0.9$ kpc,[3] a factor of ~ 5 smaller than galaxies with similar masses at $z = 0$. The number density of these massive compact galaxies is substantial, about the same as that of galaxies in the nearby Universe that are a factor of 2–3 more massive.[10]

In the present study we use the deep Gemini spectrum to measure the stellar velocity dispersion of the galaxy, by employing standard techniques for measuring the broadening of the absorption lines.[17,18] Our methodology is explained in detail in the Supplementary Information. Briefly, smoothed model spectra were fitted to the data in real space, taking observational errors into account and ignoring data with the largest uncertainties. The uncertainty in the dispersion was determined from Monte Carlo simulations of many different combinations of assumed velocity dispersions and empirical realizations of the noise. Systematic uncertainties were assessed by varying the templates (also allowing for multiple components), the masking and weighting, and the continuum filtering, and are typically much smaller than the random uncertainty. We note that the spectrum is available in electronic form as a Supplementary Dataset.

We derive a velocity dispersion $\sigma = 510^{+165}_{-95}$ km s^{-1} for the galaxy, which is very high when compared to typical early-type galaxies in the nearby Universe. Although not statistically significant, it is striking that the best-fit value exceeds the measured dispersions of all individual galaxies in the Sloan Digital Sky Survey (SDSS).[19,20] In the SDSS a significant fraction of galaxies with velocity dispersions in excess of 350 km s^{-1} are superpositions, which are easily identified with HST imaging.[20] As shown in Fig. 1b-d, 1255–0 is a single, nearly round object with an effective radius of ≈ 0.1 arcsec in HST images. The dispersion is also a factor of ~ 2 higher than a previous measurement[21] from a stacked

spectrum of 13 galaxies at $\langle z \rangle = 1.6$. A direct comparison is difficult given the uncertainties associated with stacking individual spectra, but we note that the median stellar mass of the 13 galaxies is a factor of ~ 3 smaller than that of 1255–0 and the median effective radius is a factor of 1.5 larger. The expected dispersion of these $\langle z \rangle = 1.6$ galaxies is therefore a factor of ~ 2 lower than that of 1255–0, and the two results are consistent.

The high dispersion of 1255–0 confirms that the galaxy is very massive despite its diminutive size. The relation between mass, velocity dispersion, and size can be expressed as $M_{\text{dyn}} = C\sigma^2 r_e$, with C a constant that depends on the structure of the galaxy and other parameters. Using $\log C = 5.87$, which is the value that gives $M_{\text{dyn}} \approx M_{\text{star}}$ for galaxies in the SDSS,[6] we find $M_{\text{dyn}} = 1.5^{+1.2}_{-0.5} \times 10^{11} \, M_\odot$. For $\log C = 6.07$, the value derived from kinematic data of present-day early-type galaxies,[18] the dynamical mass is $2.4^{+1.9}_{-0.8} \times 10^{11} \, M_\odot$. Both estimates are in excellent agreement with the stellar mass (Fig. 2a). Put differently, the high dispersion that we measure was *expected* (and in fact predicted[3,10]) given our extreme size and stellar mass measurements. Quantitatively, the expected dispersion assuming $M_{\text{dyn}} = M_{\text{star}}$ and $5.87 \leq \log C \leq 6.07$ is in the range $470 \, \text{km s}^{-1} - 590 \, \text{km s}^{-1}$.

At the same time, the high dispersion confirms and extends the notion that quiescent galaxies at high redshift are structurally and dynamically very different from galaxies in the present-day Universe. Figures 2b-d show where 1255–0 falls with respect to the relations between velocity dispersion, size, and dynamical mass defined by SDSS galaxies. The galaxy is offset from the local relations, consistent with previous studies which were based on stellar masses derived from photometric data.[3,6] At fixed dynamical mass the dispersion is higher by a factor of ~ 2.5 and the effective radius is smaller by a factor of ~ 6. The most dramatic offset is in the $\log \sigma - \log r_e$ plane (Fig. 2b). These two quantities are measured directly and independently, and (to first order) do not depend on stellar populations.

The extreme compactness of massive high redshift galaxies is qualitatively consistent with models in which the central parts of massive galaxies form early in highly dissipative processes,[22] although it remains to be seen whether such models can produce objects with the size and velocity dispersion of 1255–0. In particular, it may be difficult to funnel gas clumps into an extremely compact configuration without forming stars at larger radii. Regardless of the details of the model, in its star-forming phase at $z \gtrsim 3.5$ the galaxy likely had a very compact molecular gas distribution with a rotation velocity of $\sim 700 \, \text{km s}^{-1}$. The median rotation velocity of CO in submm galaxies at $z = 2 - 3.5$ has been found[23] to

