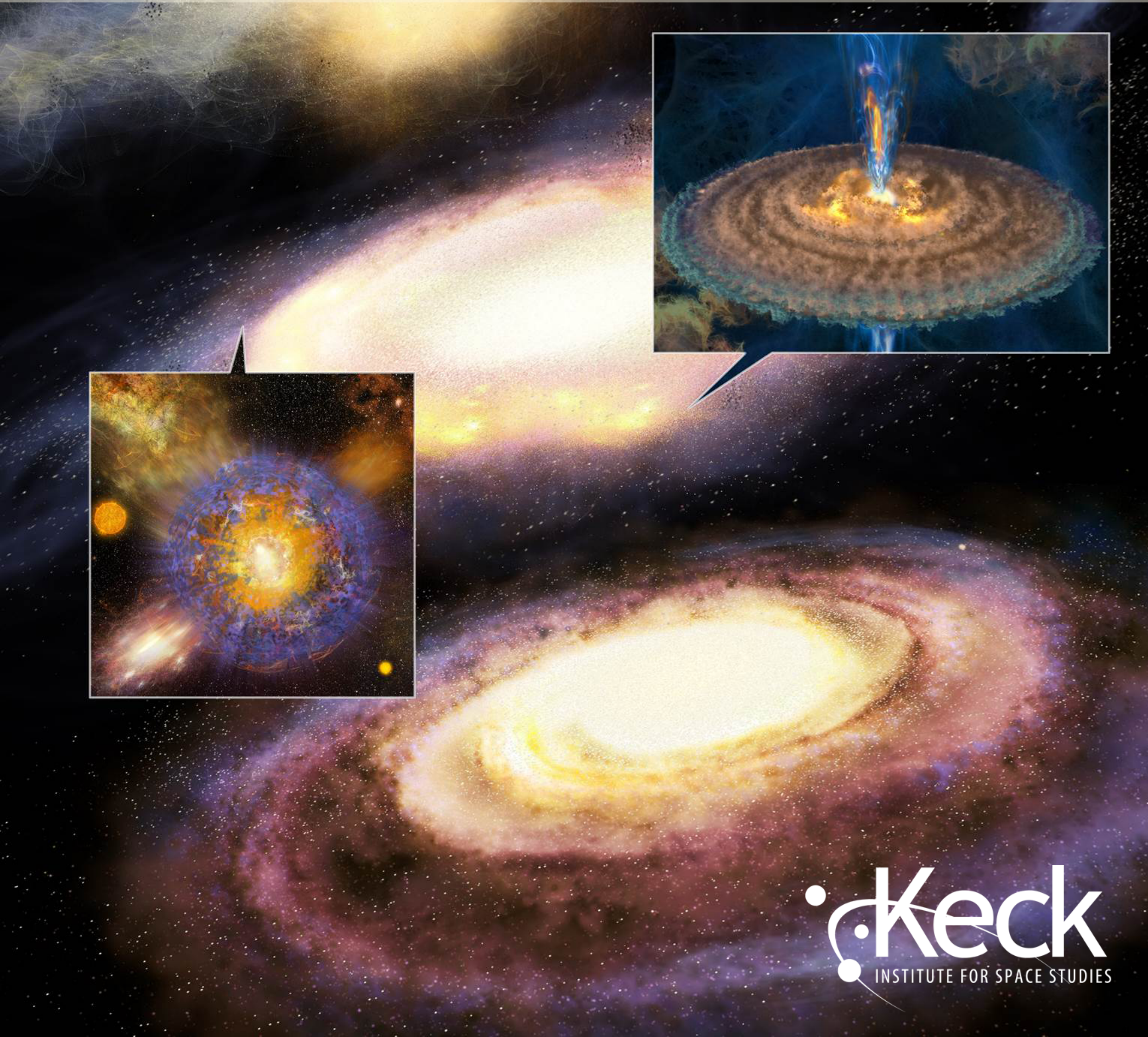
Bridging the Gap: Observations and Theory of Star Formation Meet on Large and Small Scales

