Peer Review File

Editorial Note: Parts of this Peer Review File have been redacted as indicated to maintain the confidentiality

of unpublished data.

Manuscript Title: Bipolar Outflows out to 10 kpc for Massive Galaxies at Redshift z~1

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

I read the manuscript titled “Observational Evidence of the Prevalence of Bipolar Galactic Outflows out to 10 kpc at z ≈ 1 for Massive Galaxies” with interest. The paper presents an analysis of stacking pseudo-narrow band images from VLT MUSE observations of galaxies. The galaxies are stacked after aligning their major axis and as a function of inclination angles. The authors report that spatially extended MgII emission is seen in the stacked images (out to 10 kpc).

The results as presented are not fully convincing and there are several major issues with the analysis. The results are not accounting for several factors while performing the analysis as noted below. Further no rigorous statistical approach is followed while deriving the main results. No quantitative measurement of the extended MgII emission (accounting for all the issues noted below) is provided. No error-bar or uncertainty is estimated for the results presented. The manuscript is not appropriately crediting previous work done in this field (see comments below) and would benefit from a better discussion accounting for all the previous work. Kinematic studies of down-the-barrel spectroscopy of galaxies have shown evidence that disk galaxies predominantly drive bi-polar outflows. These results have been shown at different redshifts (z~0.1-1). The relevant literature references are presented in the discussion below.

As such, I don’t recommend that this paper be published in nature.

Issues:

- The stacking results present the emission map of MgII+stellar continuum together. MgII emission is weak and will have a significant stellar continuum contribution to it. Therefore, any result presented should always have the continuum contribution subtracted from it. For each individual pseudo-NB image, a corresponding continuum image must be created (from the MUSE datacube) and subtracted before stacking all the images. All the analysis must be performed on this continuum subtracted MgII emission map.

- The paper doesn’t perform any quantitative assessment on the statistical significance of the emission lines. What is the sky background on the continuum subtracted stacked images? Contours can be calculated to show >3/5 sigma emission isophotes, which can be used to quantify the spatial extent of the emission.

- The images are affected by ground based seeing. Additional smoothing is performed on the images, thereby increasing the effective PSF. Any analysis of the spatial extent must account for these two effects. All the reported distances are back of the envelope estimates taken by eye from the images.

- Change in MgII emission as a function of wavelength: MgII emission owing to galactic wind is typically arising from back-scattered light, thereby they are typically redshifted. MgII 2796 emission line (which is used here), will be blended with blueshifted MgII 2803 absorption, particularly for strong galactic winds. Therefore, it is obvious that any MgII2796 emission will be very hard to detect for face-on galaxies owing to this “emission-filling” effect. One is better off trying to detect the weaker MgII 2803 or fluorescent FeII* emission lines because of this problem. Therefore, the manuscript’s claim that only edge on galaxies show strong MgII emission simply means that there is less emission filling for edge on galaxies relative to face on galaxies (particularly at low latitudes, where most of their signal originates).

• Mass dependence: Again, this analysis must be performed on continuum subtracted images and quantified wrt to background levels. Just looking at the images presented, it is not obvious that emission is different around edge-on and face-on galaxies. Even for the high-mass sample, there is enough extended MgII emission around the face-on galaxy at high radii (with different morphology but still it is there). It is hard to say what is real vs an artefact in these images owing to large amount of smoothing performed. I recommend that the authors avoid smoothing the data for the analysis.

• Stacks are not unbiased: If one compares Figure 1 (Full sample) and Figure 3 (mass selected sample), Figure 1 resembles the results from the high-mass sample, although more than half the sample is at log M*<9.5. This clearly shows that their stacks are biased against the low mass galaxies. This is not surprising as the high-mass galaxies are more luminous and will dominate the signal in the stacks. A better light weighted (or inverse variance weighted) stacking method might be more unbiased here. The authors should also consider performing bootstrap resampling to account for

• Stacks should be performed on continuum subtracted NB images and bootstrap resampling should be performed to assess the impact of outliers and to ensure, non-linear error propagation owing to continuum-subtraction is properly accounted for.

• The Ring structure around face-on galaxies: The ring structure (particularly around the high-mass galaxy sample) might just mean that extended MgII emission is seen around face-on galaxies as well. It is not clear how the manuscript can rule out that hypothesis. Another possibility is the following: while inspecting the images, it is very clear that there are many other galaxies (background or foreground contaminations) and extended structures (e.g., tidal tails, extended disks) around these galaxies. A biased stack of face-on galaxies would be expected to produce a ring like structure which is light contamination from all these other sources. The authors should check that is not the case. They should mask out any bright sources around the galaxy before performing the image co-adds. This will also be true for the edge-on galaxy stacks.

• Several of the azimuthal angle estimates in Extended Data Figure 3 are significantly wrong. These are a minority of galaxies so won’t have a dramatic impact on the result but should still be corrected for completeness.

• CGM- outflow connection references: In general, the references quoted in the manuscript are woefully inadequate and misses out a significant portion of the relevant literature. In particular, the discovery paper that showed that MgII absorption along minor axis of galaxies is enhanced (Bordoloi et al. 2011, ApJ, 743, 1) and several follow up papers that confirm this evidence (Kacprzak et al., 2012, ApJL 760, 7; Lan et al. ApJ 2018, 866, 1).

• Direct evidence of biconical outflows: Even at higher-z, it has been known for more than a decade that regular galaxies exhibit bi-conical outflows. This was shown by studying down-the-barrel outflow kinematics as a function of galaxy inclination angles. First at low-z using SDSS galaxies (Chen Y.M. et al., 2010, ApJ, 140, 2), at intermediate z (~0.5) by Rubin et al. 2014, ApJ, 794, 2, and at z~1 by Kornei et al., 2012, ApJ 758, 2 and Bordoloi et al. 2014, ApJ 794, 2. The results presented in the manuscript should discuss their findings in this context.

• “A correlation between galaxy MgII EW and stellar mass was first reported by…” This is a surprising statement, as this correlation has been well known for more than a decade. Several groups have independently reported it over the years (Weiner et al. 2009, ApJ 692, 1; Rubin et al 2010, ApJ, 719, 2; Kornei et al, 2012, Bordoloi et al 2014, Rubin et al. 2014). Please reword the discussion in this correct context.

- Detection of galactic winds in emission: The authors should note that spectacular detection of extended galactic wind in emission was recently reported by Rupke et al. 2019, Nature, 574, 7780 (OII) and in MgII by Shaban et al 2022, ApJ, 936, 1. There is significant evidence of detection of spatially extended outflows in emission and should be included in the discussion.