be $\approx 470\,\mathrm{km\,s^{-1}}$ (assuming $V_{\mathrm{rot}} = 0.6 \times \mathrm{FWHM}$) — a high value by most standards, but still somewhat lower than what we expect for the progenitors of galaxies such as 1255–0. There is not yet much information on the gas dynamics of massive galaxies at redshifts $z > 3.5$. The $z = 6.4$ quasar SDSS J1148+5251 has a relatively small CO linewidth of $V_{\mathrm{rot}} \approx 170\,\mathrm{km\,s^{-1}}$,[24] but it may be that quasars are biased low because their gas disks are preferentially seen face-on.[25] It is obviously not clear whether the gas was ever in a regular disk; it would be interesting to determine whether 1255–0 shows rotation, but that requires imaging (or spectroscopy) of higher spatial resolution than is currently available.

A problem that is perhaps even more vexing than the origin of galaxies such as 1255–0 is their subsequent evolution onto the local relations between size, velocity dispersion, and mass. The simplest explanation is that the mass and/or size measurements of the compact galaxies are incorrect,[3,13] but this is difficult to maintain given the dynamical measurement presented here. We are left with the conclusion that very significant structural and dynamical changes are required to bring massive, quiescent high redshift galaxies to the local relations. This cannot easily be achieved through star formation as the compact high redshift galaxies already appear to have stopped forming new stars, consistent with the old ages inferred for the stars in today's most massive galaxies. Among the models that have been proposed[8–13] minor mergers may be the most effective single mechanism, as simple virial arguments suggest that the velocity dispersion changes by a factor of $f_r^{-1/4}$ for a factor of f_r change in radius.[10,11] However, it is an open question whether mergers alone can "puff up" galaxies by the required amount, as the precise effect depends on the accretion rate, the masses, orbits, and gas content of accreted galaxies, angular momentum transfer between stars and dark matter, and on possible evolution in the contribution of dark matter to the measured kinematics.[26] Finally, we note that evolution in the velocity dispersion of galaxies would trivially imply evolution in the black hole mass – σ relation,[27,28] such that black hole masses are lower at fixed σ at high redshift.

While confirming that the velocity dispersions of compact galaxies are high, our measurement is obviously not sufficiently accurate to properly characterize the evolution of the relations in Fig. 2. A 1σ error of $25\,\%$ in the velocity dispersion implies an error of $56\,\%$ in the dynamical mass, and further progress requires dispersions with uncertainties $\lesssim 10\,\%$ for much larger samples. New spectrographs being readied for use on 8m class telescopes, combined with new wide field imaging surveys that can provide sufficiently bright targets, are expected to revolutionize this field in the next few years. As indicated here, such observations are crucial for calibrating stellar masses at high redshift and for

measuring the structural and dynamical evolution of massive galaxies from the time that their star formation was quenched to the present.

1. Trujillo, I., Förster Schreiber, N. M., Rudnick, G., Barden, M., Franx, M. *et al.* The Size Evolution of Galaxies since $z \sim 3$: Combining SDSS, GEMS, and FIRES. *Astrophys. J.* **650**, 18–41 (2006).

2. Toft, S., van Dokkum, P., Franx, M., Labbé, I., Förster Schreiber, N. M. *et al.* Hubble Space Telescope and Spitzer Imaging of Red and Blue Galaxies at $z \sim 2.5$: A Correlation between Size and Star Formation Activity from Compact Quiescent Galaxies to Extended Star-forming Galaxies. *Astrophys. J.* **671**, 285–302 (2007).

3. van Dokkum, P. G., Franx, M., Kriek, M., Holden, B., Illingworth, G. D. *et al.* Confirmation of the Remarkable Compactness of Massive Quiescent Galaxies at $z \sim 2.3$: Early-Type Galaxies Did not Form in a Simple Monolithic Collapse. *Astrophys. J.* **677**, L5-L8 (2008).

4. Cimatti, A., Cassata, P., Pozzetti, L., Kurk, J., Mignoli, M. *et al.* GMASS ultradeep spectroscopy of galaxies at $z \sim 2$. II. Superdense passive galaxies: how did they form and evolve? *Astron. Astrophys.* **482**, 21-35 (2008).

5. van der Wel, A., Holden, B. P., Zirm, A. W., Franx, M., Rettura, A. *et al.* Recent Structural Evolution of Early-Type Galaxies: Size Growth from $z = 1$ to $z = 0$. *Astrophys. J.* **688**, 48-58 (2008).

6. Franx, M., van Dokkum, P. G., Förster Schreiber, N. M., Wuyts, S., Labbé, I., & Toft, S. Structure and Star Formation in Galaxies out to $z = 3$: Evidence for Surface Density Dependent Evolution and Upsizing. *Astrophys. J.* **688**, 770-788 (2008).

