



The Young Massive Star Cluster Westerlund 2 Observed with MUSE. II. MUSEpack—A Python Package to Analyze the Kinematics of Young Star Clusters

Peter Zeidler^{1,2} , Antonella Nota^{2,3}, Elena Sabbi² , Peter Luljak^{2,4}, Anna F. McLeod^{5,6} , Eva K. Grebel⁷, Anna Pasquali⁷ , and Monica Tosi⁸

¹ Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA; zeidler@stsci.edu

² Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

³ ESA, SRE Operations Division, Spain

⁴ Harvey Mudd College, 301 Platt Boulevard, Claremont, CA 91711, USA

⁵ Department of Astronomy, University of California Berkeley, Berkeley, CA 94720, USA

⁶ Department of Physics and Astronomy, Texas Tech University, PO Box 41051, Lubbock, TX 79409, USA

⁷ Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstraße 12–14, D-68120 Heidelberg, Germany

⁸ INAF-Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Gobetti 93/3, I-40129 Bologna, Italy

Received 2019 August 9; revised 2019 September 12; accepted 2019 September 14; published 2019 October 25

Abstract

We mapped the Galactic young massive star cluster Westerlund 2 with the integral field spectrograph MUSE (spatial resolution: $0.2 \text{ arcsec px}^{-1}$, spectral resolution: $\Delta\lambda = 1.25 \text{ Å}$, wavelength range: $4600\text{--}9350 \text{ Å}$) mounted on the Very Large Telescope. We present the fully reduced data set and introduce our new Python package “MUSEpack,” which we developed to measure stellar radial velocities (RVs) with an absolute precision of $1\text{--}2 \text{ km s}^{-1}$ without the necessity of a spectral template library. This novel method uses the two-dimensional spectra and an atomic transition line library to create templates around strong absorption lines for each individual star. Automatic, multi-core processing makes it possible to efficiently determine stellar RVs of a large number of stars with the necessary precision to measure the velocity dispersion of young star clusters. MUSEpack also provides an enhanced method for removing telluric lines in crowded fields without sky exposures, and a Python wrapper for ESO’s data reduction pipeline. We observed Westerlund 2 with a total of 11 short and 5 long exposures (survey area: $\sim 11 \text{ arcmin}^2$ or 15.8 pc^2) to cover the bright nebular emission and OB stars, as well as the fainter pre-main-sequence stars ($\geq 1 M_{\odot}$). We extracted 1725 stellar spectra with a mean signal-to-noise ratio of $S/N > 5$ per pixel. Typical RV uncertainties of 4.78 km s^{-1} , 2.92 km s^{-1} , and 1.1 km s^{-1} are reached for stars with a mean $S/N > 10$, $S/N > 20$, and $S/N > 50$ per pixel, respectively. It is possible to reach RV accuracies of 0.9 km s^{-1} , 1.3 km s^{-1} , and 2.2 km s^{-1} with ≥ 5 , 3–4, and 1–2 spectral lines used to measure the RVs, respectively. The combined statistical uncertainty on the RV measurements is 1.10 km s^{-1} .

Unified Astronomy Thesaurus concepts: Radial velocity (1332); Young star clusters (1833); Spectroscopy (1558); Pre-main sequence stars (1290)

1. Introduction

The majority of stars are born in clustered environments through the collapse of giant molecular clouds (GMCs, e.g., Lada & Lada 2003). The initial mass and density of the GMC define the number of stars formed in the newly born star cluster and its stellar density. The stellar densities range from low-density systems in the Milky Way (MW), such as the Upper Scorpius OB association ($\leq 0.1 M_{\odot} \text{ pc}^{-3}$, e.g., Preibisch & Mamajek 2008), via loosely bound open clusters, to more massive systems, such as Westerlund 2 (hereafter Wd2) ($\sim 6.6 \times 10^3 M_{\odot} \text{ pc}^{-3}$, e.g., Westerlund 1961; Zeidler et al. 2015, 2016, 2017, 2018), NGC 346 ($\sim 64 M_{\odot} \text{ pc}^{-3}$, e.g., Sabbi et al. 2008), or NGC 3603 ($\sim 1.5 \times 10^4 M_{\odot} \text{ pc}^{-3}$, e.g., Sung & Bessell 2004; Stolte et al. 2004; Rochau et al. 2010; Pang et al. 2011, 2013), to the most massive young superstar clusters, such as R136 ($(4.8\text{--}24) \times 10^4 M_{\odot} \text{ pc}^{-3}$, e.g., Hunter et al. 1995; Crowther et al. 2016) in the Large Magellanic Cloud (LMC). Despite numerous studies of young star clusters in the MW and Magellanic Clouds, their formation processes and the cluster evolution in the first few million years are not well understood. On the one hand, we observe star clusters that appear to have formed in one central starburst (monolithic cluster formation, e.g., Lada et al. 1984; Bastian & Goodwin 2006), as was suggested for the central star cluster HD 97950 in NGC 3603

(Banerjee & Kroupa 2015). We also observe hierarchical cluster formation, in which a star cluster forms via subsequent merging of smaller sub-clusters (Larson 1981; Longmore et al. 2014; Dale et al. 2015), which is seen for many star clusters in the MW and Magellanic Clouds (e.g., Wd2, Zeidler et al. 2015; NGC 346, Sabbi et al. 2007, 2008) or in multiple star clusters of the H II region 30 Dor (Sabbi et al. 2012, 2013, 2016). The subsequent evolution of a newly born star cluster is regulated not only by the stellar density and total mass (the stellar mass function), but also by stellar kinematics.

While it is possible to measure stellar proper motions using long-baseline observations from ground and space, such as multi-epoch *Hubble Space Telescope* (HST) observations (e.g., E. Sabbi et al. 2019, in preparation), or the *Gaia* satellite (Gaia Collaboration et al. 2016, 2018), the development of wide-field integral field units (IFUs) has made it feasible to spectroscopically map resolved, nearby star-forming regions and young star clusters so as to also measure the radial velocities (RVs) with a similar telescope-time efficiency, thus providing 3D stellar kinematics.

The large field of view (FOV) of 1 arcmin^2 of the Multi Unit Spectrographic Explorer (MUSE, Bacon et al. 2010) mounted in the Nasmyth focus of UT4 at the Very Large Telescope (VLT) at the European Southern Observatory (ESO), Chile, has proven to be very effective in studying stellar RVs and gas kinematics,

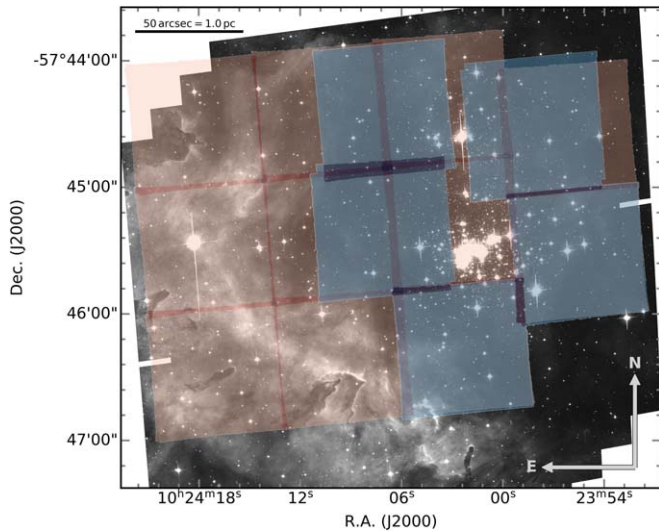


Figure 1. Coverage of the MUSE data plotted over the *HST* F814W image of Wd2. The blue fields show the short exposures, while the red fields show the long exposures. In total we covered $\sim 11 \text{ arcmin}^2$, which corresponds to 7.6 pc^2 at the distance of Wd2.

including feedback processes in star-forming regions and young star clusters (e.g., McLeod et al. 2015, 2019; Zeidler et al. 2018). Stellar RVs can be measured to an accuracy of $\sim 1 \text{ km s}^{-1}$ by cross-correlating the extracted spectra with spectral templates (Kamann et al. 2013, 2016, 2018; Zeidler et al. 2018). The spatial sampling of MUSE is $0.2 \text{ arcsec px}^{-1}$ with a spectral resolution of $\Delta\lambda = 1.25 \text{ \AA}$ when observing at optical wavelengths between 4600 and 9350 \AA .

For young star clusters, where the majority of stars are still in their pre-main-sequence phase no proper optical spectral template libraries exist. Our new method, presented in this work, is independent of spectral template libraries. We show that our Python package “MUSEpack” is able to measure RVs with an accuracy of $\sim 1 \text{ km s}^{-1}$ when applied to our VLT/MUSE, medium-resolution spectra of Wd2. The package has been developed specifically for MUSE data but can be used for any kind of IFU data set.

This paper is structured as follows: in Section 2 we introduce the work that has been done on the Galactic young star cluster Wd2. Section 3 gives an overview of the MUSE data set on Wd2. In Section 4 we provide an overview of MUSEpack, while Section 5 provides a detailed description of the data reduction process. Section 6 gives an introduction to stellar RV fitting and the performance of MUSEpack. In Section 7 we introduce the created stellar RV catalog of Wd2. In Section 8 we summarize and conclude the work.

2. Westerlund 2—A Young Massive Star Cluster

We are conducting a photometric and spectroscopic survey of the young massive star cluster (YMC) Wd2 (Westerlund 1961) in the H II region RCW49 (Rodgers et al. 1960) using 16 MUSE pointings (097.C-0044(A), 099.C-0248(A), P.I.: P. Zeidler; see Figure 1) with two different exposure times to obtain stellar RVs down to $\sim 1 M_{\odot}$ in combination with multi-band optical and near-infrared *HST* observations (ID: 13038, P.I.: A. Nota). Wd2 is a Galactic YMC located in the Sagittarius spiral arm, at a distance of $\sim 4 \text{ kpc}$ and with an age of $< 2 \text{ Myr}$. It consists of two coeval sub-clumps, namely the Main Cluster and the Northern Clump (Vargas Álvarez et al. 2013; Hur et al. 2015; Zeidler et al. 2015).

We derived the stellar mass function (Zeidler et al. 2017) and showed that Wd2 is highly mass-segregated, which is likely primordial, given its young age. In Zeidler et al. (2018), we presented six MUSE pointings of the cluster center that were observed in the first (097.C-0044(A)) of two observing runs. Using these data we showed that (1) it is possible to obtain stellar and gas RVs with an accuracy of $\sim 1 \text{ km s}^{-1}$ in nearby YMCs, (2) we can measure spectral types of main-sequence and pre-main-sequence stars, and (3) Wd2 shows a bimodal velocity structure with two peaks at $(8.10 \pm 1.53) \text{ km s}^{-1}$ and $(25.41 \pm 3.15) \text{ km s}^{-1}$. The RV dispersion of both peaks is in agreement with the RV dispersion seen in other YMCs such as Trumpler 14 (Kiminki & Smith 2018), the Carina Nebula (e.g., Smith & Brooks 2008), or NGC 3603 (Rochau et al. 2010; Pang et al. 2013). Because of the limited data of only 72 stars, we could not properly address the physical origin of such a bimodal RV distribution.