Referee #2 (Remarks to the Author):

The paper presents an interesting study of MgII line in stacked data cubes of galaxies at z~1. The authors' main objective is to examine the geometry of the MgII emission by creating narrowband images. They find that the MgII emission is extended up to 10 kpc from the primary galaxy.

One of the unique aspects of this study is that the authors stack the datacubes based on the morphology of the galaxies, specifically dividing them into edge-on (PA>55) and face-on systems (PA<30). By doing so, the authors can detect complex profiles of the MgII line in both types of galaxies. The authors claim to detect extended emission in only edge-on systems, which they attribute to the geometry of the outflows, i.e., the biconical nature of the outflows. They also find a faint ring-like emission around face-on galaxies, source of which is unclear.

The authors estimate an opening angle of 64^0 for the outflows. This study sheds light on the geometry of the outflows and their impact on the surrounding intergalactic medium at z~1. The paper is nicely written and presents key results in clear and concise manner. I recommend the manuscript for publication after addressing a few of my concerns.

Major Comments:

1. The study reports that the total integration time in the edge-on systems is almost double compared to the face-on systems. However, it is unclear whether the lack of low signal-to-noise (S/N) extended emission in the face-on galaxies is due to a smaller sample size or other factors. The authors did not mention if they reached similar surface brightness in both stacked images.

To address this question, it may be useful to randomly select the same number of galaxies from both subsamples and check whether the detection of extended emission in edge-on systems still holds. This could help rule out the possibility that the lack of extended emission in face-on galaxies is solely due to a smaller sample size.

Additionally, it would be helpful if the authors included the number of galaxies stacked in each subsample to provide a clearer picture of the statistical significance of their findings. Knowing the sample sizes can also help readers understand the sensitivity of the analysis and the degree of confidence in the results.

2. The extended data Figure 1 reveals that both face-on and edge-on systems have similar distributions of stellar mass and redshift. However, it is important to note that due to size-mass evolution, it may be challenging to accurately estimate the position angle (PA) of low stellar mass galaxies at high redshift, which could lead to them being misclassified as face-on. To further clarify this point, it would be helpful for the authors to present a scatter plot of stellar mass vs redshift for both subsamples. This would provide a clearer understanding of the distribution of galaxies in the two subsamples and highlight any potential biases that may exist due to size-mass evolution.

3. The authors demonstrate that anisotropic outflows are only detected in the high-mass sample

of edge-on galaxies, which is an interesting finding. The MgII emission is significantly brighter in the low stellar mass systems, as has been demonstrated in earlier works (e.g., Erb et al. in 2012). The MgII emission appears to be more extended than the stellar continuum, indicating that scattering of photons by the outflowing gas. Could the authors comment whether the irregular geometry of the MgII emission could be due to the irregular geometry of outflows in the low stellar mass systems as opposed to the biconical geometry in higher mass galaxies?

4. The continuum subtracted spectra from the central 1'' in the edge-on system also shows the clear multiple kinematics components. In fact, the multiple components seem to be more significant in the edge-on system than in the face-on galaxies. If the outflows were purely biconical then I would not expect to detect any outflowing gas component in the central spectrum of edge-on galaxies.

Also, could the authors use some statistical metric such as BIC-score to justify fitting of multiple components for both face-on and edge-on galaxies.

Minor Comments:

The second sentence in the abstract "These outflows tend to form bipolar shapes above and below the galactic planes and extend well into the circumgalactic medium (CGM), up to tens of kpc perpendicular to the galaxy." makes it seem like the biconical nature of outflows is well observed and accepted phenomena. It undermines the significance of their findings. Maybe reword to something like "Theoretical models assumes bipolar shapes for the outflows that extends well into the circumgalctic medium (CGM), up to tens of Kpc perpendicular to the galaxies. However, they have been…"

Author Rebuttals to Initial Comments:

Response to referee #1's comments:

I read the manuscript titled "Observational Evidence of the Prevalence of Bipolar Galactic Outflows out to 10 kpc at z ≈ 1 for Massive Galaxies" with interest. The paper presents an analysis of stacking pseudo-narrow band images from VLT MUSE observations of galaxies. The galaxies are stacked after aligning their major axis and as a function of inclination angles. The authors report that spatially extended MgII emission is seen in the stacked images (out to 10 kpc).

The results as presented are not fully convincing and there are several major issues with the analysis. The results are not accounting for several factors while performing the analysis as noted below. Further no rigorous statistical approach is followed while deriving the main results. No quantitative measurement of the extended MgII emission (accounting for all the issues noted below) is provided. No error-bar or uncertainty is estimated for the results presented. The manuscript is not appropriately crediting previous work done in this field (see comments below) and would benefit from a better discussion accounting for all the previous work. Kinematic studies of down-the-barrel spectroscopy of galaxies have shown evidence that disk galaxies predominantly drive bi-polar outflows. These results have been shown at different redshifts (z~0.1-1). The relevant literature references are presented in the discussion below.

As such, I don't recommend that this paper be published in nature.

Issues:

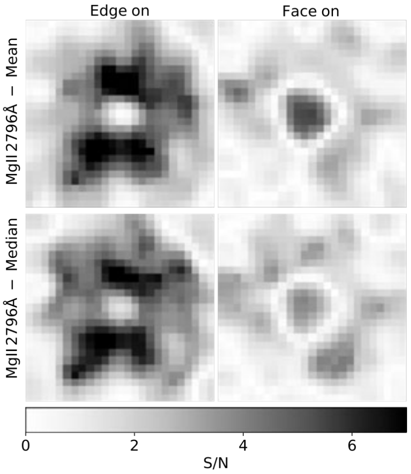
1• *The stacking results present the emission map of MgII+stellar continuum together. MgII emission is weak and will have a significant stellar continuum contribution to it. Therefore, any result presented should always have the continuum contribution subtracted from it. For each individual pseudo-NB image, a corresponding continuum image must be created (from the MUSE datacube) and subtracted before stacking all the images. All the analysis must be performed on this continuum subtracted MgII emission map.*

We appreciate the emphasis on the continuum subtraction in our study. However, we are happy to report that the continuum subtraction process was already included in the Methods (e.g., line260-line263). We confirm that before stacking, we already removed the continuum, and as a result, there is no substantial contribution from the stellar continuum in the stacking results. To make this clearer, we add a sentence in the Main Text (line64-line65).

Additionally, we would like to clarify that in our study, we stack the MUSE mini-datacubes instead of individual pseudo-NB images. It was already mentioned in the text in line62 and line259-line260.

