Title: The necessity of multiple photometric redshift estimates with quality flags

Contributors: Mara Salvato, Roberto Assef
Co-signers: Matthew Temple, W. Niel Brandt, Gordon Richards, Franz E. Bauer, Raphael Shirley, Ilsang Yoon, Yue Shen, Weixiang Yu, Luka C. Popovic, Angela Bongiorno, Qingling Ni, Bozena Czerny, James Mullaney, Claudia M. Raiteri, Xin Liu, Mark Lacy, Manda Banerji, Filippo D'Ammando, Dragana Ilic, Andjelka Kovacevic, Jan-Torge Schindler, Claudio Ricci, Yongquan Xue Ni, Ioannis Liodakis, Robert Nikutta

0. Summary Statement

We discuss the importance of:

a) providing the community with additional photometric redshifts computed with independent methods tuned for sources other than those dominated by galaxy emission.

b) a flag indicating when the methods provide different results.

c) a column for each method used for the computation of the photometric redshift, indicating the quality of the fit.

It is understood that in order to provide the community with science-ready catalogs, photometric redshifts and physical parameters will be computed for all extragalactic sources. Although the effort is laudable, as long as the photometric redshifts are computed (either via SED fitting or machine learning) under the assumption that the source is a pure galaxy, this would be detrimental to the science cases of the AGN community.. AGN and QSOs are just a fraction of the extragalactic sources, but still will number of the order of 10-100M in the 10 year lifetime of LSST (see Chapter 10 of the LSST Science Book). For all these sources the relative host/AGN contribution is a priori unknown, and this will affect the redshift determination. Physical parameters estimated with systematically uncertain redshifts and neglecting the contribution of AGN to the SED can result in a significant negative impact to science applications with LSST data.

1. Scientific Utility

Due to its unprecedented depth, the great majority of objects detected by LSST will lack spectroscopic follow-up observations. Science investigations will require that users rely on photometric redshifts for most extragalactic sources. Regardless of the method, photometric redshifts primarily rely on the mapping of Galactic-extinction corrected (see this LoR) observed colors to redshift. In particular, the redshift is determined by looking at the color of prominent features such as, e.g., the Balmer break which is present in most of the galaxies. In galaxies hosting an active nucleus, this feature is often outshone by the power-law that characterizes the continuum emission of the active nucleus, thus making the computation of the photometric redshift a real challenge. Even worse, ignoring the AGN nature of a source and treating it as a normal galaxy will most likely result in a catastrophically wrong photometric redshift. This concept is illustrated in Figure 1, where the photometric redshifts for known AGN in the COSMOS field are computed first assuming a library of templates that includes AGN, QSO and hybrids (host+AGN component with various relative contribution) and then assuming that they are normal galaxies. In the latter case, the accuracy is quite poor and the fraction of outliers extremely high, with 33% of cases for which the code failed to assign a redshift. The latter failure is actually preferred (a NULL result would alert a user that the source is peculiar) to the case of a photometric redshift computed using the wrong assumptions. That physical parameters will be computed using this redshift estimate further compounds the issue, since:

a) probably the redshift is wrong,

b) even if the redshift were not catastrophically wrong, the templates used for estimating the physical parameters do not account for the AGN component, thus over-estimating e.g., the stellar mass and star-formation rate.