7. Damjanov, I., McCarthy, P. J., Abraham, R. G., Glazebrook, K., Yan, H. *et al.* Red Nuggets at $z \sim 1.5$: Compact Passive Galaxies and the Formation of the Kormendy Relation. *Astrophys. J.*, in press (arXiv:0807.1744).

8. Naab, T., Johansson, P. H., Ostriker, J. P., & Efstathiou, G. Formation of Early-Type Galaxies from Cosmological Initial Conditions. *Astrophys. J.* **658**, 710-720 (2008).

9. Fan, L., Lapi, A., De Zotti, G., & Danese, L. The Dramatic Size Evolution of Elliptical Galaxies and the Quasar Feedback. *Astrophys. J.* **689**, L101-L104 (2008).

10. Bezanson, R., van Dokkum, P. G., Tal, T., Marchesini, D., Kriek, M. *et al.* The Relation Between Compact, Quiescent High Redshift Galaxies and Massive Nearby Elliptical Galaxies: Evidence for Hierarchical, Inside-Out Growth. *Astrophys. J.*, in press (arXiv:0903.2044).

11. Naab, T., Johansson, P. H., Ostriker, J. P. Minor Mergers and the Size Evolution of Elliptical Galaxies. *Astrophys. J.*, submitted (arXiv:0903.1636).

12. van der Wel, A., Bell, E. F., van den Bosch, F. C., Gallazzi, A., & Rix, H.-W. On the Size and Co-Moving Mass Density Evolution of Early-Type Galaxies. *Astrophys. J.*, submitted (arXiv:0903.4857).

13. Hopkins, P. F., Bundy, K., Murray, N., Quataert, E., Lauer, T., & Ma, C.-P. Compact High-Redshift Galaxies are the Core of Present-Day Massive Spheroids. *Astrophys. J.*, submitted (arXiv:0903.2479).

14. Kriek, M., van Dokkum, P. G., Labbé, I., Franx, M., Illingworth, G. D., Marchesini, D., & Quadri, R. F. An Ultra-Deep Near-Infrared Spectrum of a Compact Quiescent Galaxy at $z = 2.2$. *Astrophys. J.*, in press (arXiv:0905:1692).

15. Kriek, M., van Dokkum, P. G., Franx, M., Quadri, R., Gawiser, E. *et al.* Spectroscopic Identification of Massive Galaxies at $z \sim 2.3$ with Strongly Suppressed Star Formation. *Astrophys. J.* **649**, L71-L74 (2006).

16. Kriek, M., van Dokkum, P. G., Franx, M., Illingworth, G. D., Marchesini, D. *et al.* A Near-Infrared Spectroscopic Survey of K-selected Galaxies at $z \sim 2.3$: Redshifts and Implications for Broadband Photometric Studies. *Astrophys. J.* **677**, 219-237 (2008).

17. Franx, M., & Illingworth, G., & Heckman, T. Major and minor axis kinematics of 22 ellipticals. *Astrophys. J.*, **344**, 613-636 (1989).

18. van Dokkum, P. G., & Stanford, S. A. The Fundamental Plane at $z = 1.27$: First Calibration of the Mass Scale of Red Galaxies at Redshifts $z > 1$. *Astrophys. J.*, **585**, 78-89 (2003).

19. Bernardi, M., Sheth, R. K., Nichol, R. C, Miller, C. J., Schlegel, D. *et al.* A Search for the Most Massive Galaxies: Double Trouble? *Astron. J.*, **391**, 1191-1199 (2006).

20. Bernardi, M., Hyde, J. B., Fritz, A., Sheth, R. K., Gebhardt, K., & Nichol, R. C. A Search for the Most Massive Galaxies – II. Structure, Environment, and Formation. *Mon. Not. R. Astron. Soc.*, **391**, 1191-1199 (2008).

21. Cenarro, A., & Trujillo, I. Mild Velocity Dispersion Evolution of Spheroid-like Massive Galaxies since $z \sim 2$. *Astrophys. J.*, **696**, L43-46 (2009).

22. Dekel, A., Birnboim, Y., Engel, G., Freundlich, J., Goerdt, T., *et al.* Cold streams in early massive hot haloes as the main mode of galaxy formation. *Nature* **457**, 451-454 (2009).

23. Greve, T. R., Bertoldi, F., Smail, I., Neri, R., Chapman, S. C., *et al.* An interferometric CO survey of luminous submillimetre galaxies. *Mon. Not. R. Astron. Soc.* **359**, 1165-1183 (2005).

24. Bertoldi, F., Cox, P., Neri, R., Carilli, C. L., Walter, F., *et al.* High-excitation CO in a quasar host galaxy at $z = 6.42$. *Astron. Astrophys.* **409**, L47-L50 (2003).