2• *The paper doesn't perform any quantitative assessment on the statistical significance of the emission lines. What is the sky background on the continuum subtracted stacked images? Contours can be calculated to show >3/5 sigma emission isophotes, which can be used to quantify the spatial extent of the emission.*

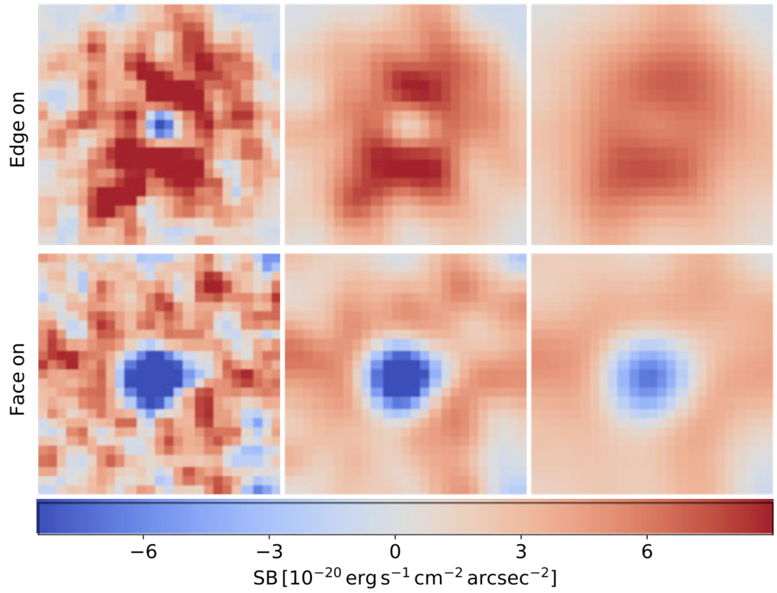
We thank the referee for emphasizing the S/N of this work. To address this concern, we have included an additional S/N image in the Methods section. The estimation of the S/N is explained in line274-line281. The noise is estimated by randomly producing pseudo-NB images near the MgII line within a wavelength range of 30A. All these random pseudo-NB images are produced using the same method as the MgII image. In this way, we take into account all the nonlinear effects in the data reduction and stacking. As we can see, we reach a S/N as high as ~5 in the bipolar region.



3• *The images are affected by ground based seeing. Additional smoothing is performed on the images, thereby increasing the effective PSF. Any analysis of the spatial extent must account for these two effects. All the reported distances are back of the envelope estimates taken by eye from the images.*

We agree that this work is limited by the spatial resolution. Besides the MUSE resolution, we also applied Gaussian smoothing. This is a common method to improve the data visualization (seen in many other works on CGM emission, e.g., Borisova et al. 2016, ApJ, 831, 39; Arrigoni Battaia et al. 2019, MNRAS, 482, 3162; Wisotzki et al. 2018, Nature, 562, 229; Leclercq et al. 2022, A&A, 663, A11). The aim of smoothing is to suppress the high-frequency noise and better present the signal at large scale. Smoothing would not under any circumstance produce the observed bipolar-shape emission. If anything, it would reduce anisotropies. We attach a plot showing the MgII pseudo-NB images with different smoothing kernel sizes.

The only distance presented in this paper is the impact parameter of the MgII outflow, which is calculated from the weighted average of flux (line362-line364). The error bar is computed using bootstrapping, based on the noise image calculated by the adjacent wavelength range, with a sampling of 100.



4• *Change in MgII emission as a function of wavelength: MgII emission owing to galactic wind is typically arising from back-scattered light, thereby they are typically redshifted. MgII 2796 emission line (which is used here), will be blended with blueshifted MgII 2803 absorption, particularly for strong galactic winds. Therefore, it is obvious that any MgII2796 emission will be very hard to detect for face-on galaxies owing to this "emission-filling" effect. One is better off trying to detect the weaker MgII 2803 or fluorescent FeII* emission lines because of this problem. Therefore, the manuscript's claim that only edge on galaxies show strong MgII emission simply means that there is less emission filling for edge on galaxies relative to face on galaxies (particularly at low latitudes, where most of their signal originates).*

We thank the referee for explaining again the emission-filling effect of the MgII line. This is exactly the physics behind this paper. In the central region of the face-on galaxies we indeed do not observe MgII in emission, but in absorption. For edge-on galaxies, in the center region (where there is stellar continuum in behind), we observe a P Cygni profile that contains the absorption and back-scattering of MgII photons. However, the major result of this paper, the bipolar MgII at a distance of 10 kpc perpendicular to the galaxy plane, is produced by MgII photons that escaped in the outflows. At this distance, there is no stellar continuum behind the outflow, and thus there would not be absorption of photons.

5• *Mass dependence: Again, this analysis must be performed on continuum subtracted images and quantified wrt to background levels. Just looking at the images presented, it is not obvious that emission is different around edge-on and face-on galaxies. Even for the high-mass sample, there is enough extended MgII emission around the face-on galaxy at high radii (with different morphology but still it is there). It is hard to say what is real vs an artefact in these images owing to large amount of smoothing performed. I recommend that the authors avoid smoothing the data for the analysis.*

We would like to reiterate that all our analyses are performed on continuum subtracted datacubes (Methods, line260-line263). The background level is similar for low-mass and high-mass subsamples, because the subsamples are divided by exactly the median stellar mass. Thus the number of objects is the same for the low- and high-mass subsamples. As we explained in #3, smoothing will not produce the anisotropic emission. We never denied that "there is enough extended MgII emission around the face-on galaxy at high radii". In this paper, we have confirmed that the extended MgII emission around the face-on galaxy at large radii is indeed real (Methods, line316-line319).

6• *Stacks are not unbiased: If one compares Figure 1 (Full sample) and Figure 3 (mass selected sample), Figure 1 resembles the results from the high-mass sample, although more than half the sample is at log M*<9.5. This clearly shows that their stacks are biased against the low mass galaxies. This is not surprising as the high-mass galaxies are more luminous and will dominate the signal in the stacks. A better light weighted (or inverse variance weighted) stacking method might be more unbiased here. The authors should also consider performing bootstrap resampling to account for*

We agree with the referee that a stack spanning a wide range of masses may be dominated by the higher mass objects. This is precisely why we divide the sample into high- and low-mass subsamples. The low-mass and high-mass subsamples are defined by exactly the median stellar mass.

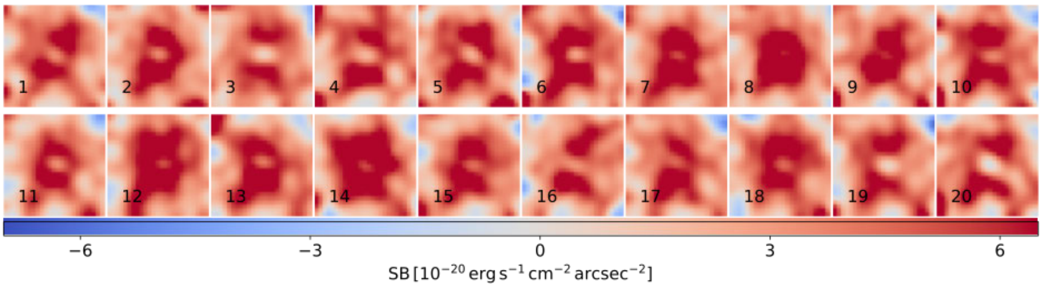
The comment on bootstrap and unbiased stacking is not very clear. We will provide a more detailed response in #7.