Study Report



Keck
INSTITUTE FOR SPACE STUDIES

Study Date: November, 3–7, 2014

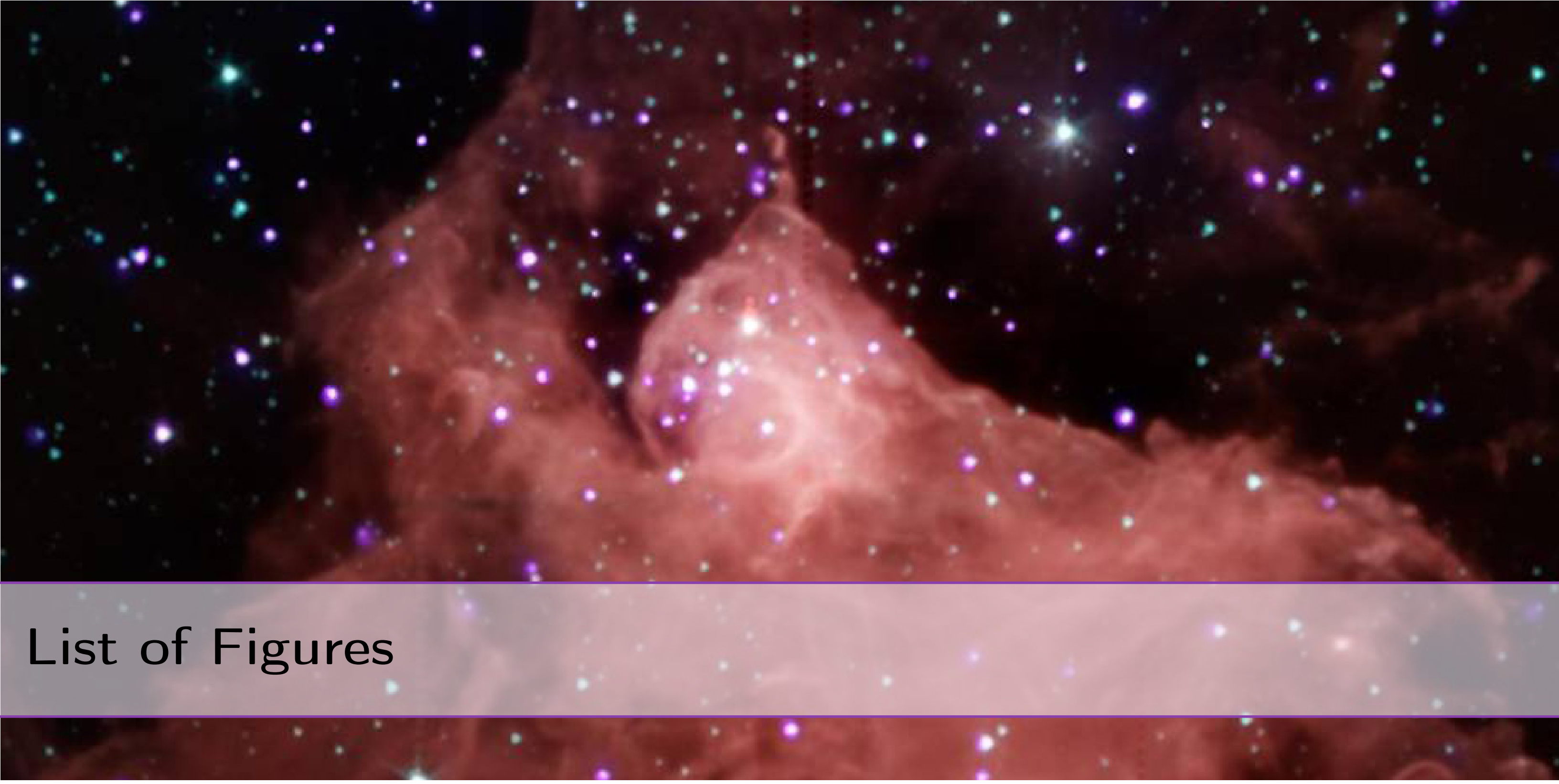
Team Leads: Jorge L. Pineda (JPL) and Philip F. Hopkins (Caltech)

Cover Image: Chuck Carter / Keck Institute for Space Studies

Header Images: Star-forming nebulae, Credit: NASA/JPL-Caltech/UCLA; Star-forming cloud Cepheus B, Credit: NASA/CXC/JPL-Caltech/PSU/CfA; Workshop participants; Messier 78, Credit: NASA/JPL-Caltech; Star formation in M17 SWex, Credit: NASA/JPL-Caltech/Penn State; Star Formation in the Circinus Molecular Cloud Complex, Credit:NASA/JPL-Caltech/UCLA; Stellar grouping R136 in the 30 Doradus Nebula, Credit: NASA, ESA, and F. Paresce (INAF-IASF, Bologna, Italy), R. O'Connell (University of Virginia, Charlottesville), and the Wide Field Camera 3 Science Oversight Committee; Star formation in the Serpens Cloud Core, Credit: NASA/JPL-Caltech/2MASS; Star-formation region NGC 602, Credit: X-ray: NASA/CXC/Univ.Potsdam/L.Oskinova et al; Optical: NASA/STScI; Infrared: NASA/JPL-Caltech; Star-forming region W5, Credit: Lori Allen, Xavier Koenig (Harvard-Smithsonian CfA) et al., JPL-Caltech, NASA.