3. The Data Set

We observed Wd2 with MUSE in extended mode,⁹ both with and without adaptive optics (AO) during the ESO periods 97 and 99 (Program ID: 097.C-0044(A), 099.C-0248(A), P.I.: P. Zeidler). In total, we acquired 16 pointings with three dither positions each, with a total telescope time (exposure time + overheads) of almost 21 hr. The dither pattern follows a 90° and 180° rotation strategy to minimize detector defects and impurities. Eleven out of the 16 pointings are obtained with a total exposure time of 660 s (hereafter: short exposures), while the remaining five pointings have a total exposure time of 10,800 s (hereafter: long exposures). An overview of all pointings is given in Table 1.

The two different exposure times were chosen to cover a large dynamic range. The short exposures (blue fields in Figure 1) cover the complete cluster area including the gas rim of the surrounding H II region to study the massive stars and the gas, while the long exposures (red fields in Figure 1) are designed to study the fainter pre-main-sequence population. The long exposures cover the immediate vicinity of the cluster core, avoiding the central regions and the gas, in order to avoid saturating a major fraction of the data cubes.

Four out of the five long exposures are observed using the wide-field AO mode of MUSE. This mode uses the four laser guide stars for ground-layer correction of the atmosphere’s distortions, resulting in an improvement of the seeing by a factor of two. Because the lasers are Na-lasers, a notch filter must be used to block the light between 5780 and 5990 \AA , which results in a gap in the spectrum.¹⁰

4. MUSEPACK

To reduce and analyze our IFU data set of the YMC Wd2, we developed MUSEpack. It is a Python-based package, which contains two main classes¹¹: MUSEreduce and RV_spectrum. MUSEpack can be downloaded from <https://github.com/pzeidler89/MUSEpack> under a GPL version 3 license

⁹ MUSE can observe with two spectral ranges, the nominal mode (4800–9350 \AA) and the extended mode (4600–9350 \AA).

¹⁰ For detailed information we refer to the VLT/MUSE handbook <https://www.eso.org/sci/facilities/paranal/instruments/muse/doc.html>.

¹¹ A class is a special construct for object-oriented programming languages that has the capability of bundling data and functionality together.

and Version 1 of this software is archived on Zenodo (Zeidler 2019). We will introduce the capabilities of the various classes and their modules throughout this paper. A full description of the package can be found on <https://musepack.readthedocs.io/en/latest/>. There, we also provide some basic examples to show the features of MUSEpack and to provide the user with an easy start.

4.1. MUSEreduce

MUSEreduce is a Python wrapper¹² for the standard reduction pipeline (Weillbacher et al. 2012, 2014), based on ESO¹³ Reflex (ESORex), which provides a user-friendly json file¹⁴ to set the input parameters for ESORex. Additionally, MUSEreduce is meant to sort the raw data and assign the appropriate calibration files to each science exposure. In contrast to ESO’s graphical user interface (GUI)-based data manager GASGANO¹⁵, MUSEreduce is Python-based, which makes it platform-independent and easy to use for remote execution via a secure shell (SSH) connection.¹⁶ MUSEreduce supports the three main observing modes of MUSE: the wide-field mode without adaptive optics (WFM-NOAO), the wide-field mode with adaptive optics (WFM-AO), and the narrow-field mode with adaptive optics (NFM-AO). We used data obtained with the three different observing modes to assess the quality of the code. For the WFM-NOAO and WFM-AO we used data of the Galactic YMC Wd2 (097.C-0044(A), 099.C-0248(A), P.I.: P. Zeidler) representative of a crowded field with a high dynamic range (Zeidler et al. 2018, and this work). We also used data of the blue compact dwarf galaxy J0291+0721 (096.B-0212(A), P.I.: B. James), a faint extended object (James et al. 2019). The data of J0291+0721 were obtained including two sky fields. To test the NFM-AO we used observations of two T-Tauri disks (60.A-9482 (A), P.I.: J. Girard), which were also obtained using sky fields (J. Girard et al. 2019, in preparation).

4.2. RV_spectrum

RV_spectrum is the main workhorse routine to fit spectral lines to extracted stellar and gas spectra and to fit RVs with an accuracy of 1–2 km s⁻¹. The novelty compared to other RV measurement techniques is that it does not require stellar spectral template libraries. To be able to use the full capabilities of this class the user should install the non-standard Python packages PampelMuse (Kamann et al. 2013), Penalized Pixel-Fitting (pPXF Cappellari & Emsellem 2004; Cappellari 2017), and pyspeckit (Ginsburg & Mirocha 2011).

5. Data Reduction

The data were reduced using MUSEreduce, a Python class within MUSEpack (see Section 4) using ESO’s standard reduction pipeline (v.2.4.2). When available, we used the reduced calibration files (MASTERFLAT, MASTERBIAS, WAVECAL, LSF, TWI-LIGHT) provided by ESO. If the reduced calibration files were not provided, we executed the respective data reduction steps to create them from the raw calibration files. All intermediate and final

products were checked by eye in order to ensure that the data reduction processes and the alignment of the dither positions were executed correctly.

During these quality checks we discovered a moiré-like pattern in the southernmost long exposure (LONG_1). After a thorough analysis we concluded that it is most likely due to the sky conditions under which the individual dither positions were obtained. While the exposures LONG_1a and LONG_1c were obtained with Moon condition “bright,” the exposure LONG_1b was observed with Moon condition “dark” (during a different night, see Table 1). We used a simulated lunar spectrum together with the sky spectrum of the exposure LONG_1b to model the lunar contamination and to remove it from the data cube. This procedure is described in detail in the Appendix.

5.1. Telluric Line Correction in Crowded H II Regions

To subtract the sky (dominated by telluric emission lines) from the science exposures, the MUSE data reduction pipeline uses a sky spectrum extracted from either a sky observation or the darkest regions of the science observation itself. This sky spectrum is convolved with the local, position-dependent line-spread function (LSF) and refitted to each spaxel together with a low-order continuum. Subsequently, it is subtracted from the data cube.

For many observations, such as Galactic star clusters, obtaining sky fields is not feasible since the telescope has to be moved by several degrees to avoid observing the sky inside the parental gas and dust cloud. The high stellar crowding within these young star clusters and the gas and dust of the H II region often do not leave enough “empty” regions to properly determine a sky spectrum, yet a proper sky subtraction is necessary to remove contamination from the mostly dominant emission lines created by the Earth’s atmosphere. Some of these telluric emission lines have the same wavelengths as typical emission lines from H II regions, which are typically the Balmer lines, the oxygen lines, and atomic and molecular transition lines of oxygen, hydrogen, and some other trace elements. As a result, the extracted sky spectra overestimate the fluxes for these emission lines, which leads to an over-subtraction with a possible shift of the centroid.

To improve this, we altered the process of how the data reduction pipeline handles the estimation of the sky spectrum under two main assumptions: (1) the fluxes of the emission lines from the H II region are much stronger than the fluxes emitted by our atmosphere, and (2) the continuum of the H II region is dominant over the telluric continuum. The sky spectrum is then extracted in the following manner.

1. The data reduction pipeline performs a regular sky line and continuum extraction (`muse_create_sky`) for the darkest spaxels of each data cube, creating the two files “SKY_LINES.fits” and “SKY_CONTINUUM.fits” that contain the fitted sky line fluxes and the continuum, respectively.
2. The fluxes of all emission lines besides the OH and O₂ lines are set to zero in the “SKY_LINES.fits” file.
3. The continuum flux is set to zero to avoid subtracting the continuum of the H II region.
4. `muse_create_sky` is run again with the new “SKY_LINES.fits” and “SKY_CONTINUUM.fits” files as input. This is necessary to properly determine the fluxes of the OH and O₂ emission lines that are blended with other emission lines (e.g., [O I] λ 5577 and OH(1601)X-X(0701) Q11(03.5f) at $\lambda = 5577.58273$ Å).

¹² A Python script that calls functions in other programming languages.

¹³ <https://www.eso.org/sci/software/pipelines/muse/>

¹⁴ A json or JavaScript Object Notation file is an open-standard, human-readable text file format containing data objects as attribute-value pairs.

¹⁵ <https://www.eso.org/sci/software/gasgano.html>

¹⁶ An SSH connection is a cryptographic network protocol to connect network services securely over an unsecured network.