7• *Stacks should be performed on continuum subtracted NB images and bootstrap resampling should be performed to assess the impact of outliers and to ensure, non-linear error propagation owing to continuum-subtraction is properly accounted for.*

All our analyses are performed on continuum subtracted datacubes (Methods, line260-line263).

To make sure that the anisotropic Mg II emission is prevalent, and to confirm that the stacking is not dominated by a few outliers, we had already done a test by showing the distribution of MgII fluxes at the bipolar region for each individual galaxy (Methods, line289-line299). Now following the referee's suggestion, we have also performed a bootstrap analysis by randomly stacking 60 galaxies from the edge-on sample. Several examples are shown below. This further demonstrates the prevalence of the anisotropic MgII emission.

Furthermore, we thank the referee for emphasizing the S/N of this work. A S/N image is provided in #1, which takes into account all the non-linear effects.



8• The Ring structure around face-on galaxies: The ring structure (particularly around the high-mass galaxy sample) might just mean that extended MgII emission is seen around face-on galaxies as well. It is not clear how the manuscript can rule out that hypothesis. Another possibility is the following: while inspecting the images, it is very clear that there are many other galaxies (background or foreground contaminations) and extended structures (e.g., tidal tails, extended disks) around these galaxies. A biased stack of face-on galaxies would be expected to produce a ring like structure which is light contamination from all these other sources. The authors should check that is not the case. They should mask out any bright sources around the galaxy before performing the image co-adds. This will also be true for the edge-on galaxy stacks.

We already masked out all the neighboring galaxies before data stacking (Methods, line261-line262). Most of the undetected objects (randomly distributed faint neighbors) should be removed by median stacking of the datacube. Therefore, the background or foreground contaminations should not be an issue.

For the ring structure around face-on galaxies, we did not rule out the possibility that "*extended MgII emission is seen around face-on galaxies as well*". On the contrary, in the text (line81-line84), we list several possibilities. Our primary result is the bipolar morphology of the MgII emission around the edge-on galaxies, which is absent around the face-on galaxies.

9• *Several of the azimuthal angle estimates in Extended Data Figure 3 are significantly wrong. These are a minority of galaxies so won't have a dramatic impact on the result but should still be corrected for completeness.*

We thank the referee for bringing this issue to our attention. Several of the snapshots were difficult to assess due to the contrast of the image. We corrected this problem. We would be grateful if the referee could point out if there are others. The small number of galaxies will indeed not significantly impact the median stacking analysis.

10• *CGM- outflow connection references: In general, the references quoted in the manuscript are woefully inadequate and misses out a significant portion of the relevant literature. In particular, the discovery paper that showed that MgII absorption along minor axis of galaxies is enhanced (Bordoloi et al. 2011, ApJ, 743, 1) and several follow up papers that confirm this evidence (Kacprzak et al., 2012, ApJL 760, 7; Lan et al. ApJ 2018, 866, 1).*

We thank the referee for providing these references. We will respond to this comment in #13.

11• *Direct evidence of biconical outflows: Even at higher-z, it has been known for more than a decade that regular galaxies exhibit bi-conical outflows. This was shown by studying down-the-barrel outflow kinematics as a function of galaxy inclination angles. First at low-z using SDSS galaxies (Chen Y.M. et al., 2010, ApJ, 140, 2), at intermediate z (~0.5) by Rubin et al. 2014, ApJ, 794, 2, and at z~1 by Kornei et al., 2012, ApJ 758, 2 and Bordoloi et al. 2014, ApJ 794, 2. The results presented in the manuscript should discuss their findings in this context.*

We acknowledge that we neglected to discuss studies on down-the-barrel absorption, which, just like those absorption studies based on background quasars that we did mention, also show bimodality that can be explained by a bipolar outflow pattern. There is also no information on the impact parameter of the down-the-barrel absorbers, and therefore, the scale of the probed outflow is unknown. In our work, we image for the first time the bipolar outflow (using an IFU, so even in 3D) and demonstrate it to be a common phenomenon, which we believe represents the most compelling direct evidence to date for the ubiquitous presence of galactic outflows in the circumgalactic medium.

We further respond to this comment in #13.

12• *“A correlation between galaxy MgII EW and stellar mass was first reported by…” This is a surprising statement, as this correlation has been well known for more than a decade. Several groups have independently reported it over the years (Weiner et al. 2009, ApJ 692, 1; Rubin et al 2010, ApJ, 719, 2; Kornei et al, 2012, Bordoloi et al 2014, Rubin et al. 2014). Please reword the discussion in this correct context.*

We are grateful for this correction. Although Weiner et al. 2009 and Rubin et al. 2010 do observe a correlation between MgII absorption EW and stellar mass, their correlation does not extend to the low-stellar mass regime where the transition between MgII absorption and MgII emission occurs. However, Kornei et al. 2012 do observe such a transition, albeit with large scatter. Therefore, we reword the corresponding sentence in the text and include Kornei et al. 2012 as reference.

13• *Detection of galactic winds in emission: The authors should note that spectacular detection of extended galactic wind in emission was recently reported by Rupke et al. 2019, Nature, 574, 7780 (OII) and in MgII by Shaban et al 2022, ApJ, 936, 1. There is significant evidence of detection of spatially extended outflows in emission and should be included in the discussion.*

We thank the referee for providing the references in #10, #11 and #13. We acknowledge that these previous works on individual galaxies are relevant and should have been mentioned in our discussion. Limited by the maximum number of references in the main text, we add them in a new paragraph in the Methods section (line346-line358).

These previous works mentioned by the referee can be classified into three categories: absorption studies against background quasars, studies of down-the-barrel absorption, and emission maps of the galactic outflow. The first two categories, while providing valuable insight into the phenomenon of galactic outflows, are indirect observations that are short of information on the impact parameter or shape of the galactic outflows. The emission maps are limited to only a few individual cases that may present unusual patterns of outflow. In particular, they do not provide a clear connection between extended emission and the galaxy azimuthal angle.

In contrast, our work presents direct 3D imaging of the cool and metal-enriched outflows, providing evidence that biconical outflows are common among the massive galaxies. We believe that this is a significant advancement over the previous studies mentioned and provides a more comprehensive understanding of galactic outflows.