25. Narayanan, D., Li, Y., Cox, T. J., Hernquist, L., Hopkins, P., *et al.* The Nature of CO Emission from $z \sim 6$ Quasars. *Astrophys. J. Supp.* **174**, 13-30 (2008).

26. Boylan-Kolchin, M., Ma, C.-P., Quataert, E. Red mergers and the assembly of massive elliptical galaxies: the fundamental plane and its projections. *Mon. Not. R. Astron. Soc.* **369**, 1081-1089 (2006).

27. Ferrarese, L., Merritt, D. A Fundamental Relation between Supermassive Black Holes and Their Host Galaxies. *Astrophys. J.* **539**, L9-L12 (2000).

28. Gebhardt, K., Bender, R., Bower, G., Dressler, A., Faber, S. M., *et al.* A Relationship between Nuclear Black Hole Mass and Galaxy Velocity Dispersion. *Astrophys. J.* **539**, L13-L16 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Author Contributions P.v.D. wrote the Gemini proposal, did the observations, measured the velocity dispersion, wrote the paper and led the interpretation. M.K. reduced the Gemini spectrum, determined the stellar mass, and contributed to the interpretation. M.F. independently measured the velocity dispersion and contributed to the analysis and interpretation.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. Correspondence and requests for materials should be addressed to P.v.D. (pieter.vandokkum@yale.edu).
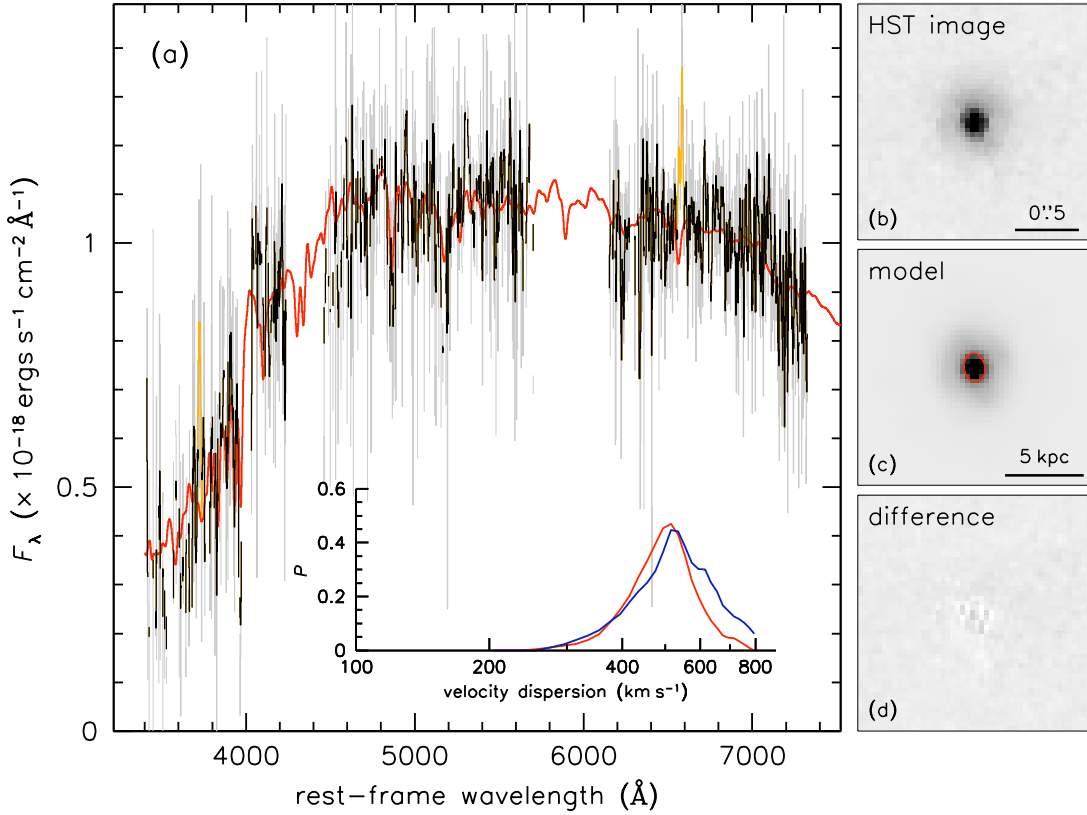
Figure 1 — Spectrum and HST images of 1255–0 at $z = 2.186$. **a**, Spectrum that was used to measure the velocity dispersion. Light grey shows the spectrum at a resolution of 5 Å ($\approx 100\,\mathrm{km\,s^{-1}}$), which was used for the actual measurement. A smoothed version of the same data (using a 25 Å boxcar filter) is shown in black. Regions around detected emission lines are shown in orange and were excluded from the fits. The most prominent absorption lines are Hβ at $\lambda 4861$ Å and Mg at $\lambda 5172$ Å. The best-fitting stellar population synthesis model,[14] smoothed to the best-fitting velocity dispersion, is shown in red. The inset shows the results of Monte Carlo simulations to determine the uncertainty in the best-fitting velocity dispersion. The curves show how often a