Table 1
Overview of the Observations

Coordinates		Date	OB-Name	OB-ID	t_{exp} (s)	Seeing (arcsec)	Airmass	AO
R.A.	Decl.							
10:24:02.25	−57:46:17.36	2018 Mar 6	LONG_1a	1661895	3600	0.41–0.43	1.255–1.370	no
10:24:02.25	−57:46:17.36	2017 Dec 21	LONG_1b	1661899	3600	0.71–0.58	1.251–1.201	no
10:24:02.25	−57:46:17.36	2018 Mar 6	LONG_1c	1661903	3600	0.54–0.54	1.383–1.593	no
10:24:07.17	−57:45:22.75	2018 Mar 10	LONG_2a_AO	1987342	3600	0.94–0.51	1.466–1.309	yes
10:24:07.17	−57:45:22.75	2018 Mar 11	LONG_2b_AO	1987346	3600	1.00–0.96	1.282–1.212	yes
10:24:07.17	−57:45:22.75	2018 Mar 11	LONG_2c_AO	1987350	3600	0.53–0.49	1.206–1.197	yes
10:23:55.47	−57:45:31.39	2018 Mar 12	LONG_3a_AO	1987371	3600	0.59–0.46	1.197–1.236	yes
10:23:55.47	−57:45:31.39	2018 Mar 11	LONG_3b_AO	1987367	3600	0.44–0.57	1.200–1.249	yes
10:23:55.47	−57:45:31.39	2018 Mar 12	LONG_3c_AO	1987375	3600	0.59–0.58	1.279–1.413	yes
10:23:58.21	−57:44:34.58	2018 Mar 13	LONG_4a_AO	1987380	3600	0.43–0.47	1.468–1.310	yes
10:23:58.21	−57:44:34.58	2018 Mar 13	LONG_4b_AO	1987384	3600	0.54–0.43	1.294–1.217	yes
10:23:58.21	−57:44:34.58	2018 Mar 13	LONG_4c_AO	1987388	3600	0.57–0.42	1.211–1.196	yes
10:24:06.89	−57:44:24.14	2018 Mar 14	LONG_5a_AO	1987393	3600	0.41–0.57	1.241–1.198	yes
10:24:06.89	−57:44:24.14	2018 Mar 14	LONG_5b_AO	1987397	3600	0.88–0.64	1.195–1.211	yes
10:24:06.89	−57:44:24.14	2018 Mar 20	LONG_5c_AO	1987401	3600	0.67–0.57	1.335–1.237	yes
10:23:56.11	−57:44:33.14	2018 Jan 21	SHORT-1	1661920	220	1.26–1.50	1.234–1.240	no
10:23:56.11	−57:44:33.14	2018 Jan 21	SHORT-1	1661920	220	1.47–1.30	1.247–1.253	no
10:23:56.11	−57:44:33.14	2018 Jan 21	SHORT-1	1661920	220	1.50–1.41	1.240–1.246	no
10:23:55.48	−57:45:31.43	2018 Jan 21	SHORT-1	1661920	220	1.30–1.24	1.254–1.260	no
10:23:55.48	−57:45:31.43	2018 Jan 21	SHORT-1	1661920	220	1.24–1.21	1.261–1.268	no
10:23:55.48	−57:45:31.43	2018 Jan 21	SHORT-1	1661920	220	1.35–1.25	1.269–1.276	no
10:24:17.45	−57:45:29.30	2018 Jan 25	SHORT-2	1661962	220	0.52–0.62	1.310–1.320	no
10:24:17.45	−57:45:29.30	2018 Jan 25	SHORT-2	1661962	220	0.51–0.52	1.300–1.309	no
10:24:17.45	−57:45:29.30	2018 Jan 25	SHORT-2	1661962	220	0.52–0.62	1.310–1.320	no
10:24:16.83	−57:46:27.62	2018 Jan 25	SHORT-2	1661962	220	0.71–0.60	1.332–1.342	no
10:24:16.83	−57:46:27.62	2018 Jan 25	SHORT-2	1661962	220	0.60–0.51	1.343–1.354	no
10:24:16.83	−57:46:27.62	2018 Jan 25	SHORT-2	1661962	220	0.62–0.60	1.320–1.330	no
10:24:18.08	−57:44:30.98	2018 Jan 25	SHORT-2	1661962	220	0.59–0.57	1.273–1.281	no
10:24:18.08	−57:44:30.98	2018 Jan 25	SHORT-2	1661962	220	0.57–0.49	1.281–1.290	no
10:24:18.08	−57:44:30.98	2018 Jan 25	SHORT-2	1661962	220	0.61–0.59	1.265–1.272	no
10:24:02.27	−57:46:17.40	2016 Jun 3	SHORT-MID_1	1323877	220	0.74–0.61	1.270–1.277	no
10:24:02.27	−57:46:17.40	2016 Jun 2	SHORT-MID_1	1323877	220	0.93–0.67	1.260–1.266	no
10:24:02.27	−57:46:17.40	2016 Jun 2	SHORT-MID_1	1323877	220	0.83–0.85	1.251–1.257	no
10:24:02.90	−57:45:19.12	2016 Jun 2	SHORT-MID_1	1323877	220	0.75–0.79	1.242–1.248	no
10:24:02.90	−57:45:19.12	2016 Jun 2	SHORT-MID_1	1323877	220	0.78–0.85	1.235–1.240	no
10:24:02.90	−57:45:19.12	2016 Jun 2	SHORT-MID_1	1323877	220	0.79–0.96	1.228–1.232	no
10:24:03.53	−57:44:20.80	2016 Jun 2	SHORT-MID_1	1323877	220	1.10–0.81	1.221–1.226	no
10:24:03.53	−57:44:20.80	2016 Jun 2	SHORT-MID_1	1323877	220	0.82–0.67	1.211–1.214	no
10:24:03.53	−57:44:20.80	2016 Jun 2	SHORT-MID_1	1323877	220	0.98–0.95	1.216–1.220	no
10:24:09.55	−57:46:22.51	2016 Jun 3	SHORT-MID_2	1323880	220	0.52–0.73	1.393–1.406	no
10:24:09.55	−57:46:22.51	2016 Jun 3	SHORT-MID_2	1323880	220	0.53–0.63	1.412–1.426	no
10:24:09.55	−57:46:22.51	2016 Jun 3	SHORT-MID_2	1323880	220	0.66–0.49	1.375–1.388	no
10:24:10.18	−57:45:24.19	2016 Jun 3	SHORT-MID_2	1323880	220	0.68–0.69	1.359–1.370	no
10:24:10.18	−57:45:24.19	2016 Jun 3	SHORT-MID_2	1323880	220	0.70–0.65	1.343–1.354	no
10:24:10.18	−57:45:24.19	2016 Jun 3	SHORT-MID_2	1323880	220	0.66–0.62	1.328–1.338	no
10:24:10.81	−57:44:25.91	2016 Jun 3	SHORT-MID_2	1323880	220	0.63–0.64	1.301–1.310	no
10:24:10.81	−57:44:25.91	2016 Jun 3	SHORT-MID_2	1323880	220	0.71–0.75	1.314–1.324	no
10:24:10.81	−57:44:25.91	2016 Jun 3	SHORT-MID_2	1323880	220	0.78–0.58	1.289–1.298	no

Note. Overview of the MUSE observations of Wd2 (Program ID: 097.C-0044(A), 099.C-0248(A)). In Columns 1 and 2 the coordinates are given. Column 3 shows the observation dates. Columns 4 and 5 indicate the name of each exposure and the observing block (OB) ID, respectively. Column 6 lists the exposure time. Column 7 indicates the seeing and Column 8 shows the airmass. Column 9 shows whether this pointing was obtained using WFM-AO. OBs that had to be repeated do not appear in this table. We grouped the three exposures that are combined into one pointing.

5. In the last step the sky subtraction is handled in `muse_scipost` by only applying the local LSF but not refitting the fluxes for each spaxel (`skymethod=subtract-model`).

To run this method we need to set the “sky: modified” keyword in the `config.json` file to “True.” It is also important that the subtraction method is “subtract-model” to avoid

refitting the spectral lines to each spaxel. It is possible to properly recover many of the nebular emission lines, such as [O I] $\lambda 5577$, $H\alpha$, or $H\beta$ (see Figure 2). For emission lines such as [O I] $\lambda 5577$, the telluric contamination by blends can be significant, especially for long exposure times, which means that condition 1 is no longer satisfied. Therefore, we recommend to use short exposure times to determine fluxes

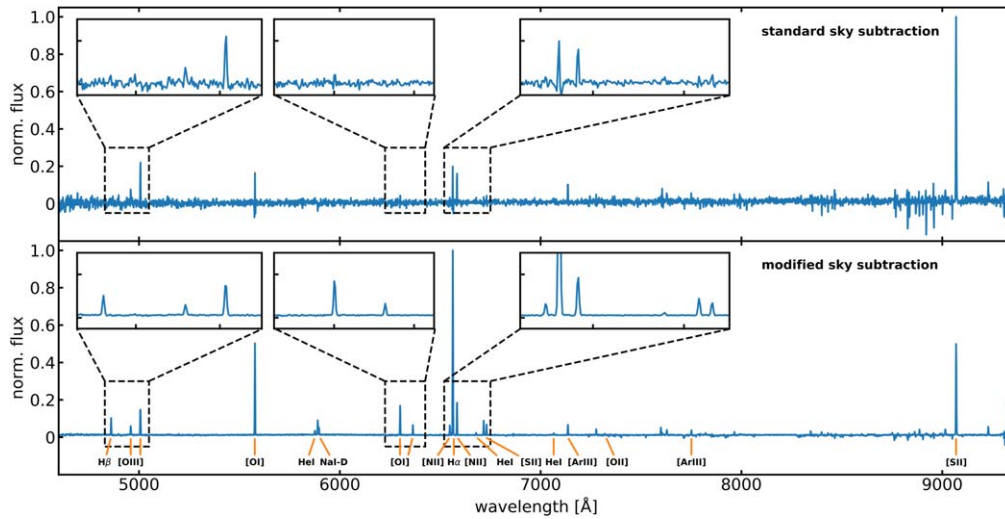


Figure 2. A sample spectrum extracted from a less luminous part of the H II region. Top: the spectrum as it is extracted using the standard method provided by the MUSE data reduction pipeline. Bottom: the spectrum as it is extracted with our modified version, as described in Section 5.1. Main nebular emission lines are marked. It becomes clear that many main nebular lines, such as [O I] λ 5577, H α , H β , or the sodium doublet, are properly recovered.

of the gas emission to minimize the residual telluric contamination in blends, and thus to properly measure the flux of nebular emission lines. Stellar spectra are not affected by this since a local, star-by-star background subtraction is performed (see Section 5.4).

Tests have shown that if a sky field is used instead of the science exposure to obtain the sky spectrum, e.g., J0291+0721 (096.B-0212(A), P.I.: B. James), ESO’s standard procedure for sky correction is recommended. The reason for this is that a refit of the fluxes of the sky emission lines to the science data may be necessary, caused by possible different observing conditions between the science and the sky exposures.

5.2. Source Detection and Extraction

To extract the stellar spectra from the reduced MUSE data cubes we used the Python package PampelMuse (Kamann et al. 2013). PampelMuse performs a position- and wavelength-dependent fit to the point-spread function (PSF) for each layer of the MUSE data cube. As input a high-resolution photometric stellar source catalog must be provided.¹⁷ This deep high-resolution catalog provides the necessary information to PampelMuse to de-blend the spectra of the majority of stars, especially in crowded regions, or to flag spectra accordingly where de-blending is no longer possible. These spectra were not used in the further analysis. We used the *HST* multi-band photometric star catalog of the Wd2 region obtained by Zeidler et al. (2015). When available, we used the *F814W* photometry as a brightness estimate, since this band roughly covers the wavelength range of the MUSE data set, otherwise (e.g., for saturated sources) we use one of the other wide-band filters. For some very bright sources all filters were saturated. These stars were assigned $F814W = 10 \text{ mag}$ ¹⁸ to provide a catalog as complete as possible. To create an input catalog in the necessary PampelMuse format, the MUSEpack module

`cubes.pampelmuse_cat` can be used. MUSE is a seeing-limited instrument, which means that each point source covers multiple spaxels on the detector and each spectrum is extracted from multiple spaxels, which mitigates wavelength calibration uncertainties.

5.3. Flux Offsets between Different Pointings

PampelMuse also calculates the stellar magnitudes by convolving the extracted stellar spectra with the throughput curve of the filter given in the source catalog. In the case of the Wd2 data the input magnitudes correspond to the Wide Field Camera (WFC) *F814W* filter of the Advanced Camera for Surveys (ACS, Ryon et al. 2019). The ACS *F814W* filter covers a wavelength range from 6900 to 9500 Å while the longest wavelength of the MUSE spectra is 9350 Å. Therefore, the ACS *F814W* covers about 150 Å more of the stellar SEDs than the MUSE data, and it is expected that the magnitudes extracted from the MUSE spectra are somewhat fainter than the ACS magnitudes. Additionally, the MUSE data set was not observed under photometric conditions, which also introduces some flux uncertainties.

To confirm the reliability of the MUSE data reduction we analyzed the stellar magnitudes and we find that the MUSE *F814W* magnitudes are up to $\sim 0.7 \text{ mag}$ brighter than the ACS *F814W* magnitudes. Between different observing blocks (OBs) these offsets vary between 0 and 0.77 mag, which corresponds to a flux difference of up to a factor of two (see Figure 3). We also measured the flux differences using sky emission to ensure that the offsets are not caused by stellar binary systems or blends. The variation between different pointings within one OB is typically small ($< 0.1 \text{ mag}$). These small variations within an OB can be explained by changes in the observing conditions (clouds, temperature, etc.) since the observations were not taken under photometric conditions. The large offsets between different OBs are most likely caused by the calibration of the spectrophotometric standard star¹⁹ including the time at which the standard star was observed in relation to the science

¹⁷ The catalog should be as complete as possible and go at least 2 mag deeper than the MUSE detection limit. Additionally, a much higher spatial resolution than MUSE is recommended to detect blends.

¹⁸ This magnitude estimate is sufficient for the source detection, and PampelMuse will properly determine the magnitudes based on the extracted spectra and the *HST* *F814W* throughput curve.

¹⁹ The stars used for this calibration are automatically taken from the reference list of standard stars: <https://www.eso.org/sci/observing/tools/standards/spectra.html>.