Response to referee #2's comments:

The paper presents an interesting study of MgII line in stacked data cubes of galaxies at z~1. The authors' main objective is to examine the geometry of the MgII emission by creating narrowband images. They find that the MgII emission is extended up to 10 kpc from the primary galaxy.

One of the unique aspects of this study is that the authors stack the datacubes based on the morphology of the galaxies, specifically dividing them into edge-on (PA>55) and face-on systems (PA<30). By doing so, the authors can detect complex profiles of the MgII line in both types of galaxies. The authors claim to detect extended emission in only edge-on systems, which they attribute to the geometry of the outflows, i.e., the biconical nature of the outflows. They also find a faint ring-like emission around face-on galaxies, source of which is unclear.

The authors estimate an opening angle of 64^0 for the outflows. This study sheds light on the geometry of the outflows and their impact on the surrounding intergalactic medium at z~1. The paper is nicely written and presents key results in clear and concise manner. I recommend the manuscript for publication after addressing a few of my concerns.

Major Comments:

1. *The study reports that the total integration time in the edge-on systems is almost double compared to the face-on systems. However, it is unclear whether the lack of low signal-to-noise (S/N) extended emission in the face-on galaxies is due to a smaller sample size or other factors. The authors did not mention if they reached similar surface brightness in both stacked images.*

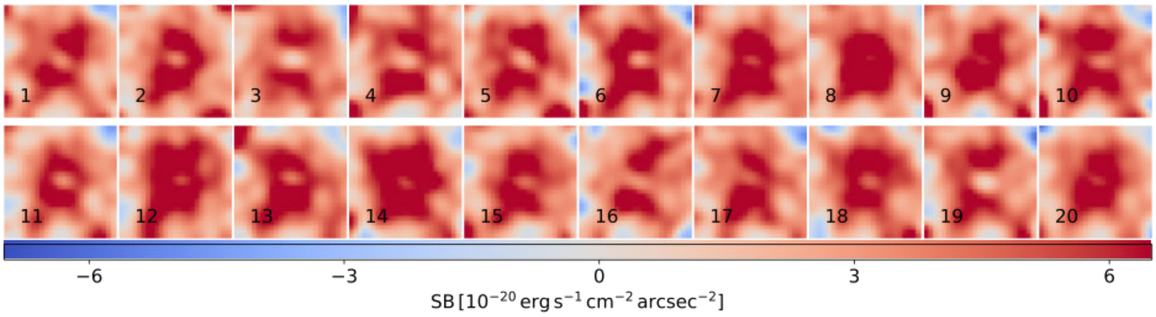
To address this question, it may be useful to randomly select the same number of galaxies from both subsamples and check whether the detection of extended emission in edge-on systems still holds. This could help rule out the possibility that the lack of extended emission in face-on galaxies is solely due to a smaller sample size.

Additionally, it would be helpful if the authors included the number of galaxies stacked in each subsample to provide a clearer picture of the statistical significance of their findings. Knowing the sample sizes can also help readers understand the sensitivity of the analysis and the degree of confidence in the results.

We thank the referee for pointing out this concern. The cumulative integration time for edge-on and face-on galaxy subsamples are different. The integration time for the two subsamples are 2790h and 1363h, respectively. We add a new plot of the S/N of the extended emission, as is presented in Extended Data Fig. 4, which clearly shows that the S/N for the face-on stacking is lower. We would also like to clarify that despite the lower S/N, the extended emission around the face-on galaxies, shown as a ring pattern, is also robust (Methods, section 4).

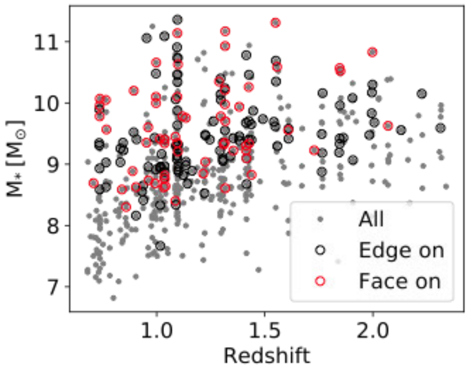
To further address this concern, we conduct a small bootstrap test. We randomly stack 60 galaxies from the edge-on subsample to match the total integration time of the face-on subsample. Several examples are shown in the figure below (panel 1-19). In panel 20, we stack the 75 galaxies at z<1.5, to further respond to the question in #2 about the morphology of galaxies at higher redshift. The main conclusion we can draw from this test is that the anisotropic MgII emission around the edge-on galaxies is prevalent. Even if the stack depth of the edge-on subsample is as low as that of the face-on subsample, we can still clearly see the bipolar shape, in contrast to the face-on galaxies that show absorption in the center and a “ring” at the outer part.

The number of galaxies at different subsamples is listed in line245 - line 246. We have 115 edge-on and 62 face-on galaxies.



2. *The extended data Figure 1 reveals that both face-on and edge-on systems have similar distributions of stellar mass and redshift. However, it is important to note that due to size-mass evolution, it may be challenging to accurately estimate the position angle (PA) of low stellar mass galaxies at high redshift, which could lead to them being misclassified as face-on. To further clarify this point, it would be helpful for the authors to present a scatter plot of stellar mass vs redshift for both subsamples. This would provide a clearer understanding of the distribution of galaxies in the two subsamples and highlight any potential biases that may exist due to size-mass evolution.*

Below we provide a plot of the redshift versus stellar mass of the edge-on (black circles), face-on (red circles) and all galaxies detected by MUSE at this redshift range (gray dots). As expected, this distribution is highly observationally biased, with only massive galaxies being well resolved at higher redshifts. To demonstrate that this observational effect does not bias our result, we stack only the galaxies at z<1.5, as is shown by panel 20 in the previous plot in Comment#1. We can still observe the bipolar MgII emission around the edge-on galaxies.



3. *The authors demonstrate that anisotropic outflows are only detected in the high-mass sample of edge-on galaxies, which is an interesting finding. The MgII emission is significantly brighter in the low stellar mass systems, as has been demonstrated in earlier works (e.g., Erb et al. in 2012). The MgII emission appears to be more extended than the stellar continuum, indicating that scattering of photons by the outflowing gas. Could the authors comment whether the irregular geometry of the MgII emission could be due to the irregular geometry of outflows in the low stellar mass systems as opposed to the biconical geometry in higher mass galaxies?*

It is a very insightful question regarding the comparison of the high-mass and low-mass systems. Our observations show that only high-mass systems exhibit strong bipolar emission, while low-mass systems show more isotropic emission (Fig. 3). Indeed, the isotropic emission around low mass systems could indicate a more isotropic geometry of outflows. For example, at higher redshift, the Lya haloes are well explained by isotropically expanding shells (e.g., Verhamme A. et al. 2006, A&A, 460, 397; Dijkstra M., Kramer R., 2012, MNRAS, 424, 1672 and their references). We know that the radiative transfer of MgII and Lya photons are very similar. If we believe that our lower-mass galaxies at z~1 and Lya emitters at z>3 are comparable systems, we could explain the observed isotropic MgII emission by the same physical model. However, we should also note that for lower-mass galaxies, the outflows are generally less strong. It is also possible that weak bipolar MgII emission is embedded in the strong isotropic MgII halo or even the PSF.