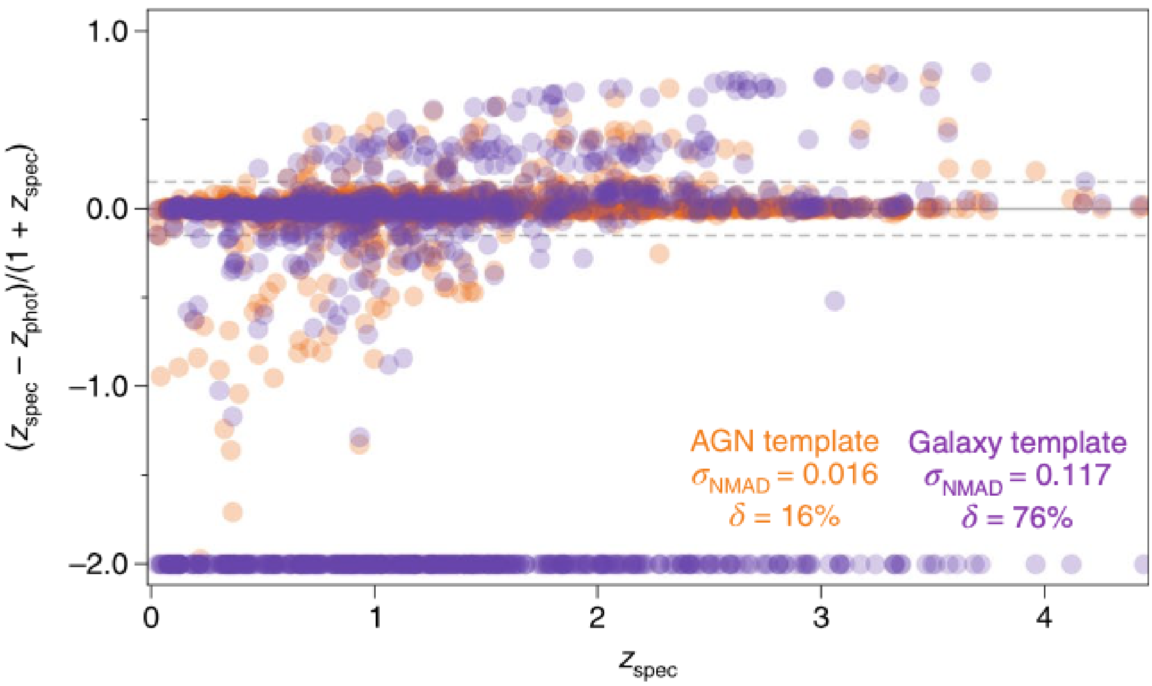

Figure 1: Difference between photo-z obtained using either galaxy (violet) or AGN (orange) templates for sources that are AGN detected in COSMOS and with a reliable spec-z. For the 33% of sources for which no normal galaxy template could be used to compute the photo-z, we artificially set the photo-z to −2. Credit : Salvato, Ilbert & Hoyle, 2019.

The current plan for the release of the catalogs by Rubin, outlined in the DM photometric redshifts roadmap (*DMTN-049_2*), will suffer from this as photometric redshifts and physical parameters will be computed assuming that all the sources are inactive galaxies.

We require that photometric redshifts are computed with at least one additional method that takes into account the possibility for the source to be an AGN, or, even better, a code able to accommodate requirements for galaxies and AGN at the same time. Codes based on SED fitting that can consider also templates (and priors) typical of AGN already exist (e.g., EAZY, LePHare), with some specific code also provided by approved in-kind contributions (e.g., GER-MPE; PI: Salvato M.). While we refer here to SED fitting codes, it is worth noting that, thanks to the increased number of sources with spectroscopic redshifts that can be used for training, machine learning is also now able to compute reliable photometric redshifts for galaxies and AGN at once. The adoption of at least another code able to also include templates and priors typical of AGN is not only fundamental for all the community that is interested in science related to AGN, but also for communities that require high reliability of their redshift and physical parameter estimates or identification of AGN for removal, such as DESC.

The current roadmap considers the possibility of outputting flag parameters, but does not provide specific requirements on these. A goodness-of-fit or equivalent estimate for machine learning methods should be included, allowing the user to identify objects for which the photometric redshifts and physical parameters should not be trusted. Furthermore, if multiple algorithms were to be used, a flag highlighting the quality of the fit for each method. Providing the users with multiple photometric redshift estimates and quality flags will allow them to make informed decisions and extract most efficiently the science from the LSST data release catalogs, benefiting the entire project community and the observatory science output.

3. References

Arnouts, S., Cristiani, S., Moscardini, L., et al, 1999,MNRAS, 310

Ilbert, O., Arnouts, S., McCracken, H., et al, 2006, A&A, 457

Brammer, van Dokkum, Coppi, 2008, ApJ, 686

Salvato, M., Ilbert, O., & Hoyle, B. 2019, Nature Astronomy, 3, 212