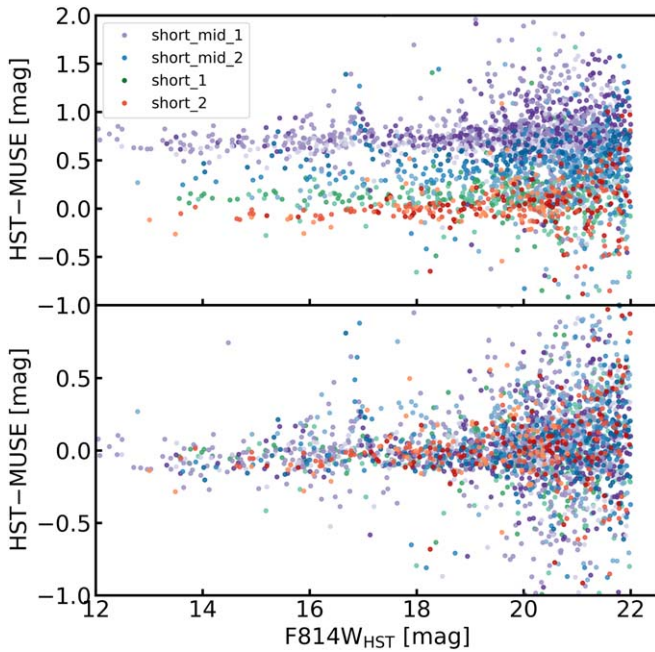


Figure 3. The offsets of the $F814W$ stellar magnitudes for the four short-exposure OBs (top panel). The different shades for each color represent one pointing within the OB. In the bottom panel we show the corrected magnitudes.

observations, e.g., OB SHORT-MID_1 was observed at the beginning of the night June 2/3 (22:48:04–00:10:06 UT), while the corresponding standard star was observed at the end of the same night (09:52:36–10:08:02 UT), partly into twilight (09:58:00–11:11:00 UT).

To be able to compare fluxes between the various OBs of our data set, we correct for these offsets in the following manner.

1. We determine the median difference between the HST and MUSE $F814W$ magnitudes using a 3σ clipping to account for stellar variability and photometric uncertainties.
2. We transform the magnitude difference to a flux-scaling factor.
3. We correct each spaxel in the data cube for both the data and the variance cube.
4. We rerun CUBEFIT and GETSPECTRA of PampelMuse to obtain the spectra with the corrected fluxes.

The fluxes in the data cubes may be corrected using the `cubes.wcs_cor` routine (see Section 5.6 for more details) with “correct_flux=True.” In the bottom panel of Figure 3 we show the corrected magnitudes, whose differences are now close to zero. All long exposures, as well as the SHORT-1 and SHORT-2, show smaller variations (0.04–0.14 mag). These OBs were obtained in run 099.C-0248(A), where the observing conditions were generally better than in run 097.C-0044(A), which was executed as a filler program.

One can say that the effective level of stellar crowding in the data is determined by the spectral type and magnitude of the stars and the depth of the data.

5.4. The Background Subtraction

The extracted stellar spectra still show contamination from the emission of the local H II region. Therefore, similarly to a photometric data reduction (e.g., Zeidler et al. 2015) a local, star-by-star background subtraction has to be performed. This is a more challenging task than photometry for IFU

Table 2
The Extracted Stellar Spectra

S/N	Short exposure	Long exposure	Unique
>50	187	193	259
20–50	283	300	423
10–20	192	359	427
5–10	263	529	616
≥ 5	925	1381	1725
<5	2656	2395	3318

Note. The number of stellar spectra we extracted from the short (Column 2) and long exposures (Column 3) for different S/N ranges. For completeness we also provide the number of stars with $S/N < 5$, which are not considered in the further analyses. Column 4 contains the total number of detected stars. If a star was detected in both the long and short exposures, we use the spectrum with the higher S/N.

observations. The PSF, the signal-to-noise ratio (S/N), and the stellar spectral energy distribution (SED) depend on wavelength. As a result, the background emission and the effective level of stellar crowding in the data, determined by the spectral type and magnitude of the stars, are a function of wavelength. PampelMuse estimates the background flux local to each star by dividing each spectral bin into equally sized squares and weighting the median flux in each square by the distance to the stars. The estimated background flux is subtracted from the stellar flux individually for each spectral layer. This method works well for regions with a low crowding or a background that is marginally changing within a few PSFs. Wd2 is a crowded cluster with a highly variable background (gas clouds) on small spatial scales. This leads to an over- or underestimation of the background, especially at wavelengths where stellar absorption features overlap with nebular emission lines (e.g., Balmer lines or helium lines). We are currently working on a method with which we will be able to clean any residual contamination from the spectra in a secondary step. This will be published in a future paper and we will add the method to MUSEpack. For the further analysis in this work we are using the stars for which we extracted spectra where the background has been well subtracted.

5.5. The Extracted Stars

To extract the stellar spectra we decided to use a lower S/N limit of 5. Throughout the paper the S/N of the spectra always refers to the mean S/N per pixel. In total, we extracted 924 stellar spectra from the short exposures and 1381 stellar spectra from the long exposures. Combining the two catalogs, we analyzed 1726 stars with $S/N \geq 5$. In Table 2 we provide an overview of the number of stars detected in the short and long exposures, as well as in the combined catalog. If a star was detected in both the long and short exposure, we use the spectrum with the higher S/N. For completeness we also provide the number of stars detected with $S/N < 5$, which are not considered in any further analyses. One has to keep in mind that the long exposures (five in total) only cover $\sim 45\%$ of the area covered by the short exposures (11 in total). Additionally, the long exposures do not cover the cluster core.

5.6. The World Coordinate System—Creating the Mosaic

While pointing and guiding of the VLT are excellent (also for exposure times of 1 hr), the absolute world coordinate system (WCS) is less accurate, leading to coordinate offsets between

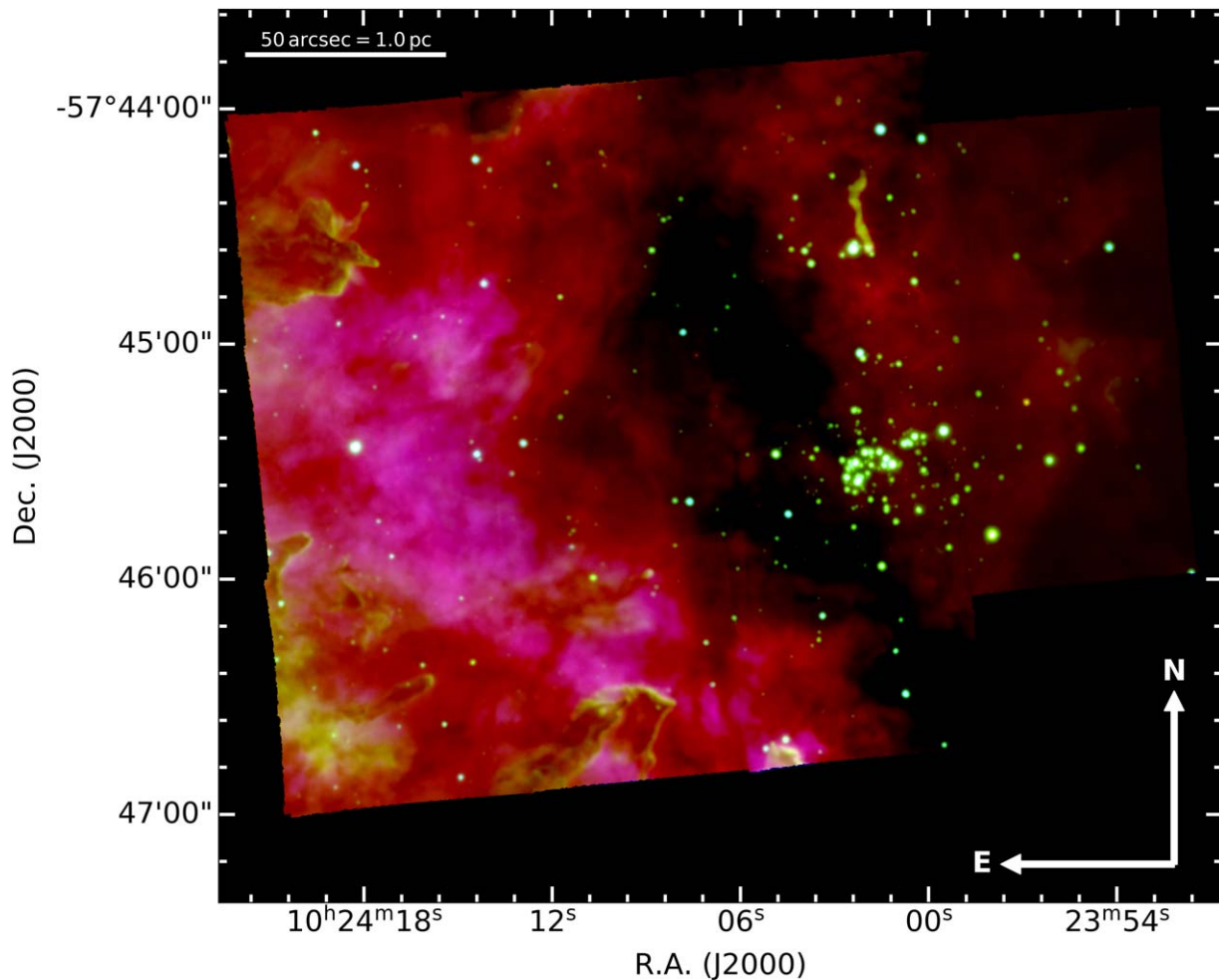


Figure 4. Mosaic of the MUSE data set of Wd2 (short exposures only). To create this image the $H\alpha$ (red), $[N II]\lambda 6583$ (green), and $[O III]\lambda 5007$ (blue) nebular lines were used.

individual exposures. This is corrected by the MUSE data reduction pipeline when combining individual dither positions but not between individual OBs. To properly create mosaics and compare the data to other observations (e.g., *HST*, *Spitzer*), a common WCS is needed. We decided to adopt the *Gaia* DR2 WCS. Therefore, as the first step we matched our photometric *HST* catalog (Zeidler et al. 2015) to the *Gaia* DR2 catalog (Gaia Collaboration et al. 2016, 2018) using the software CataPack.²⁰ CataPack uses a 2D histogram technique to match multiple catalogs and to determine the coordinate transformation including a polynomial of degree n (defined by the user) for the coordinate transformation. The latter is then subsequently used to correct the coordinates of the *HST* catalog and the corresponding headers of the Wd2 MUSE images. Although the accuracy of the *Gaia* DR2 catalog is still limited in crowded regions, the stellar density in the outskirts of Wd2 is low enough for an accurate determination of the WCS. In total we used 1239 stars with a median uncertainty for the positions of the stars of 0.345 mas in R.A. and 0.319 mas in decl.

PampelMuse determines the shift, rotation, and scale between the MUSE data cubes and the photometric input catalog. We use the PampelMuse solution to obtain a uniform transformation between the photometric stellar catalog and each individual data cube. This transformation is added as a new, primary WCS to

the fits headers of the data cubes. For these transformations we used the `cubes.wcs_cor` module from MUSEpack.

With this common WCS it is possible to properly mosaic layers of the individual cubes. The modules `cubes.line-maps` and `cubes.mosaics` provide the necessary processing. In Figure 4 we show the RGB mosaic representing the short exposures using the slices for the $H\alpha$ (red), $[N II]\lambda 6583$ (green), and $[O III]\lambda 5007$ (blue) spectral lines.

