Science aim: Provide the crucial infrared link between cosmological surveys in the optical and the radio to enable a leap forward in understanding the origin and evolution of the Universe

Scientific Importance: Studies of Large-scale Structure (LSS) - the evolution in the distributions of galaxies on scales ranging from a few Mpc up to several hundred Mpc - are a cornerstone to understanding our cosmic origins. LSS provides a key test of fundamental physical phenomena, including: general relativity and its alternatives, the validity of cosmic inflation, the equation of state of dark energy, and the density contrast evolution of dark matter. In recognition of this, upcoming observatories will survey thousands of square degrees, including: Euclid WFIRST, DESI and eBOSS in the optical, and the SKA in the radio, with studies of LSS as one of their primary goals.

Even with this planned suite of optical and radio LSS surveys, there remains a fundamental knowledge gap. Different populations of galaxies trace the underlying mass distribution, and the geometry of the Universe, in different ways, with different levels of `bias'. Thus, *optical and radio surveys give a highly incomplete picture of LSS*. Optical light almost entirely misses the obscured phases of galaxy assembly that are so numerous at $z > 1$, hence their use is largely limited to tracers at $z<1.5$ (the exception being optical quasars, but their rarity at any epoch is a further limiting factor). Radio observations on the other hand, while independent of obscuration and sensitive to star-forming galaxies up to high redshifts that are missed in the optical, suffer from a different limitation: it is hard to characterize sources (i.e separate star forming galaxies by SFR, or isolate AGN), and to obtain redshifts. Hence, many radio surveys rely on optical spectroscopy for distances. Thus, even *combining* future optical and radio surveys still means that the ultimate potential of LSS to test fundamental physics is stymied by having only a small number of biased mass tracers that do not work well together.

The key requirements for using galaxies as probes of LSS are thus (1) to have carefully selected samples whose properties, at least statistically, are well understood (e.g. the LRG samples from BOSS), and (2) to have multiple *separate* galaxy populations to trace LSS across a broad redshift range. The second requirement is particularly important; having multiple independent tracers means that the ratio of the scale dependence of their biases is then not limited by sample variance.

The Far-Infrared Surveyor will have a profound impact on studies of LSS, by providing a fundamental link between the optical and radio surveys that incorporates the strongest elements from both. IR-luminous, dusty star-forming galaxies are the ultimate complement to optical tracers, as: i) they are largely missing from the optical catalogs; ii) they are highly biased; iii) their redshift kernel extends to higher redshifts than optically selected galaxies. They are also a powerful complement to radio tracers, as i) IR galaxies are an overlapping but distinct tracer population to those in SKA surveys, and ii) the IR provides the crucial source characterization that is so hard to get in the radio, including the separation of AGN and star-forming systems and approximate SFRs. Moreover, the scale of the planned spectroscopic surveys with e.g. Euclid, eBOSS, SKA-MIDI, at a few thousand square degrees with of order 1 source per square degree, is complementary to what can be achieved with the FIR Surveyor. Thus, combining optical, infrared and radio LSS surveys gives a unique and extremely powerful suite of multi-tracers - a highly complete census of well-characterized star-forming and passively evolving systems to $z=3$, across large cosmological volumes. This allows for a large number of fundamental investigations, including:

- Search for signals of primordial non-gaussianity as a test of the validity of cosmic inflation
- Combine optical and FIR selection to define multiple different samples for cosmic magnification tests (two non-overlapping populations at different redshifts, to see how one magnifies the other).
- Weak lensing measurements, both on their own and in concert with WFIRST and radio WL studies, where cross-correlating morphology measures can improve fidelity across all wavelengths.
- Measure the alignment of the IR dipole to test the assumption that the LSS and CMB frames coalign.
- Cross-correlating FIR samples with CMB maps. The CMB lensing kernel is broad, with significant support from $z \sim 0.5 - 4$. Cross-correlating these maps with FIR samples (which must extend to $z>2$) then gives a clean way to measure the bias of the galaxy population.
- Measure the integrated Sachs-Wolfe (ISW) effect through CMB cross-correlation. The ISW effect is a useful independent probe of dark energy, and can also be used to test GR.

Measurements Required: For LSS the key requirement is a large volume containing enough objects for the multi-tracer approach. Scaling from existing forecasts for *spectroscopic* LSS surveys, the minimum is about 5000 sq deg, with 0.5 sources per sq arcmin, to redshifts of z=2. This assumes all targets have spec-z's. Higher redshifts though are particularly advantageous given that most of the optical surveys will be limited to z<2, and the z>2 Universe is where a FIRS is particularly powerful. So, a reasonable requirement is to extend the redshift range to z<2.5. The ideal would be the same or greater source density, higher z range, over 20,000 sq deg.

Longevity/Durability*:* No other existing or planned observatory can perform meaningful studies of LSS in the IR. Herschel did not reach the depths required, with further barriers raised by its large beam and difficulty in getting redshifts. Similar barriers are faced by all existing ground-based FIR facilities. Even projecting to a hypothetical 50+m ground-based observatory with large-format detectors a key limitation remains - the narrow atmospheric windows in the FIR mean it is hard to image at more than ~2 separate wavelengths, especially over the IR SED peak in the rest-frame at z~2, and impossible to obtain complete and unbiased redshift samples. JWST and ALMA cannot survey (by a few orders of magnitude) the required areas.

Performance Requirements

Parameter	Unit	Required value	Desired Value	Comments
Wavelength/band	um	80 - 500	40 - 800	Req: 3-4 continuum bands over the rest-frame peak in the IR SED, plus a few accessible bright lines for redshifts, for sources up to z=3. Ideal: More continuum bands/lines improve source characterization
Number of targets		0.5 sources per sq arcmin up to z=2.5	2 sources per sq arcmin up to z=3	Based on a 5000 sq deg area, matching SKA-MID area and source density, and on complementarity with Euclid spec surveys. Both assume spec-z's are available for most sources.
Survey area	deg.^{2}	5000 sq deg	20000 sq deg	Req: baseline synergy with SKA-MID1, Euclid, and WFIRST.. Ideal: allows all of the studies listed above.
Angular resolution	arcsec	5" at 100um	<0.5" at 100um	Req: small enough PSF at longer wavelengths so that confusion does not curtail the multi-tracer work, but assuming spec-z's are available for most sources. Ideal: allows weak lensing at 100um
Spectral res	$\lambda/\delta\lambda$	R=500	R=1000	Driver is accurate FS line redshifts, which doesn't push to high values. BL redshifts would hurt with e.g. RSD.
Continuum Sensitivity	mJy	0.03	0.03	Req: Continuum detections in two bands for sources with CII-based redshifts
Spectral line sensitivity (1 s)	W m^{-2}	1.5e-19	5e-20	Req: redshifts for ~1 starburst per sq arcmin up to z=2.5, which means detect a CII line flux of ~1.4e-19 Wm-2 Ideal: Push to either z=4 or 2 sources per sq arcmin, either is about 5e-20 W m-2
Field of Regard		half-sky	all-sky	

Key references: McDonald P., Seljak U., 2009, J. Cosmol. Astropart. Phys., 0910, 007;
Seljak U., 2009, Phys. Rev. Lett., 102, 021302; Zhao et al 2016 - useful forecasts for eBOSS