I mention in the main text (line102) the possibility that the outflows in the low-mass systems are more irregular. It is a pity that we can not discuss more in the paper. Due to the S/N of the data and the lack of good theoretical models, extra discussion on the isotropic emission could introduce more confusion rather than clarity to the readers. Nevertheless, we believe that our current observations provide valuable insights into the nature of outflows in different mass regimes, and we hope our results will inspire more investigations in the future.

4. *The continuum subtracted spectra from the central 1” in the edge-on system also shows the clear multiple kinematics components. In fact, the multiple components seem to be more significant in the edge-on system than in the face-on galaxies. If the outflows were purely biconical then I would not expect to detect any outflowing gas component in the central spectrum of edge-on galaxies.*

Also, could the authors use some statistical metric such as BIC-score to justify fitting of multiple components for both face-on and edge-on galaxies.

We acknowledge that the face-on and edge-on samples do not consist of PURELY face-on and edge-on galaxies, which we believe may be the primary reason for the potential multiple components in the edge-on system. Nonetheless, it is a common practice to decompose the absorption line into a quiescent component at the galaxy systemic velocity and an outflow component (e.g., Chen Y.M. et al., 2010, ApJ, 140, 2).

To further address this issue, we conduct several tests to compare the goodness of fit for the face-on and edge-on samples. The Chi2 values for the 2-gaussian fits of the face-on and edge-on samples are 0.25, 3.20, respectively. The R2 values for the two samples are 0.98, 0.92, respectively. The BIC values are 86.1, 122.0, respectively. We have also tried to fit the face-on and edge-on systems with one single gaussian. The BIC values are 114.1 and 128.6, respectively. These tests reveal that: (1) two-component fitting of the spectra is found to be better than one single gaussian component, though it is marginal in the case of the edge-on system, and (2) the two-component fitting of the face-on system is found to be better than that of the edge-on galaxies. However, as we have mentioned previously, it is reasonable that the spectra of the edge-on system contain multiple kinematic components, because the edge-on galaxies do cover a wide range of inclination angles. Moreover, for the edge-on system, the MgII absorption is much weaker.

Furthermore, in the text we include the error bars of the outflow velocity (Methods, line 337 and line 343) by generating 100 mock spectra with noise, which could also highlight the robustness of the decomposition.

Minor Comments:

The second sentence in the abstract “These outflows tend to form bipolar shapes above and below the galactic planes and extend well into the circumgalactic medium

(CGM), up to tens of kpc perpendicular to the galaxy." makes it seem like the biconical nature of outflows is well observed and accepted phenomena. It undermines the significance of their findings. Maybe reword to something like "Theoretical models assumes bipolar shapes for the outflows that extends well into the circumgalctic medium (CGM), up to tens of Kpc perpendicular to the galaxies. However, they have been…"

We thank the referee for highlighting the significance of this work. We have revised this sentence.

Reviewer Reports on the First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

I read the author response to the revised manuscript titled "Observational Evidence of the Prevalence of Bipolar Galactic Outflows out to 10 kpc at z ≈ 1 for Massive Galaxies".

The manuscript still has several issues that need addressing.

The manuscript attempts to quantify the significance of detection (for Figure 1) by producing a SNR map of the stacks. This is presented in Extended Data Figure 4. However, this information is not used to quantify the findings presented in the main article. The reader is still left to visually inspect Figure 1 and Extended Data Figure 4 to infer what regions might be statistically significant. This is a very subjective approach.

I would strongly recommend that the authors follow my original suggestion of creating significance contours and overlaying them on all their stacks. Each stack (Figures 1, 2, 3) will have their own significance maps and that should be accordingly propagated throughout the paper, and not relegated to the methods section only.

This lack of statistical rigor in the "main article" is still my primary concern for this manuscript. As presented, the main result of ~10 kpc extended emission is derived by visually inspecting the figure. It should be quantified for the pixels which are detected with statistical significance and proper error bars inferred.

The authors should also quantify the uncertainty coming into this distance estimate owing to the smoothing kernel used and atmospheric seeing.

Only in the method section the manuscript quotes a formally derived distance of 9+/-2 kpc by taking all the weighted flux. This should be revisited for every stack and computed for statistically significant pixels only.

Same concern holds for Figure 2 and Figure 3 in the main article.

The manuscript should also infer the size of the stellar emission region in the stacks and report them so that one may compare the size of the MgII emission with the stellar continuum emission directly.

Extended Data Fig 6 and Extended data section 3: All extracted spectra should be quantified for robustness of the signal. Rest frame emission line equivalent width for each emission line (with associated errors) should be computed and for each panel, the corresponding statistically significant emission should be marked out.

Minor Issues:
In the main article it says that MgII line can be observed from the ground at z>0.5, that is not true. One can push all the way down to the atmospheric cutoff and detect MgII at z~0.1 (see e.g., Kacprazk et al, ApJ Letters, 777, 1, 2013).

Page 3, Line 97: Figure is mislabeled. Perhaps the authors meant Figure 3?

Mass outflow rate discussion: Since this is really a back of the envelope calculation with major uncertainties in almost all the important parameters assumed, this discussion should not be in the main article and moved to the methods section only.

Further, while I appreciate that these are crude back of the envelope calculations, one must be mindful that (1) instrument resolution effects are not accounted for (2) MgII outflowing gas covering fraction different galaxies will be different, making non-linear contribution to the measurements. clearly the absorption line ratios seen in Figure 4 are off for MgII doublet. Please add a few sentences discussing these caveats.

Stacking method: When the cubes are shifted into the rest frame before stacking, is the impact of cosmological dimming accounted for? Please clarify in the text. Since the redshift range span between ~ 0.5<z<2.3 this will be important in the stacks.

Extended Data Figure 7: Kindly quantify the emission strength by computing the EW consistent with my recommendation for other Figures.

Several of the edge on galaxy sample stamps look like they are undergoing mergers (e.g., bottom row 3rd object, 3rd row last object). I would recommend that the authors check any such objects for companions. They may want to exclude them from the stacks to avoid further confusion. Again, I doubt that will change the main result of the paper, but will get rid of any dubious objects.

Referee #2 (Remarks to the Author):

I am satisfied with the author's response and changes made to the revised manuscript. I recommend the revised manuscript for publication.