6. The RVs

Measuring accurate RVs in YMCs is challenging because the typical velocity dispersion in these clusters is of the order of 5 km s^{-1} . Traditionally, it has been necessary to obtain high-resolution stellar spectra with fiber, slit, or echelle spectrographs. The development of large IFUs allows us, for the first time, to map these resolved star clusters spectroscopically with a similar efficiency to photometric surveys. However, their typically low spectral resolution presents a challenge. Nevertheless, Kamann et al. (2013, 2016, 2018) showed for Galactic globular cluster (GC) stars that using MUSE spectra and a spectral library allows one to determine the RVs of stars to an accuracy of $1\text{--}2 \text{ km s}^{-1}$. This method allowed them to study the rotation profile of these old star clusters, but because no extensive optical pre-main-sequence spectral libraries exist, applying this to young clusters is challenging

²⁰ <http://www.bo.astro.it/%7Epaolo/Main/CataPack.html>

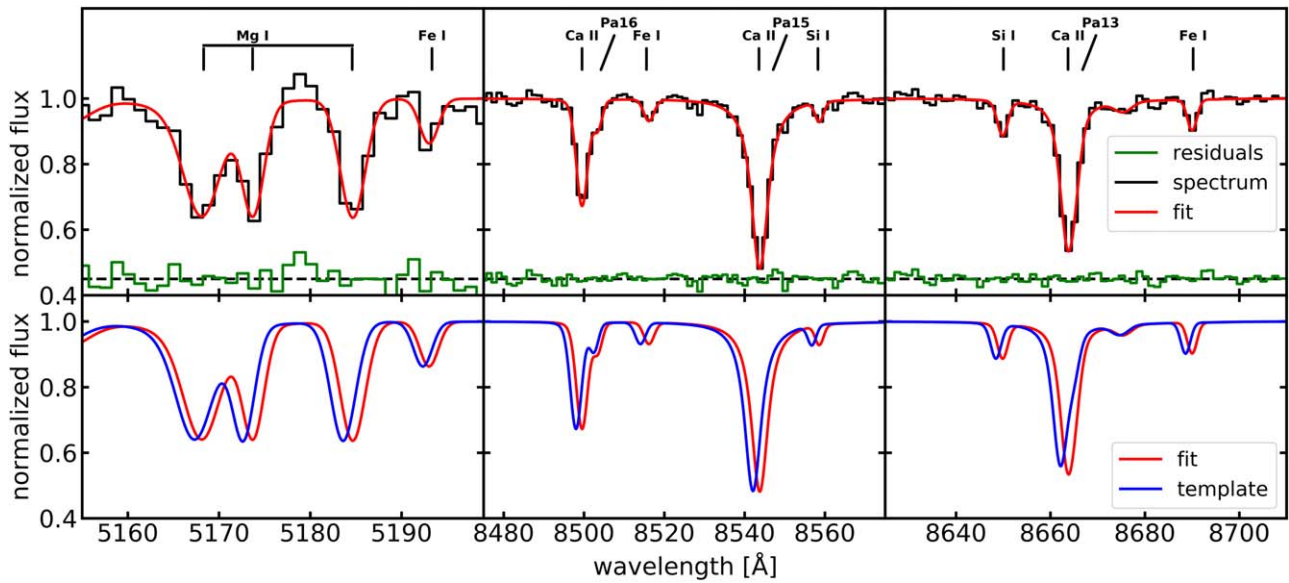


Figure 5. An example of fits of the Mg I $\lambda\lambda 5367, 5172, 5183$ and Ca II $\lambda\lambda 8498, 8542, 8662$ spectral lines. In the top row the extracted and rectified stellar spectrum is plotted (black) including the fitted spectral lines in red. The green line represents the fit residuals (offset by 0.45 from zero in normalized flux units). Additionally, important stellar absorption lines are marked. In the bottom row we show the fitted spectrum in red and the resulting template at rest-frame wavelengths in blue.

The novelty of our method is to create a template library based on the observed spectra, as well as a spectral line library. An introduction to this method was presented in Zeidler et al. (2018). It uses strong stellar spectral lines in combination with a Monte Carlo (MC) approach to measure the RVs. If not stated otherwise, from now on we use wavelengths in angstrom and the rest-frame for air, since this is the standard for the reduced MUSE data cubes, although MUSEpack does not depend on a specific rest-frame.

6.1. The Spectral Templates

To create stellar spectral templates we use the information provided by the Atomic Spectra Database of the National Institute of Standards and Technology (NIST²¹). The templates are created around strong stellar absorption lines, such as Mg I $\lambda\lambda 5367, 5172, 5183$, Na I $\lambda\lambda 5889, 5895$, and Ca II $\lambda\lambda 8498, 8542, 8662$ for stars with a cool photosphere, and He I and He II lines for stars with a hotter photosphere. We are not using the Balmer lines because the majority of stars are still in their pre-main-sequence phase and as a result these lines may be affected by accretion processes (e.g., Zeidler et al. 2016).

6.1.1. The Method

The templates are created using the `line_fitting` instance of the MUSEpack class `RV_spectrum`, which uses `pyspeckit` (Ginsburg & Mirocha 2011) as the core fitting routine. The spectral resolution of the MUSE data requires that `line_fitting` properly handles blended lines. For a proper velocity fit it is necessary that the line of interest (primary line) is stronger than the blend in order to properly fit the centroid. To create a spectral template the following steps must be executed.

- (1) The creation of a spectrum instance using the `RV_spectrum` class. This instance includes the wavelength, flux, and flux uncertainty array as well as all methods needed to run the RV fit.
- (2) The reading of a catalog (`RV_spectrum.catalog`) of lines that shall be fitted and used to measure RVs. This

catalog must contain: the unique line name, the rest-frame wavelength of the spectral line, the start and end points of the fitting range, and the initial order of the polynomial used to fit the continuum. The order of the polynomial may be increased subsequently while fitting the spectra.

- (3) The execution of the main fitting module (`line_fitting`), which attempts to fit the spectral line profiles (using a Voigt profile, see Section 5.1 of Zeidler et al. 2018) with a common continuum, taking into account wavelength limits, blends, and the instrument dispersion. The `line_fitting` module is capable of running on multiple cores (one core per primary line). After the fit has converged, a fitted spectrum as well as the template are created and added to the instance. To create the template, `line_fitting` uses the continuum and the fitted line parameters (A, σ, γ) but the rest-frame wavelengths from the spectral line catalog.

In Figure 5 we show an example of the fitted template for the Mg I $\lambda\lambda 5367, 5172, 5183$ and Ca II $\lambda\lambda 8498, 8542, 8662$ spectral lines, similar to what can be plotted by executing `RV_spectrum.plot`. In black we show the stellar spectrum and in red the fitted spectral lines (top row). In blue we show the created template (bottom row), which is shifted to the rest-frame wavelengths. The green line represents the fit residuals. The small offset of amplitude between the template and the fit is a result of the wavelength-dependent shift, which may change the amplitude of the superposition of the primary line and the blend. This does not influence the location of the centroid of the primary line. We allow for small amplitude adjustments while fitting the RVs (see Section 6) to obtain a more stable cross-correlation.

6.1.2. Important Parameters

The `line_fitting` module accepts multiple input parameters. In this section we introduce the mandatory and most important ones, but for a detailed description of all available parameters we refer to <https://musepack.readthedocs.io/en/latest/>.

The two mandatory arguments are a spectral line catalog (format: unique line name, wavelength) and an array of names of the primary lines that shall be fitted.

²¹ <https://www.nist.gov/pml/atomic-spectra-database>

The `line_fitting` module also takes multiple important (but not mandatory) keyword arguments. A “list with blends” (format: prime name, prime wavelength, blend name, blend wavelength, maximum line ratio) may be provided. The maximum line ratio is the maximum ratio of the primary to the blended line. Setting “autoadjust” to “True” allows the wavelength range “limits,” determining the minimum and maximum allowed central wavelengths of the fitted lines, to be adjusted during execution depending on the last fitted centroid of the primary line. For all other lines the wavelength shift ($\Delta\lambda/\lambda$) is taken into account accordingly. If “fwhm_block” is set to “True,” it limits the fitted line width such that it is never smaller than the dispersion of the instrument to prevent unreasonable and impossible narrow lines. “adjust_preference” sets the order in which “niter” and “contorder” are adjusted in the cases where either the fit does not converge after “niter” iterations or it diverges. If set to “contorder,” the polynomial order used to fit the continuum is increased in steps of one (to a maximum of “max_contorder”). If the maximum continuum order is reached the wavelength range of the fit (as set in the input catalog) is changed by ± 5 Å until the maximum number of adjustments (“max_ladjust”) is reached. If “adjust_preference” is set to “wavelength” the order of adjustments is reversed. If both parameters reach their set maxima without convergence a warning is issued.

6.2. Measuring RVs

After the steps in Section 6.1 have been executed the spectral instance also contains the stellar spectrum to be fitted, the template at rest-frame wavelength, as well as a continuum function. The RV fit is initiated by executing `RV_spectrum.rv_fit`. To measure the RVs via cross-correlating the spectrum and the template we use pPXF (Cappellari & Emsellem 2004; Cappellari 2017). The `RV_spectrum.rv_fit` instance requires as input the initial guesses in the form [RV, dispersion], similar to pPXF. The dispersion is the difference between the template and the spectrum, which in our case is by definition zero, since the template is created with the fit parameters for each line. We recommend to execute `RV_spectrum.rv_fit_peak` before `RV_spectrum.rv_fit`, which provides an RV measurement based on the line centroids only, which has proven to be a good initial guess for the RV fit. The RV fit instance executes the following scheme to determine the stellar RVs.

1. The RV of each primary line is measured. Therefore, pPXF is executed *niter*/2 times (default *niter* = 10,000) and for each iteration the uncertainties of the spectrum are randomly reordered. To determine the RV of each line a normal distribution is fit to the RV distribution, while its fitted mean value is considered to be the RV of this line.
2. If there are three or more absorption lines the median between the RVs and the median absolute deviation (MAD) is determined in order to perform a 3σ clip. This ensures that lines with “odd” profiles are removed from the final RV fit. Additionally, only lines with a certain significance are used. This significance is defined by how many standard deviations the spectral line has to be above the local uncertainty (Poisson noise) of the spectrum (amplitude $> \text{significance} \times \sigma$).²² We determined that a significance of 5 is a good value but the user is able to set it as a parameter in the RV fit module.

3. The final step determines the RV by repeating step 1, using all lines together that were classified as good in the previous step. This time pPXF is executed *niter* times. At the end a Gaussian is fit to the resulting RV distribution of the *niter* MC fits. The RV of the star is the mean of this distribution, while the 1σ standard deviation is defined as the RV uncertainty.

Each of the above pPXF runs is independent of the others so these steps can be executed using multi-core processing. The final results of the RV fit are also stored in the spectral object and properly logged.

6.3. The Debug Mode

The `RV_spectrum` class may be executed in debug mode. By activating the debug mode, all tasks are executed on a single core. The spectral line fit is executed interactively, where each iteration is plotted, all fit parameters are printed to the terminal, and the user must confirm before continuing to the next iteration. This mode is useful for following the individual steps of spectral line fitting (e.g., if the fit keeps diverging) or for tracing back error messages, since this can be tricky when running Python on multiple cores.