Editing and Formatting: Meg Rosenburg

Contents



List of Figures

Participants

Lee Armus
Caltech Campus

Guillermo A. Blanc
Carnegie Observatories

John Carpenter
Caltech

C. Darren Dowell
JPL/Caltech

Paul Goldsmith
JPL/Caltech

David Guszejnov
Caltech Campus

Chihomi Hara
University of Tokyo

Christopher Hayward
Caltech Campus

Philip Hopkins
Caltech Campus

Ji-hoon Kim
Caltech Campus

Evan Kirby
Caltech Campus

Janice Lee
Space Telescope Science Institute

Adam Leroy
National Radio Astronomy Observatory

Sharon Meidt
Max Planck Institute for Astronomy

Leonidas Moustakas
JPL/Caltech

Eric Murphy
Caltech Campus

Karen Olsen
Niels Bohr Institute

Jorge L. Pineda
JPL/Caltech

Gwen Rudie
Carnegie Observatories

Nick Scoville
Caltech Campus

Kartik Sheth
National Radio Astronomy Observatory

Allison Strom
Caltech Campus

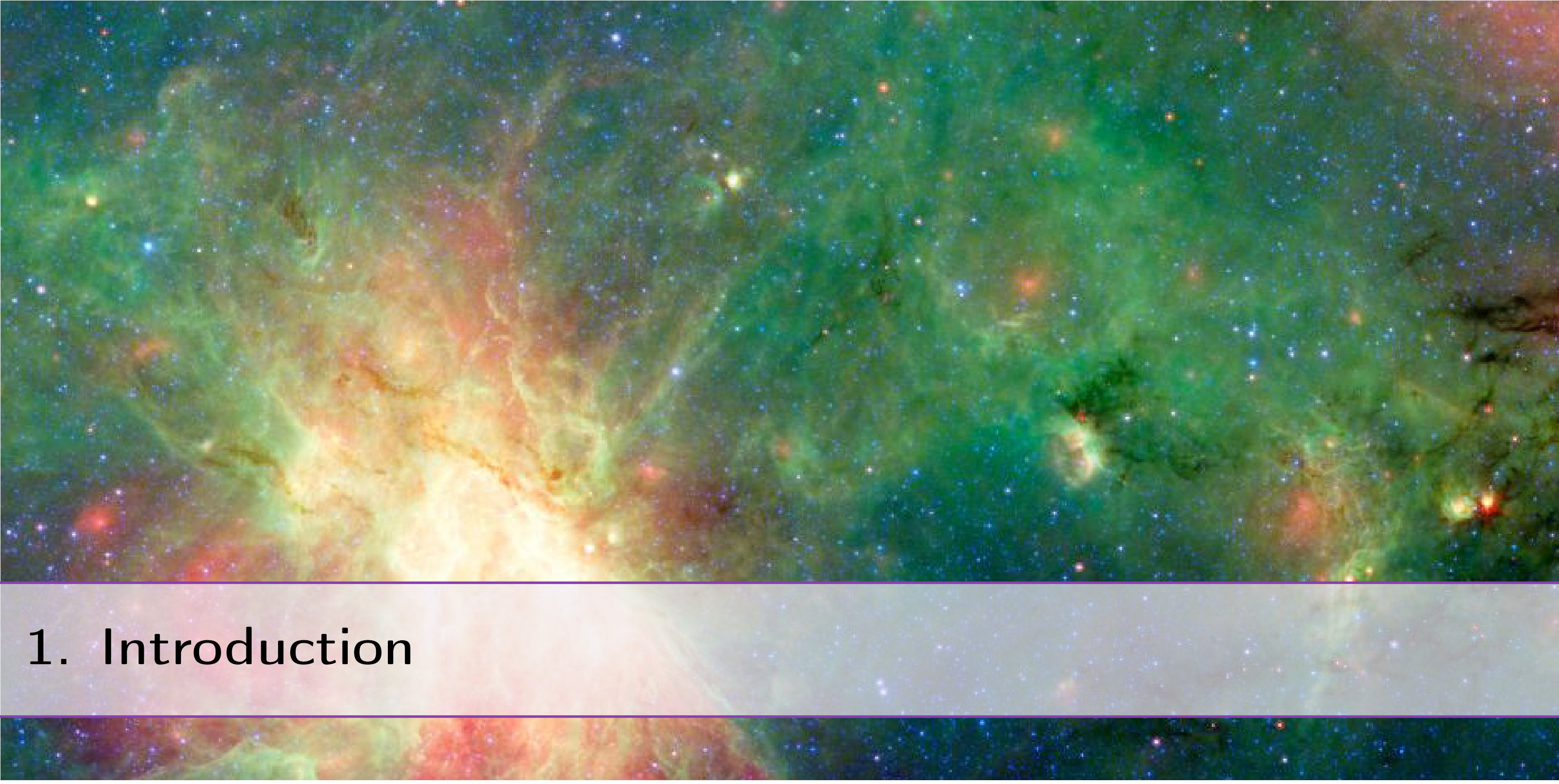
Jonathan Tan
University of Florida

Julie Wardlow
University of Copenhagen

Harold Yorke
JPL/Caltech

Foreword

This report presents the discussion that took place at the "Bridging the Gap: Observations and Theory of Star Formation Meet on Large and Small Scales" workshop hosted at the Keck Institute of Space Studies (KISS) at the California Institute of Technology, November, 3–7, 2014. The workshop gathered experts on the study of galaxy evolution with expertise in observations, theory, and instrumentation to discuss how to connect the physics of galaxy and star formation at large scales to studies of the local universe, both observational and theoretical. The group discussed several critical questions for understanding galaxy evolution over a wide range of scales, and for enabling future progress in the field. From this set of questions, we propose a set of critical goals for advancement and a list of recommendations that will allow the accomplishment of these goals.



1. Introduction

The drive to understand galaxy formation and evolution over the lifetime of the universe has justified vast space-based and ground-based telescope facilities, as well as the development of new technologies. The details of star formation, and the modes by which that activity couples to the broader galactic environment, occurs on small spatial scales. These scales can only be traced with great sophistication in the local universe, as witnessed by observations using the Spitzer Space Telescope, the Herschel Space Observatory, the Stratospheric Observatory for Infrared Astronomy (SOFIA), and the Atacama Large Millimeter Array (ALMA). A critical realization of the last decade, however, is that the large