dispersion of $510\,\mathrm{km\,s^{-1}}$ is measured given the true dispersion and noise. The two curves are for two different methods of simulating noise: shuffling the residuals of the fit in the wavelength direction (blue curve), and extracting "empty" 1D spectra from the 2D spectrum (red curve). **b-d**, The HST NICMOS2 image of the galaxy in the H_{160} filter, the best-fitting model of the galaxy (with the effective radius indicated in red), and the residual obtained by subtracting the model from the data. The galaxy is a single, very compact object with an effective radius of 0.78 kpc. Its coordinates are $\alpha = 12^{\mathrm{h}}54^{\mathrm{m}}59.6^{\mathrm{s}}$, $\delta = +01°11^{\mathrm{m}}30^{\mathrm{s}}$ (J2000), its K band observed magnitude is 19.26 (Vega) and its R band observed magnitude is 24.98 (Vega).[16] Alternative names that have been used for this object are 1256-151[15] and 1256-0.[3,16]
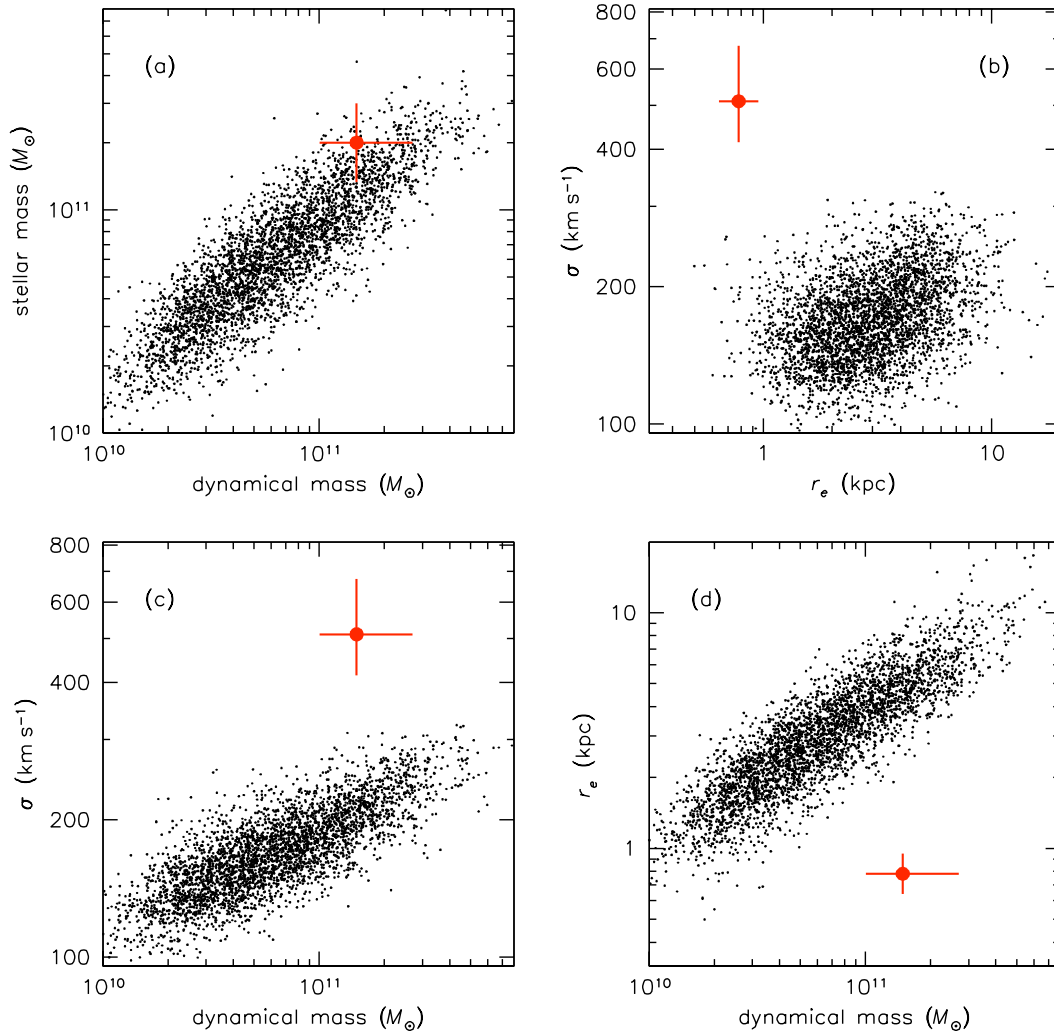
Figure 2 — **Properties of 1255–0 compared to nearby galaxies. a**, Relation between stellar mass and dynamical mass. Small symbols are galaxies in the SDSS[6] in the redshift range $0.05 - 0.07$, and the large red symbol is galaxy 1255–0 at $z = 2.186$. Our definition of dynamical mass, $\log M_{\mathrm{dyn}} = 5.87 + 2\log(\sigma) + \log(r_e)$, leads to a one-to-one correspondence between stellar mass and dynamical mass for SDSS galaxies. Despite its small size 1255–0 has a very high mass, similar to elliptical galaxies today. The dynamical mass is consistent (within 1σ) with the stellar mass that was estimated[14] from fitting stellar population synthesis models to the photometry. **b-d,** Relations between velocity dispersion, effective radius, and dynamical mass. Note that these three panels do not depend on stellar populations (except indirectly through the fact that the spectrum and the Hubble Space Telescope image are weighted by luminosity, not mass). It is clear that the structure and kinematics of 1255–0 are fundamentally different from those of nearby galaxies, and significant evolution is required to bring this object to the local relations.