Author Rebuttals to First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

I read the author response to the revised manuscript titled “Observational Evidence of the Prevalence of Bipolar Galactic Outflows out to 10 kpc at z ≈ 1 for Massive Galaxies”.

The manuscript still has several issues that need addressing.

Response to the referee:
We thank the referee for the helpful suggestions. We find these suggestions to be pertinent and significant, as they assist us in further justifying the statistical significance of our results. Below we will address each point raised by the referee. All the changes are marked in blue in the new manuscript.

The manuscript attempts to quantify the significance of detection (for Figure 1) by producing a SNR map of the stacks. This is presented in Extended Data Figure 4. However, this information is not used to quantify the findings presented in the main article. The reader is still left to visually inspect Figure 1 and Extended Data Figure 4 to infer what regions might be statistically significant. This is a very subjective approach.

I would strongly recommend that the authors follow my original suggestion of creating significance contours and overlaying them on all their stacks. Each stack (Figures 1, 2, 3) will have their own significance maps and that should be accordingly propagated throughout the paper, and not relegated to the methods section only.

Response to the referee:
We agree with the referee's suggestion that utilizing significance contours is a better approach for visualizing the detection significance. We upgrade Figures 1,2 and 3 by adding significance contours of 2,4 and 6- sigma levels. Also, we have removed the S/N map (previously Extended Data Figure 4) from the revised version.

In the Methods section (line268-line275), we include an explanation of the procedures to calculate the noise. We produce 50 mock NB images at the wavelength range near the MgII line (+/-80A). All these mock NB images are smoothed using the same kernel as in Figures1,2 and 3. We calculate the standard deviation of these images to generate a noise image. This noise image is then utilized to obtain the S/N map contours presented in Figures 1, 2, and 3 in the revised manuscript.

This lack of statistical rigor in the “main article” is still my primary concern for this manuscript. As presented, the main result of ~10 kpc extended emission is derived by visually inspecting the figure. It should be quantified for the pixels which are detected with statistical significance and proper error bars inferred.

The authors should also quantify the uncertainty coming into this distance estimate owing to the smoothing kernel used and atmospheric seeing.

Only in the method section the manuscript quotes a formally derived distance of 9+/-2 kpc by taking all the weighted flux. This should be revisited for every stack and computed for statistically significant pixels only.

Same concern holds for Figure 2 and Figure 3 in the main article.

Response to the referee:
As the anisotropic MgII emission represents the major discovery in our work, we have included a new section in the Methods (line283-line311) to analyze the extent of MgII emission along the minor axis.

We calculate the distance of the MgII emission to the galaxy disk by measuring the flux-weighted average of all pixels with S/N higher than 6-sigma in Figure 1. This measurement provides an estimate of the spatial extent of the MgII emission. To assess the uncertainty associated with this measurement, we generate 100 mock NB images by adding Gaussian noise. The error bar on the distance was then estimated based on these mock NB images.

For the mean- and median-stacked datacubes, the distance of the MgII emission was found to be 9.6 +/- 1.7 kpc and 9.4 +/- 2.0 kpc, respectively.

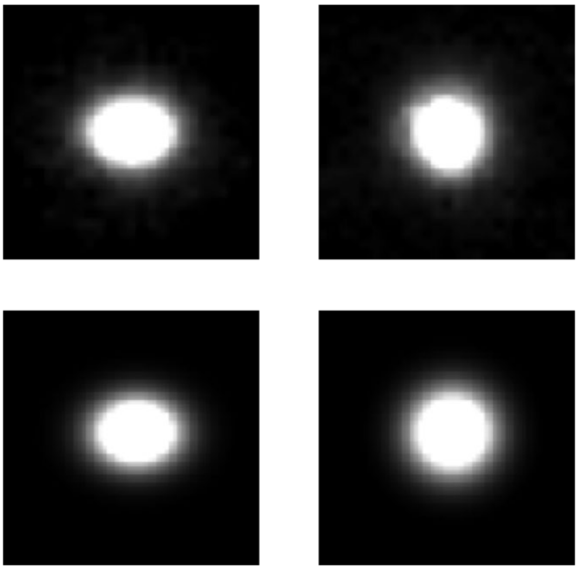
Additionally, we perform the same distance measurement for the left panel (high-mass subsample) of Figure 3. In this case, the distance of the MgII emission was determined to be 10.2 +/- 2.5 kpc.

The manuscript should also infer the size of the stellar emission region in the stacks and report them so that one may compare the size of the MgII emission with the stellar continuum emission directly.

Response to the referee:
The size of the galaxy's stellar continuum is an important comparison. We fit a 2D Sersic surface brightness model to the stacked MUSE white light image (line 290-293, line296-298). Before fitting, we convolved the image with the same Gaussian kernel as applied to the MgII maps. Below we present an example of the

fitting. The upper panels display the stacked MUSE white light image, and the lower two panels show the fitted Sersic surface brightness model. For the edge-on galaxy (full sample), we measure the effective radius of the minor axis 3.2+/-0.3 kpc.

We also perform the same measurement for high-mass subsample. The minor axis effective radius of the edge-on galaxy is 3.6+/-0.3 kpc.



Extended Data Fig 6 and Extended data section 3: All extracted spectra should be quantified for robustness of the signal. Rest frame emission line equivalent width for each emission line (with associated errors) should be computed and for each panel, the corresponding statistically significant emission should be marked out.

Response to the referee:
We agree that all the spectra in the manuscript should be quantified for robustness of the signal. To address this, we indicate the error range in Figure 4 and Extended Data Figures 5 and 6. The error of each spectrum is estimated by considering the flux adjacent to the MgII line with a range of +/-80A.

Regarding Extended Data Figure 5, we acknowledge that the figure contains multiple subplots and each subplot contains two spectra, which can make it challenging to read. To enhance clarity, we emphasize the panels where the MgII line peak flux of the edge-on galaxy (black spectra) exceeds 2 sigma by showing these panels with deeper colors, while showing the remaining panels with lighter colors.

Regarding the equivalent width (EW) of the lines, in our initial calculations, we did not have information about the continuum because we subtracted it at the very beginning. However, to investigate the EW of the MgII emission and absorption, we have now constructed a median MgII continuum image. With the median filtering (line254), we have obtained a continuum datacube. We generate pseudo NBs from each continuum datacube, while appropriately masking the neighbors. We performed median stacking of the NBs, resulting in a final MgII continuum image.