6.4. The Accuracy and Reliability of the RV Measurements

To show that our new method of fitting RVs is reliable and it is indeed possible to determine RVs to an accuracy of $1\text{--}2\text{ km s}^{-1}$, we performed multiple tests and sanity checks. We checked many of the spectral fits by eye to determine the goodness-of-fit and whether weak lines or lines where the spectral fit did not optimally converge were properly discarded.

In Section 5.1 of Zeidler et al. (2018) we already showed that the RVs determined with MUSEpack agree well with the existing literature values for eight stars ($\Delta\text{RV} = (3.7 \pm 4.08)\text{ km s}^{-1}$). Furthermore, all of the long exposures spatially fully overlap with the short exposures. There is a magnitude range in which stars are faint enough not to saturate in the long exposures but are still luminous enough that the extracted spectra from the short exposures have a high enough S/N to properly measure the stellar RV. 137 stars fulfill this criterion and the median RV offset between the long and short exposures is $(-0.27 \pm 4.64)\text{ km s}^{-1}$ (see left panel of Figure 6). Next, we used stars that are detected in the overlap regions of different data cubes. In total 17 stars are observed in multiple cubes, and even though all of them are at the edges of the IFU where the wavelength calibration is less reliable, they agree very well with each other (median offset: $(0.32 \pm 3.04)\text{ km s}^{-1}$, see middle panel of Figure 6). Last but not least, we checked the correlation between the RV determined by the centroid only and the full RV fit. Even though the uncertainties of the RVs determined with the line centroids are large their median offset is $(1.02 \pm 1.90)\text{ km s}^{-1}$ (see right panel of Figure 6). We tested whether there is any correlation between the RVs and the absorption line species used to derive them. In order to do so, we used the same set of stars and determined the RVs using different sets of lines (Ca II, Mg I, and for the short exposures Na II), and within the uncertainties the RVs are in agreement.

For some stars the deviation between the RV measured with the line centroids only and the full RV fit is quite large. After checking the spectral and RV fits we determined that the MC results jumped between two distinct values. Therefore, we decided to use this as an additional quality criterion to create the

²² We assume that the local S/N of the local spectrum only changes slowly with wavelength.

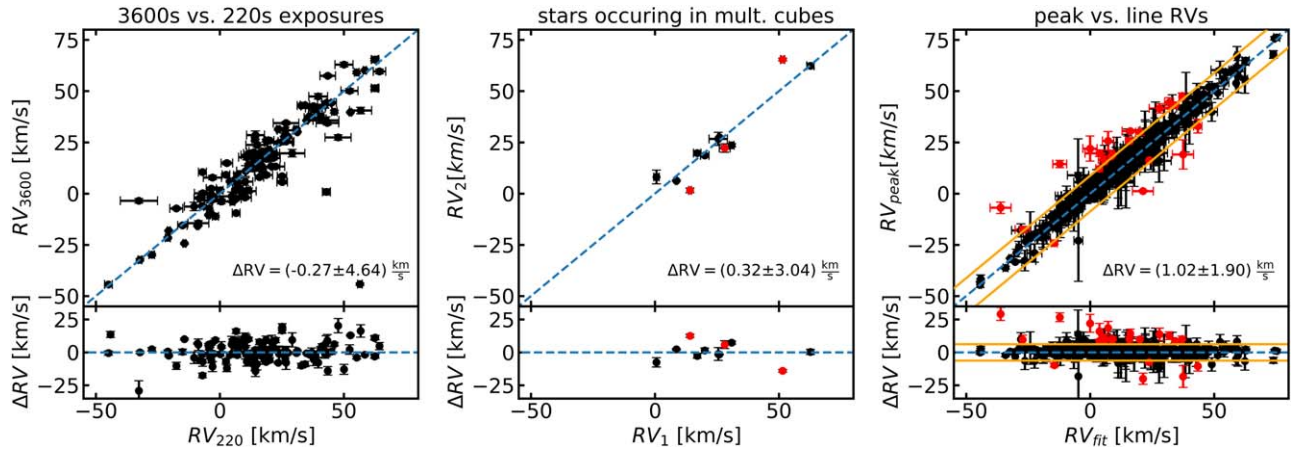


Figure 6. The accuracy of the RV measurements. Left panel: we show all stars that are detected in both the 220 s and the 3600 s exposures ($\Delta RV = (-0.27 \pm 4.64) \text{ km s}^{-1}$). Center panel: we show all stars that are detected in multiple cubes, where the tiles of the mosaic overlap ($\Delta RV = (0.32 \pm 3.04) \text{ km s}^{-1}$). Right panel: we show the peak vs. the fitted RV measurements ($\Delta RV = (1.02 \pm 1.90) \text{ km s}^{-1}$). The orange lines represent the 3σ difference between peak RV and fit RV used as an additional selection criterion. All stars marked in red do not fulfill this criterion.

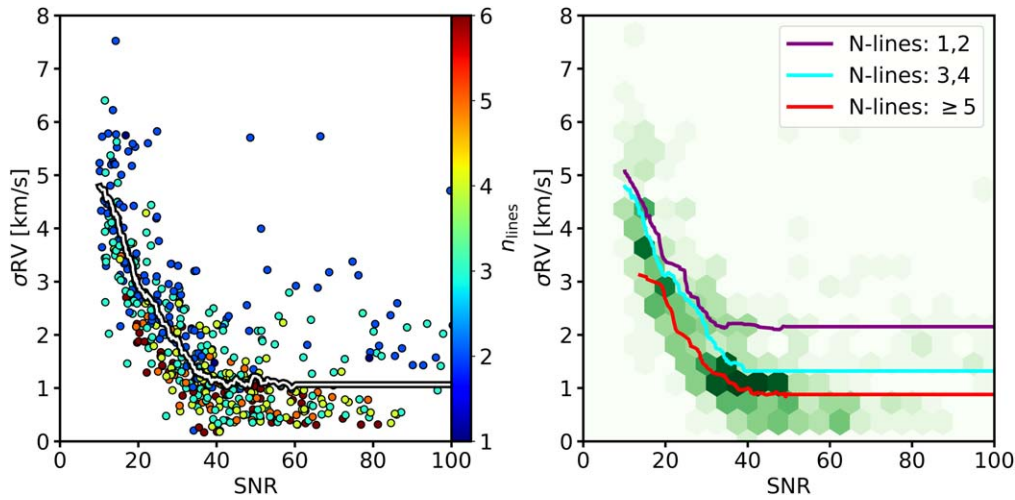


Figure 7. The distribution of the RV uncertainty vs. the S/N of the stellar spectra. Left: each individual star is color-coded by the number of lines used to perform the final velocity fit. The more lines are used and the higher the S/N of the spectra is, the smaller is the RV uncertainty. The black–white line represents the running mean of the RV uncertainty. Right: a density plot reflecting the $\sigma(RV)$ –S/N plane overplotted with the running mean of the RV uncertainties depending on the number of lines used for the final RV fit: 1 or 2 lines (purple), 3 or 4 lines (cyan), or ≥ 5 lines (red).

RV catalog. We flagged all stars whose difference between peak RV and fit RV (including error bars) is larger than 3σ (orange lines in the right frame of Figure 6). 30 out of 416 stars, or 7.2%, are affected (red stars in the right frame of Figure 6). This additional criterion can only be applied if a statistically significant number of stars is available, but for a small number of stars it is feasible (and recommended) to manually check the results.

The difference between peak RV and fit RV may be a result of binaries, which we are currently unable to resolve and detect. Due to the limited spectral resolution of MUSE, a minimum projected orbital RV of $\sim 50 \text{ km s}^{-1}$ is necessary to resolve double-peak binaries. Our inspections of the flagged spectra have not shown any hints for such sources. In a future work we plan to analyze possible binaries in those regions where sources are detected in multiple data cubes and incorporate an automatic method to flag those potential binaries. Currently, we suggest to manually check those flagged spectra.

Besides the measured uncertainties for each individual source, the above described tests lead to a combined statistical

uncertainty of

$$\sigma = \sqrt{\sigma_{1,s}^2 + \sigma_{\text{mult}}^2 + \sigma_{\text{peak}}^2} = 1.10 \text{ km s}^{-1}. \quad (1)$$

This shows that measuring RVs to a level of accuracy of $1\text{--}2 \text{ km s}^{-1}$ without the existence of spectral libraries is indeed feasible with this method.

We analyzed the correlation between the RV uncertainties, the number of lines used to measure the RVs, and the S/N of the input spectra. In the left panel of Figure 7 we show the correlation of the RV uncertainties versus S/N of the spectra for each individual star, color-coded with the final (cleaned) number of lines (see Section 6.2) used to measure the final RVs. The black and white line is the running mean of the correlation. In the right panel of Figure 7 we present a smoothed histogram of the $\sigma(RV)$ –S/N relation overplotted with the running mean if 1 or 2 (purple), 3 or 4 (cyan), or ≥ 5 (red) stellar absorption lines were used to determine the RV. In Table 3 we list the typical RV uncertainties for a specific S/N and a specific number of absorption lines used.

Table 3
The RV Uncertainty

<i>n</i> lines	S/N				
	10	20	30	40	≥ 50
All	4.78	2.92	1.93	1.16	1.06
1 and 2	5.07	3.40	2.40	2.19	2.15
3 and 4	4.78	3.10	2.10	1.32	1.30
≥ 5	...	2.65	1.41	1.00	0.88

Note. The dependence of the RV uncertainty on the S/N of the spectra used per number of lines used. All RV uncertainties are given in km s^{-1} . For ≥ 5 lines not enough data were available to obtain a statistically significant RV uncertainty value for a S/N of 10.

Overall, the RV accuracy increases with increasing S/N and increasing number of used lines, as expected. The accuracy appears to converge for all S/N larger than ≈ 50 to typical RV accuracies of 0.9 km s^{-1} , 1.3 km s^{-1} , and 2.2 km s^{-1} with ≥ 5 , 3–4, and 1–2 spectral lines, respectively (see Table 3). For $\text{S/N} < 10$ the typical RV uncertainty exceeds 5 km s^{-1} , and therefore we suggest a lower S/N limit of 10 for RV measurements.

7. The Wd2 MUSE RV Catalog

As described in the previous section, we measured the RVs for both the long and the short exposures for all stars with $\text{S/N} > 10$ for which we extracted clean spectra (see Section 5.4). We merged the catalogs from the long and the short exposures, and in the cases where there are multiple RV measurements for the same source, we will use the RV measurement for which the S/N of the spectrum was the highest. We keep the individual measurements, which we will use in a future work to look for RV shifts caused by binary systems, and we will cross-correlate them with the *HST* multi-epoch eclipsing binary study by E. Sabbi et al. (2019, in preparation). The current final catalog²³ comprises 388 individual stars, of which 117 are cluster members and 271 are field stars, based on the photometric selection using various color–magnitude diagrams (CMDs) (Zeidler et al. 2015).