Supplementary Information: Robustness of the velocity dispersion

The S/N ratio of the spectrum ranges from 5–8 per resolution element between the sky lines, which is about a factor of two lower than what is typically used for velocity dispersion measurements. We are helped by the unusually large rest-frame wavelength coverage of the spectrum, the large value of the dispersion, and the fact that in this case a measurement with a 1σ uncertainty of $\sim 20\,\%$ is already highly informative. However, a concern is that systematic effects can dominate when the S/N is low. Here we describe several tests we performed to test the robustness of the measurement.

Templates: The default template is the best-fit to the rest-frame UV – optical SED. This is a Solar metallicity model with an age of 2.1 Gyr, an e-folding time $\tau = 0.3\,$Gyr, and extinction $A_V = 0.25\,$mag. We also fitted single stellar populations (SSPs) with a range of ages and metallicities. We note that the linestrength is not fixed but a free parameter in the fit. The youngest models that were used have ages of $0.5\,$Gyr, despite the fact that such models provide a poor fit to the overall SED. Composite models were also tried, comprised of a $0.5\,$Gyr component in addition to a maximally-old population. Finally we fitted several stellar spectra that we used previously at lower redshifts, even though they span a smaller wavelength range than the galaxy spectrum. The velocity dispersion is fairly insensitive to the choice of template, varying mostly well within the quoted uncertainty, particularly when the absorption-line redshift is constrained to be within several hundred $\mathrm{km\,s^{-1}}$ of the emission-line redshift. This is illustrated in Supp. Fig. 1, which shows the reduced χ^2 as a function of velocity dispersion for the default template and for three examples of alternative templates.

Fitting region: The S/N ratio of the full spectrum is just sufficient for a dispersion measurement, and fitting over smaller wavelength regions gives results that are much less robust. Most of the signal is contributed by the observed H band: in the J band the S/N is lower, and in the K band the absorption lines are weaker. Fitting to the observed $J+H$ bands gives $\sigma = 480\,\mathrm{km\,s^{-1}}$ and fitting to the $H + K$ bands gives $\sigma = 590\,\mathrm{km\,s^{-1}}$, both with large uncertainties. We illustrate which wavelength regions contribute most to our measurement in Supp. Fig. 2. The only rest-frame region that "prefers" a significantly smaller dispersion is just blueward of the $4000\,\text{Å}$ break. This region has relatively low S/N in our spectrum and it is notoriously sensitive to template mismatch.

Fitting method: The spectrum was fit in real space, weighting by the S/N ratio and

disregarding data with the highest noise (mostly the wavelength regions between the atmospheric windows). Emission lines were excluded; exluding the Hβ absorption line has no effect on the measured dispersion. Template spectra were convolved to the instrumental resolution of $\sigma_{\mathrm{instr}} \approx 140\,\mathrm{km\,s^{-1}}$ (taking their intrinsic resolution into account) and subsequently broadened to a finely sampled grid of velocity dispersions. Continuum filtering was done with a 5$^{\mathrm{th}}$ order polynominal; we verified that the results are insensitive to the order of the fit. The redshift and the normalization of the residual of the continuum fit (i.e., the linestrength) are free parameters in the fit. Two of us (P.v.D. and M.F.) independently wrote software to fit the spectrum, and our results are fully consistent.

Determination of uncertainty: The quoted uncertainty was determined from Monte Carlo simulations. We created a finely sampled grid of "true" dispersions. For each dispersion σ_{true} a model spectrum was created. Next, 500 realizations of the model spectrum were created by adding realistic noise. These 500 noisy spectra were fit with the same procedure as used for the actual data. The template was allowed to vary in the simulations. For $\sigma_{\mathrm{true}} < 510\,\mathrm{km\,s^{-1}}$ the probability $P(\sigma_{\mathrm{true}})$ is then the fraction of simulations that give a best-fit dispersion in excess of $510\,\mathrm{km\,s^{-1}}$, and for $\sigma_{\mathrm{true}} > 510\,\mathrm{km\,s^{-1}}$ it is the fraction of simulations that give measured dispersions below $510\,\mathrm{km\,s^{-1}}$.