[Redacted text]

Based on the MgII emission map and the continuum map, we have calculated the distribution of the MgII EW. As expected, the MgII EW at the edge of the Galaxy can be significantly high (>50A) due to the absence of continuum in these areas. Also, the arithmetic division of the MgII emission and the continuum (which is a very small number in the outer region) can greatly increase the noise.Therefore, in the manuscript we do not provide the EW for each panel of ExtendedData Figure 5. However, for the key spectra with high S/N MgII emission/absorption lines (all spectra in Figure 4 and Extended Data Figure 6), we provide the MgII EW as a reference (line345, line352, lines360-362, and also in the caption of the corresponding figure).

[Redacted figure]

Minor Issues:
In the main article it says that MgII line can be observed from the ground at z>0.5, that is not true. One can push all the way down to the atmospheric cutoff and detect MgII at z~0.1 (see e.g., Kacprazk et al, ApJ Letters, 777, 1, 2013).

Response to the referee:
We thank the referee for pointing out this problem. We have made the change in the manuscript as suggested (line54).

Page 3, Line 97: Figure is mislabeled. Perhaps the authors meant Figure 3?

Response to the referee:
We are sorry for this mistake. We have corrected it (line95).

Mass outflow rate discussion: Since this is really a back of the envelope calculation with major uncertainties in almost all the important parameters assumed, this discussion should not be in the main article and moved to the methods section only.

Further, while I appreciate that these are crude back of the envelope calculations, one must be mindful that (1) instrument resolution effects are not accounted for (2) MgII outflowing gas covering fraction different galaxies will be different, making non-linear contribution to the measurements. clearly the absorption line ratios seen in Figure 4 are off for MgII doublet. Please add a few sentences discussing these caveats.

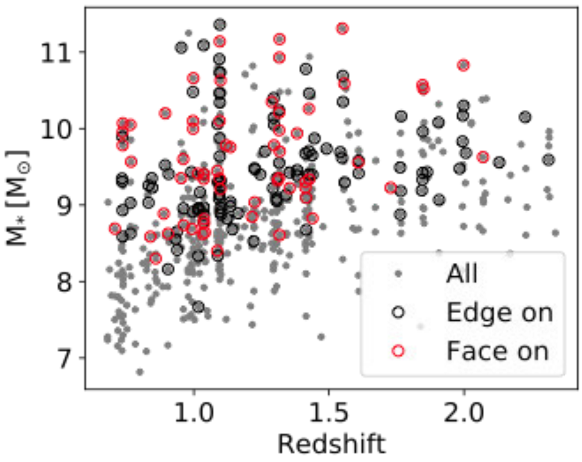
Response to the referee:
We moved the paragraph regarding the order-of-magnitude estimation from the main article to the methods section.

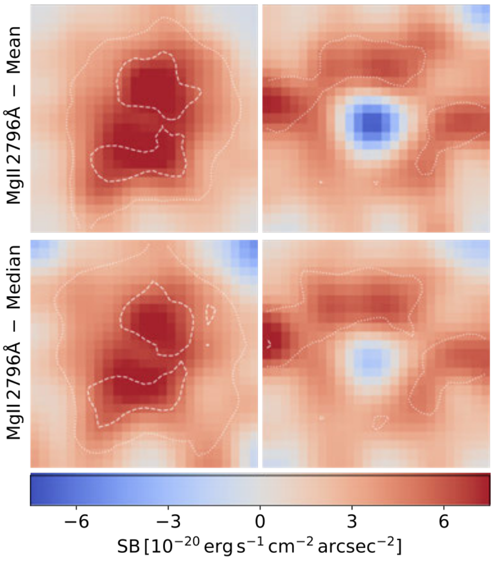
We thank the referee for pointing out these caveats for the order-of-magnitude estimation. We have included a dedicated discussion on this topic in the manuscript (line414-418).

Stacking method: When the cubes are shifted into the rest frame before stacking, is the impact of cosmological dimming accounted for? Please clarify in the text. Since the redshift range span between ~ 0.5<z<2.3 this will be important in the stacks.

Response to the referee:
We did not correct for cosmological dimming when we did the stacking, because it would have a non-linear effect on the error of each cube. We clarified this in line258-259.

Additionally, in response to a similar previous suggestion from the other referee, we conducted an additional test to address the potential influence of redshift evolution in our sample. Below we provide a plot of the redshift versus stellar mass of the edge-on (black circles), face-on (red circles) and all galaxies detected by MUSE at this redshift range (gray dots). As expected, this distribution exhibits a slight observational bias, with only massive galaxies being well resolved at higher redshifts. To demonstrate that this observational effect does not bias our result, we stack only the galaxies at z<1.5 (below, right figure). Despite the lower signal-to-noise ratio (S/N), we can still observe the bipolar MgII emission around the edge-on galaxies, further supporting the robustness of our findings.





Extended Data Figure 7: Kindly quantify the emission strength by computing the EW consistent with my recommendation for other Figures.

Response to the referee:
We calculate the EW of MgII 2796 emission/absorption lines in Extended Data Figure 6 (Extended Data Figure 7 in the old version) (line345, line352) and Figure 4 (lines360-362). The EW of these lines are also reported in the captions of the corresponding figures.

Several of the edge on galaxy sample stamps look like they are undergoing mergers (e.g., bottom row 3rd object, 3rd row last object). I would recommend that the authors check any such objects for companions. They may want to exclude them from the stacks to avoid further confusion. Again, I doubt that will change the main result of the paper, but will get rid of any dubious objects.

Response to the referee:
We removed all pairs of galaxies with low values of projected angular separations (2") and small line-of-sight relative radial velocities (1000km/s). We have taken a conservative approach, considering that previous studies such as Ventou et al. 2017, A&A, 608, 9 used a line-of-sight relative radial velocity difference of 500 km/s. As a result, we exclude 3 edge-on galaxies and 2 face-on galaxies from our sample. The revised sample now consists of 112 edge-on galaxies and 60 face-on galaxies (line236-237).

With this updated sample, we redid the stacks and updated the MgII spectra and maps in the manuscript. It is important to note that the differences between the old and new stacking results are almost negligible. The numbers reported in the manuscript have been adjusted slightly to reflect these changes.

Regarding the two objects specifically mentioned by the referee, we confirm that the first object (previously located at the bottom row, 3rd object of the Extended Data Figure 3 of the old version) has been removed from the sample due to the presence of a close neighbor with a velocity difference less than 1000 km/s. Another object (3rd row, last object in the Extended Data Figure 3 in the old version of the manuscript, 3rd row second last object in the new version) is not undergoing a merger. The nearby bright neighbor with a very small projected separation is a galaxy at z=0.3.

Referee #1 (Comments on 2nd Revision):

I am satisfied with the author's response and changes made to the revised manuscript. I recommend the revised manuscript for publication.

Response to the referee:
We thank the referee for the valuable suggestions and the recognition of our work. The suggestions have undoubtedly contributed to the improvement of our manuscript. Thank you for your time and consideration.