In Figure 8 we show the $F814W$ versus $F814W - F160W$ CMD based on the *HST* photometry (gray points). We marked all stars that are detected in the MUSE data set color-coded with their S/N. The left panel of Figure 8 shows the stars extracted from the short exposures while the right panel shows the stars extracted from the long exposures. Some of the most luminous stars are not detected in the long exposures, either because the MUSE data saturated or because they were outside the area covered by the long exposures (especially true for the cluster center). $\text{S/N} > 5$, which we consider as a detection threshold, is reached for $F814W = 19.39 \text{ mag}$ in the short exposures and $F814W = 20.49 \text{ mag}$ in the long exposures, corresponding to stellar masses of $1.9 M_{\odot}$ and $1.0 M_{\odot}$, respectively. $\text{S/N} > 10$, which we have demonstrated (see Section 6.4) is a lower limit for the reliable measurement of stellar RVs, is reached for $F814W = 18.62 \text{ mag}$ in short exposures and $F814W = 20.02 \text{ mag}$ in the long ones or stellar masses of $2.4 M_{\odot}$ and $1.9 M_{\odot}$, respectively. All stellar masses are based on a 1 Myr isochrone at a distance of 4.16 kpc with a mean visual extinction of

$A_V = 6.12 \text{ mag}$ (Zeidler et al. 2015). A more detailed overview of the brightness and mass limits is given in Table 4.

8. Summary and Conclusions

In this work we introduced the VLT/MUSE IFU data set of Wd2, a Galactic YMC. This paper is intended to showcase “MUSEpack,” our newly developed Python package, whose main purpose is to measure RVs of stars extracted from IFU data sets in the absence of stellar spectral template libraries.

We obtained 16 data cubes of MUSE data covering 11 arcmin^2 with two different exposure times to study the gas, the luminous upper main-sequence stars, as well as the fainter pre-main-sequence stars down to $F814W = 20.49 \text{ mag}$ ($1.0 M_{\odot}$, see Table 4). The long exposures covering the pre-main-sequence stars are located around the cluster center (see Figure 1) to avoid saturation. Using PampelMuse to extract the stellar spectra, we extracted 1602 stars with $\text{S/N} > 5$ (see Table 2). The internal kinematics of Wd2 will be presented in a future work.

To reduce the data we use the MUSEpack class `MUSEr-educe`, which is a Python wrapper for ESO’s data reduction pipeline. It is fully written in Python 3 and therefore platform-independent, and also it does not require a GUI, making it convenient to run remotely even with slower internet connections or on a computer cluster. As part of `MUSEr-educe` we developed a modified sky subtraction to perform proper sky subtraction without a sky field in crowded regions (e.g., resolved, nearby H II regions or nearby galaxies). The modifications avoid the oversubtraction of spectral emission lines that are part of both the science target and the Earth’s atmosphere, visible as telluric lines (see Section 5.4 and Figure 2). We tested `MUSEr-educe` with the WFM-NOAO and WFM-AO data of Wd2, with WFM-NOAO data of the blue compact dwarf galaxy J0291+0721, a faint extended object (James et al. 2019), as well as NFM-AO data of a T-Tauri disk (J. Girard et al. 2019, in preparation), to ensure that it properly supports all currently provided science modes.

In the Appendix we showed that it is possible to subtract contamination from the Moon caused by executing the observations under “bright” conditions. By fitting a lunar model to the data it is possible to correct the data cubes to an extent at which the lunar contamination is only minimal. Although there is still some flux contamination left, which has an influence on the absolute flux measurements, stellar RVs can be measured with the same accuracy as under “dark” conditions.

We showed that it is possible to measure stellar RVs with an accuracy of $1\text{--}2 \text{ km s}^{-1}$ based on MUSE data without the need for a spectral template library using the MUSEpack class `RV_spectrum` (see Section 6). This new method uses strong stellar absorption lines and fits atomic line libraries to small wavelength ranges around the stellar absorption lines. The fit is performed iteratively by automatically adopting updated input parameters based on the previous iteration within certain user-based limits. After convergence is reached a template spectrum is created based on the fitted line parameters and the rest-frame wavelengths of the spectral lines. In a subsequent step, the template is repeatedly cross-correlated against the input spectrum using an MC approach to determine the stellar RV and its uncertainty. Due to this MC approach RV accuracies of $\sim 2 \text{ km s}^{-1}$ can be reached for the majority of stars, despite the limited spectral resolution of MUSE. The statistical uncertainty is 1.10 km s^{-1} . Although S/N-dependent, the RV accuracy converges for all S/N larger than ≈ 50 , with typical RV

²³ This is an on-going project and the catalog is subject to future updates, especially after attempting to improve the local stellar background subtraction.

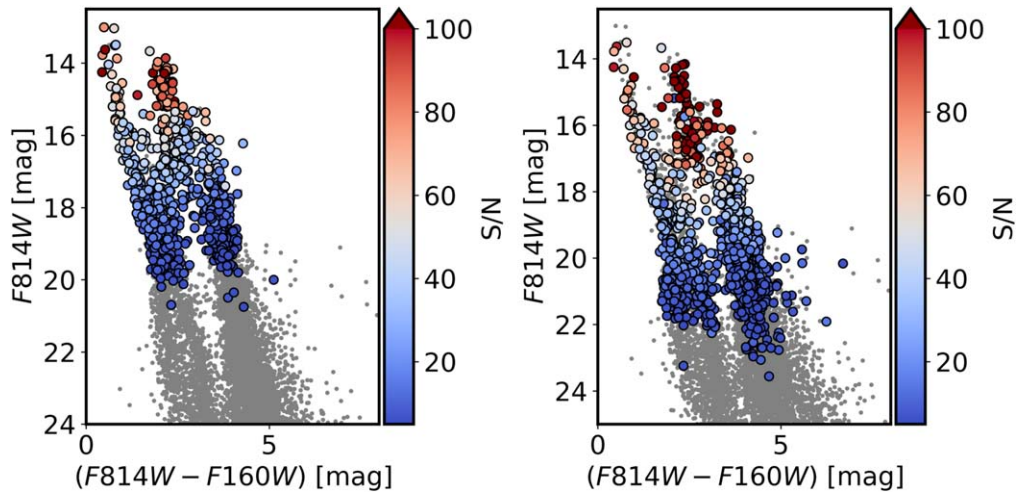


Figure 8. The sources detected with MUSE overplotted on the *HST* $F814W$ vs. $F814W - F160W$ CMD (gray dots) for the short (left panel) and long (right panel) exposures. The color represents the S/N of the spectra with a minimum S/N of 5.

Table 4
The Magnitude Limits

S/N	Short		Long	
	$F814W$ (mag)	Mass (M_{\odot})	$F814W$ (mag)	Mass (M_{\odot})
100	14.55	12.7	16.12	10.9
80	15.15	9.7	17.61	8.5
40	16.56	3.6	18.40	4.4
20	17.67	3.1	19.33	3.0
10	18.62	2.4	20.02	1.9
5	19.38	1.9	20.49	1.0

Note. The typical magnitude limits are for a given S/N of the stellar spectra for both the short and long exposures. The stellar masses are based on a 1 Myr isochrone at a distance of 4.16 kpc with a mean visual extinction of $A_V = 6.12$ mag (Zeidler et al. 2015).

accuracies of 0.9 km s^{-1} , 1.1 km s^{-1} , and 2.2 km s^{-1} with ≥ 5 , 3–4, and 1–2 spectral lines, respectively. For $S/N < 10$ the typical RV uncertainty exceeds 5 km s^{-1} , and therefore we suggest a lower S/N limit of 10 for RV measurements.

With the presented method it is now possible to efficiently measure stellar RVs in YMCs to determine their RV profile and dispersion, a key property in understanding their evolution in the first few million years. We will present the results of the RV study for Wd2 in a future paper.

MUSEpack is made available for download on Github <https://github.com/pzeidler89/MUSEpack> under a GPL version 3 license and Version 1 of this software is archived on Zenodo (Zeidler 2019). A detailed manual of the main Classes and all of the additional side-modules can be found on <https://musepack.readthedocs.io/en/latest/>. More modules will be subsequently added in the future, such as an enhanced local background subtraction for stars in crowded fields.

We thank the anonymous referee for their helpful comments and suggestions to improve the quality of the paper. We thank B. James for providing the data set J0291+0721 (096.B-0212(A)) and her patience in being the first user of the MUSEpack and especially MUSEreduce, which highly improved the user-friendliness of the package and tracing errors in the code. We thank J. Girard for

providing the NFM-AO data set (60.A-9482(A), P.I.: J. Girard), which allowed us to also test this mode. We thank N. Lützgendorf for helping us developing the Monte Carlo module MUSEpack.ppxf_MC. We thank L. Coccatto for his patience and help in solving the issues with the lunar contamination.

P.Z. acknowledges support by the Forschungsstipendium (ZE 1159/1-1) of the German Research Foundation, particularly via the project 398719443.

E.K.G. and A.P. acknowledge support by Sonderforschungsbereich 881 (SFB 881, “The Milky Way System”) of the German Research Foundation, particularly via subproject B05.

These observations are associated with program #14807. Support for program #14807 was provided by NASA through a grant from the Space Telescope Science Institute. This work is based on observations obtained with the NASA/ESA *Hubble Space Telescope*, at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555.

Based on observations collected at the European Southern Observatory under ESO programmes 096.B-0212(A), 097.C-0044(A), 099.C-0248(A), 60.A-9482(A).

This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

Software: PampleMuse (Kamann et al. 2013), ESORex (Freudling et al. 2013), pyspeckit (Ginsburg & Mirocha 2011), MUSE pipeline (v.2.2.0 Weilbacher et al. 2012, 2014), Astropy (The Astropy Collaboration et al. 2018), Matplotlib (Hunter 2007), pPXF (Cappellari & Emsellem 2004; Cappellari 2017), CataPack (Paolo Montegriffo).

Facilities: VLT:Yepun (MUSE), *HST*(WFC3, ACS), *Gaia* (Space).

Appendix Observing under “Bright” Conditions

During the quality checks we discovered a moiré-like pattern in the southernmost long exposure (LONG_1). This pointing was observed during two different nights (see Table 1).

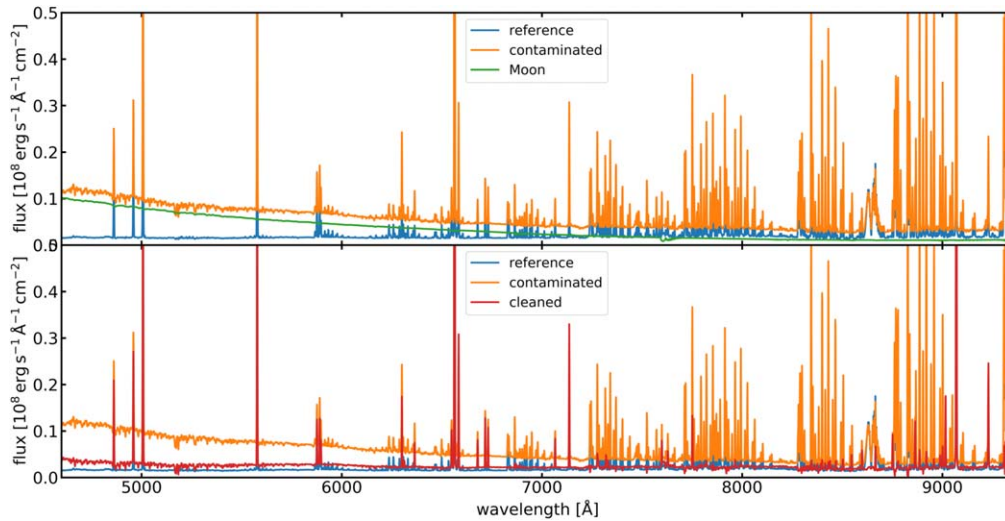


Figure 9. We show the removal of the lunar contamination from cube LONG_1a (LONG_1c is analogous). Top: the spatially collapsed spectrum of the contaminated data cube (orange) and the clean cube (LONG_1b) used as reference (blue). The stars were masked in both cubes to avoid a bias toward absorption lines. The green spectrum is the best fitting lunar spectrum. Bottom: the spatially collapsed spectrum of the contaminated data cube (orange) and the clean cube (LONG_1b) used as reference (blue). The stars were masked in both cubes to avoid a bias toward absorption lines. In red we show the spectrum of cube LONG_1a after applying the correction using the ESO data reduction pipeline.