The curves in the inset of Fig. 1a show the results of the simulations for two different ways to create realistic noise spectra. The simulations shown by the blue curve randomly redistributed the residuals from the best fit in the wavelength direction. This approach has the advantages that the noise characteristics of the data are exactly maintained and that the contribution of template mismatch to the residuals is included. We shuffled blocks of 5 pixels (25 Å) rather than individual pixels to retain the effects of correlated noise. The simulations that gave rise to the red curve were created by combining random rows of the two-dimensional spectrum. This approach has the advantage that the wavelength dependence of the noise is maintained. As shown in Fig. 1a both approaches give very consistent results. The quoted uncertainty was calculated from the blue curve as it is slightly wider than the red curve. Both simulations give a slightly larger uncertainty than the formal error from the χ^2 distribution (shown in Supp. Fig. 2). Finally, we note that the true uncertainty may be (even) larger than the quoted uncertainty due to some unknown systematic effect; as is well known, systematic effects become increasingly important at low S/N ratios. We provide the spectrum in electronic form as a Supplementary Dataset.
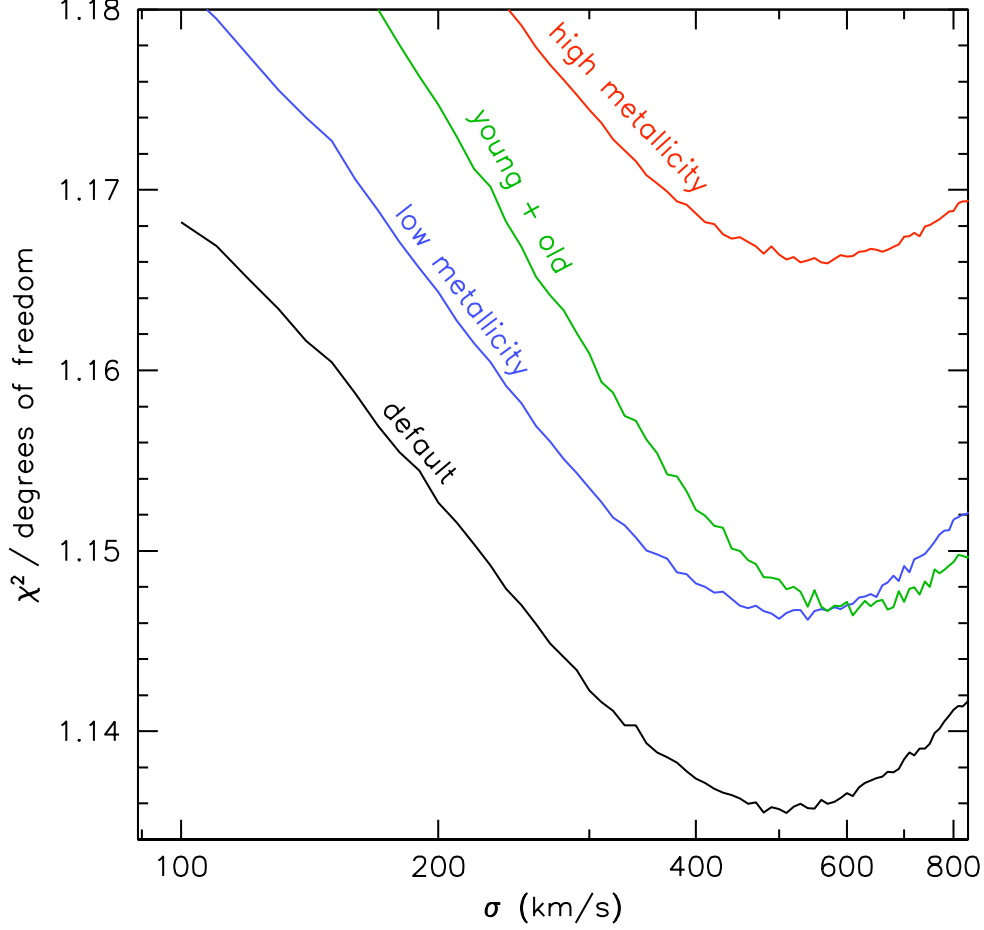
Supplementary Figure 1 — **Reduced χ^2 as a function of velocity dispersion.**
The black curve shows results for the default template, which is the best-fit stellar population
synthesis model to the spectrum and broad-band photometry. There is a well-defined minimum
at $\approx 500\,\mathrm{km\,s}^{-1}$. The blue, red, and green curves are examples of alternative templates, and
illustrate the effects of changing the template. The blue curve is for a single-age stellar population
with age $2\,\mathrm{Gyr}$ and metallicity $0.4\times$ Solar; the red curve is for a $2\,\mathrm{Gyr}$ population with metallicity
$2.5\times$ Solar; and the green curve is for a composite model comprised of a $0.5\,\mathrm{Gyr}$ old component in
addition to a maximally-old compoment. These alternative templates have higher minimum χ^2
values than the default template, as expected. Importantly, the minima occur at approximately
the same dispersion as for the default template.
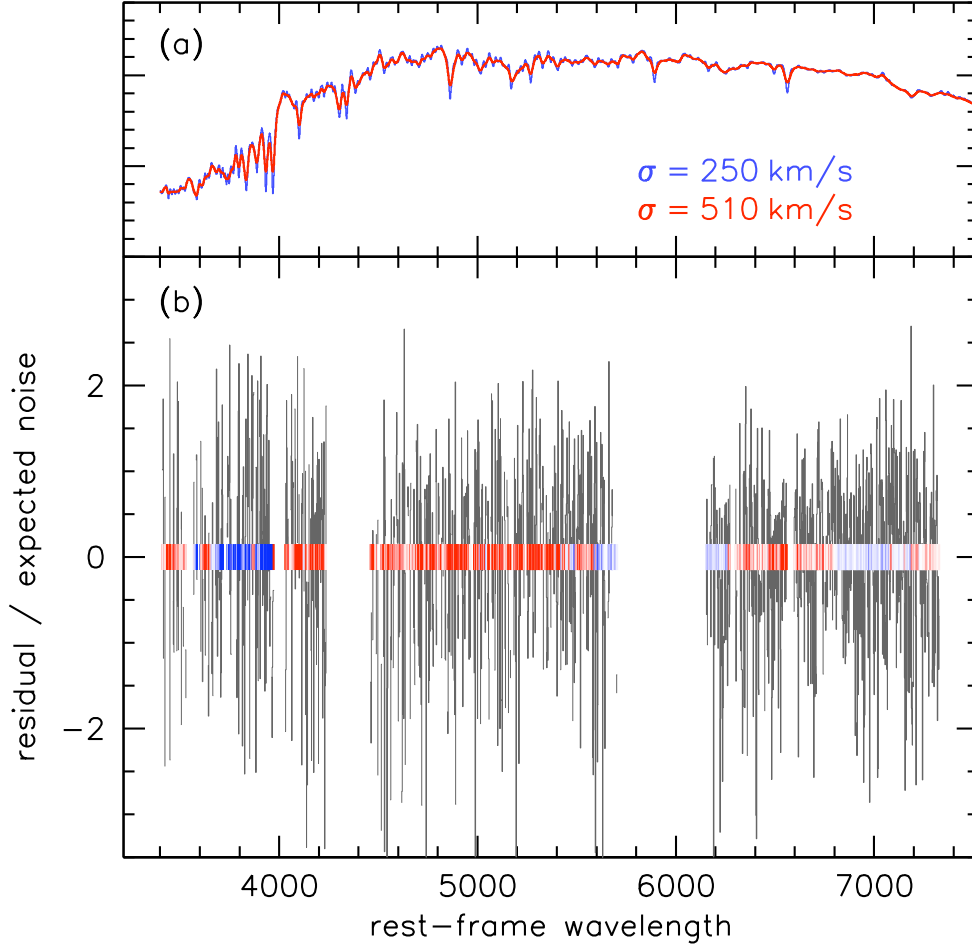
Supplementary Figure 2 — **Analysis of residuals from the fit. a**, Models with two different velocity dispersions. The red model has the best-fit velocity dispersion of $510\,\mathrm{km\,s^{-1}}$ and the blue model has a dispersion that is a factor of two lower. The differences between the models are subtle for any individual feature but are significant when the entire spectrum is considered. **b**, Residuals from the best-fit spectrum, normalized by the expected noise, are shown in grey. The residuals are well-behaved, and the reduced χ^2 for the best-fit is close to 1. The colored band shows which of the two models provides the best fit as a function of wavelength. Each vertical bar is a median of the nearest 50 datapoints (corresponding to $\approx 100\,\text{Å}$ in the rest-frame). The hue (red or blue) indicates which model fits best and the intensity indicates the absolute size of the difference between the residual from the low dispersion fit and the residual from the high dispersion fit, with white implying that both fits are equally good. With the exception of the region just blueward of the $4000\,\text{Å}$ break the high-dispersion model provides a better fit than the low-dispersion model. This demonstrates that our results are not driven by a small wavelength region or by other (obvious) wavelength-dependent effects.