LONG_1a and LONG_1c were obtained with Moon condition “bright” but exposure LONG_1b was observed with Moon condition “dark.” A detailed analysis of the individual cubes and consultation with ESO showed that, even with a Moon distance of 67.83° , observing during “bright” conditions causes lunar contamination (e.g., Patat 2004). This contamination is an additive component in the spectrum, which is not properly handled by the MUSE data reduction pipeline.

A.1. Removing the Lunar Spectrum

To correct for the lunar contamination, we attempted to subtract the lunar component from the individual data cubes LONG_1a and LONG_1c. The lunar spectrum observed with ground telescopes is in principle a solar spectrum reflected by the Moon and transmitted through the Earth’s atmosphere. The transmission efficiency of the Earth’s atmosphere depends on wavelength, the observing conditions (temperature, wind, and observation site), the position of the Moon on the sky, its phase (which influences the albedo), as well as the Moon angle (the angular separation between the Moon and the observed target) (for more details we refer to Patat 2004). This means that it is very difficult to model the lunar component for a specific observation (e.g., Krisciunas & Schaefer 1991). To estimate the sky contribution ESO provides an online tool “SkyCalc” (Sky Model Calculator²⁴), with which it is possible to estimate the flux contribution of the sky for any specific observing conditions. We used the lunar model for the observing conditions of LONG_1a and LONG_1c and fitted this model together with an extracted, high-S/N sky spectrum of the LONG_1b cube. We used pPXF, which also allows for small RV corrections between individual templates, to find the best matching linear combination of these two spectra. To better match the continuum we added a fourth-order Legendre polynomial series. We then combined the weighted lunar spectrum and the Legendre polynomial and added it as a

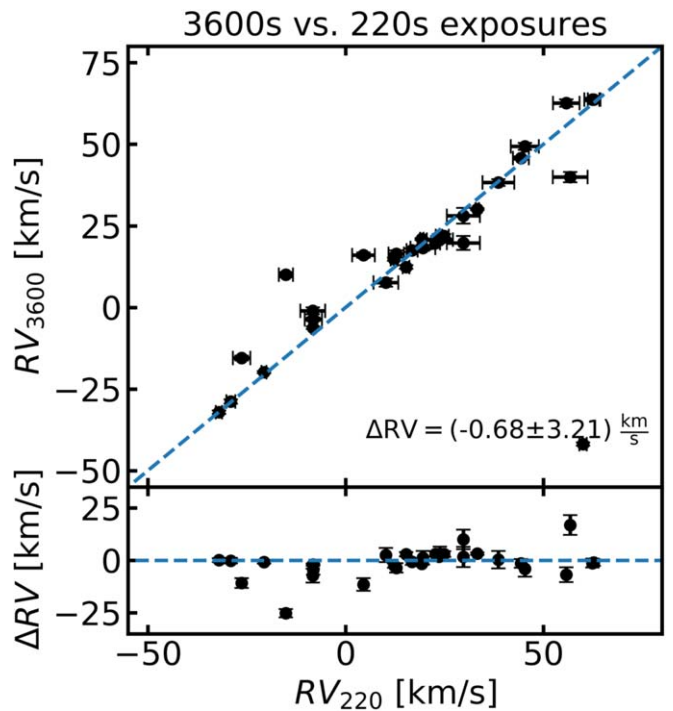


Figure 10. Comparison of stellar RV between the long (LONG_1) and short (SHORT-MID_1) exposures to determine the reliability of the data cube after the removal of the lunar spectrum.

continuum spectrum to the MUSE data reduction process. This ensures that the continuum spectrum gets properly convolved with the local LSF.

A.2. Quality Checks of the Correction

In Figure 9 we show the spatially collapsed spectrum of the contaminated data cube (orange) and the clean cube (LONG_1b) used as reference (blue) as well as the lunar spectrum (green spectrum, top panel) and the spectrum after applying the correction

²⁴ <https://www.eso.org/observing/etc/bin/gen/form?INS.MODE=swspectr+INS.NAME=SKYCALC>

using the ESO data reduction pipeline (red spectrum, bottom panel). Figure 9 shows the spectra for data cube LONG_1a. The correction for LONG_1c is analogous. Although this procedure does not fully remove the lunar component (see bottom panel of Figure 9) it significantly improves the reduced data, allowing the MUSE data reduction pipeline to properly combine the three dither positions.

We compared the stellar RVs of the long (LONG_1) and short (SHORT-MID_1) exposures (for a detailed description of how the RVs are determined, we refer to Section 6) to determine whether the lunar residuals influence the RV measurements. In Figure 10 we show the RVs of the stars detected in both data cubes. The mean offset between the two exposure times is $\Delta RV = (-0.68 \pm 3.21) \text{ km s}^{-1}$, which agrees very well with the full sample ($\Delta RV = (0.55 \pm 3.04) \text{ km s}^{-1}$, see left panel of Figure 6).

ORCID iDs

Peter Zeidler  <https://orcid.org/0000-0002-6091-7924>
 Elena Sabbi  <https://orcid.org/0000-0003-2954-7643>
 Anna F. McLeod  <https://orcid.org/0000-0002-5456-523X>
 Anna Pasquali  <https://orcid.org/0000-0001-5171-5629>
 Monica Tosi  <https://orcid.org/0000-0002-0986-4759>

References

- Bacon, R., Accardo, M., Adjali, L., et al. 2010, *Proc. SPIE*, **7735**, 773508
 Banerjee, S., & Kroupa, P. 2015, *MNRAS*, **447**, 728
 Bastian, N., & Goodwin, S. P. 2006, *MNRAS*, **369**, L9
 Cappellari, M. 2017, *MNRAS*, **466**, 798
 Cappellari, M., & Emsellem, E. 2004, *PASP*, **116**, 138
 Crowther, P. A., Caballero-Nieves, S. M., Bostroem, K. A., et al. 2016, *MNRAS*, **458**, 624
 Dale, J. E., Ercolano, B., & Bonnell, I. A. 2015, *MNRAS*, **451**, 987
 Freudling, W., Romaniello, M., Bramich, D. M., et al. 2013, *A&A*, **559**, A96
 Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *A&A*, **616**, A1
 Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, *A&A*, **595**, A1
 Ginsburg, A., & Mirocha, J. 2011, PySpecKit: Python Spectroscopic Toolkit, Astrophysics Source Code Library, ascl:1109.001
 Hunter, D. A., Shaya, E. J., Holtzman, J. A., et al. 1995, *ApJ*, **448**, 179
 Hunter, J. D. 2007, *CSE*, **9**, 90
 Hur, H., Park, B.-G., Sung, H., et al. 2015, *MNRAS*, **446**, 3797
 James, B., Kumari, N., & Emerick, A. 2019, *MNRAS*, submitted
 Kamann, S., Husser, T.-O., Brinchmann, J., et al. 2016, *A&A*, **588**, A149
 Kamann, S., Husser, T.-O., Dreizler, S., et al. 2018, *MNRAS*, **473**, 5591
 Kamann, S., Wisotzki, L., & Roth, M. M. 2013, *A&A*, **549**, A71
 Kiminki, M. M., & Smith, N. 2018, *MNRAS*, **477**, 2068
 Krisciunas, K., & Schaefer, B. E. 1991, *PASP*, **103**, 1033
 Lada, C. J., & Lada, E. A. 2003, *ARA&A*, **41**, 57
 Lada, C. J., Margulis, M., & Dearborn, D. 1984, *ApJ*, **285**, 141
 Larson, R. B. 1981, *MNRAS*, **194**, 809
 Longmore, S. N., Kruijssen, J. M. D., Bastian, N., et al. 2014, in *Protostars and Planets VI*, ed. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Klessen (Tucson, AZ: Univ. Arizona Press), 291
 McLeod, A. F., Dale, J. E., Evans, C. J., et al. 2019, *MNRAS*, **486**, 5263
 McLeod, A. F., Dale, J. E., Ginsburg, A., et al. 2015, *MNRAS*, **450**, 1057
 Pang, X., Grebel, E. K., Allison, R. J., et al. 2013, *ApJ*, **764**, 73
 Pang, X., Pasquali, A., & Grebel, E. K. 2011, *AJ*, **142**, 132
 Patat, F. 2004, *Msngr*, **118**, 11
 Preibisch, T., & Mamajek, E. 2008, in *Handbook of Star Forming Regions, Volume II: The Southern Sky*, ed. B. Reipurth (San Francisco, CA: ASP), 235
 Rochau, B., Brandner, W., Stolte, A., et al. 2010, *ApJL*, **716**, L90
 Rodgers, A. W., Campbell, C. T., & Whiteoak, J. B. 1960, *MNRAS*, **121**, 103
 Ryon, J. E. 2019, *ACS Instrument Handbook*, Version 18.0 (Baltimore, MD: STScI)
 Sabbi, E., Anderson, J., Lennon, D. J., et al. 2013, *AJ*, **146**, 53
 Sabbi, E., Lennon, D. J., Anderson, J., et al. 2016, *ApJS*, **222**, 11
 Sabbi, E., Lennon, D. J., Gieles, M., et al. 2012, *ApJL*, **754**, L37
 Sabbi, E., Sirianni, M., Nota, A., et al. 2007, *AJ*, **133**, 44
 Sabbi, E., Sirianni, M., Nota, A., et al. 2008, *AJ*, **135**, 173
 Smith, N., & Brooks, K. J. 2008, in *Handbook of Star Forming Regions, Volume II: The Southern Sky*, ed. B. Reipurth (San Francisco, CA: ASP), 138
 Stolz, A., Brandner, W., Brandl, B., Zinnecker, H., & Grebel, E. K. 2004, *AJ*, **128**, 765
 Sung, H., & Bessell, M. S. 2004, *AJ*, **127**, 1014
 The Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, **156**, 123
 Vargas Álvarez, C. A., Kobulnicky, H. A., Bradley, D. R., et al. 2013, *AJ*, **145**, 125
 Weibacher, P. M., Streicher, O., Urrutia, T., et al. 2012, *Proc. SPIE*, **8451**, 84510B
 Weibacher, P. M., Streicher, O., Urrutia, T., et al. 2014, in *ASP Conf. Ser. 485, Astronomical Data Analysis Software and Systems XXIII*, ed. N. Manset & P. Forshay (San Francisco, CA: ASP), 451
 Westerlund, B. 1961, *ArA*, **2**, 419
 Zeidler, P. 2019, MUSEpack, v1.0, Zenodo, doi:10.5281/zenodo.3433996
 Zeidler, P., Grebel, E. K., Nota, A., et al. 2016, *AJ*, **152**, 84
 Zeidler, P., Nota, A., Grebel, E. K., et al. 2017, *AJ*, **153**, 122
 Zeidler, P., Sabbi, E., Nota, A., et al. 2015, *AJ*, **150**, 78
 Zeidler, P., Sabbi, E., Nota, A., et al. 2018, *AJ*, **156**, 211