scales and small scales are strongly coupled, and cannot be treated in isolation. At the same time, the scales studied in the local universe are "sub-grid" for the purpose of cosmological simulations that make valiant efforts to include the physics of star formation and its feedback to the local environment. The fundamental uncertainties in how this sub-grid physics is incorporated into the larger picture are by far the greatest limitation in understanding galaxy formation and evolution.

The main goal of cosmological simulations is to understand the formation and evolution of the universe over a wide range of scales going from the Hubble volume to sub-light-year scales within galaxies. However, due to computational limitations, the physics of star formation in galaxies and the effects this process has on the formation and evolution of galaxies are often greatly simplified. For instance, accounting for the effect of stellar feedback in cosmological simulations is a crucial factor in galaxy formation and evolution. Without stellar feedback in galaxies, the gas would rapidly cool and collapse, converting all available gas into stars within a dynamical time. This consequence is in sharp conflict with observations—there are vastly fewer stars in our universe than the models would predict. Until recently, numerical simulations have been unable to regulate star formation efficiently enough to reproduce observations, with many consequences: models with too many stars also predict they form at the wrong times in the universe's history, that there are the wrong abundances of heavy elements, that galaxies look nothing like the Milky Way, and even that the observable properties of dark matter are (apparently) discordant with new precision-cosmology measurements. All of these problems may, in fact, be due to the fundamental problem of understanding how small and large scales interact.

New efforts have therefore focused on incorporating some treatment of the relevant physical processes governing the formation of stars and the mechanical and radiative feedback that these stars impart to the progenitor gas (e.g., *Hopkins et al.*, 2014). Because of computational limitations, however, it is impossible to include all processes and it is imperative to use observational results to determine which of those processes are the most relevant for these simulations. It is also important to study how the effect of these different star formation and feedback mechanisms in the evolution of galaxies depends on the different conditions found in between and within galaxies.

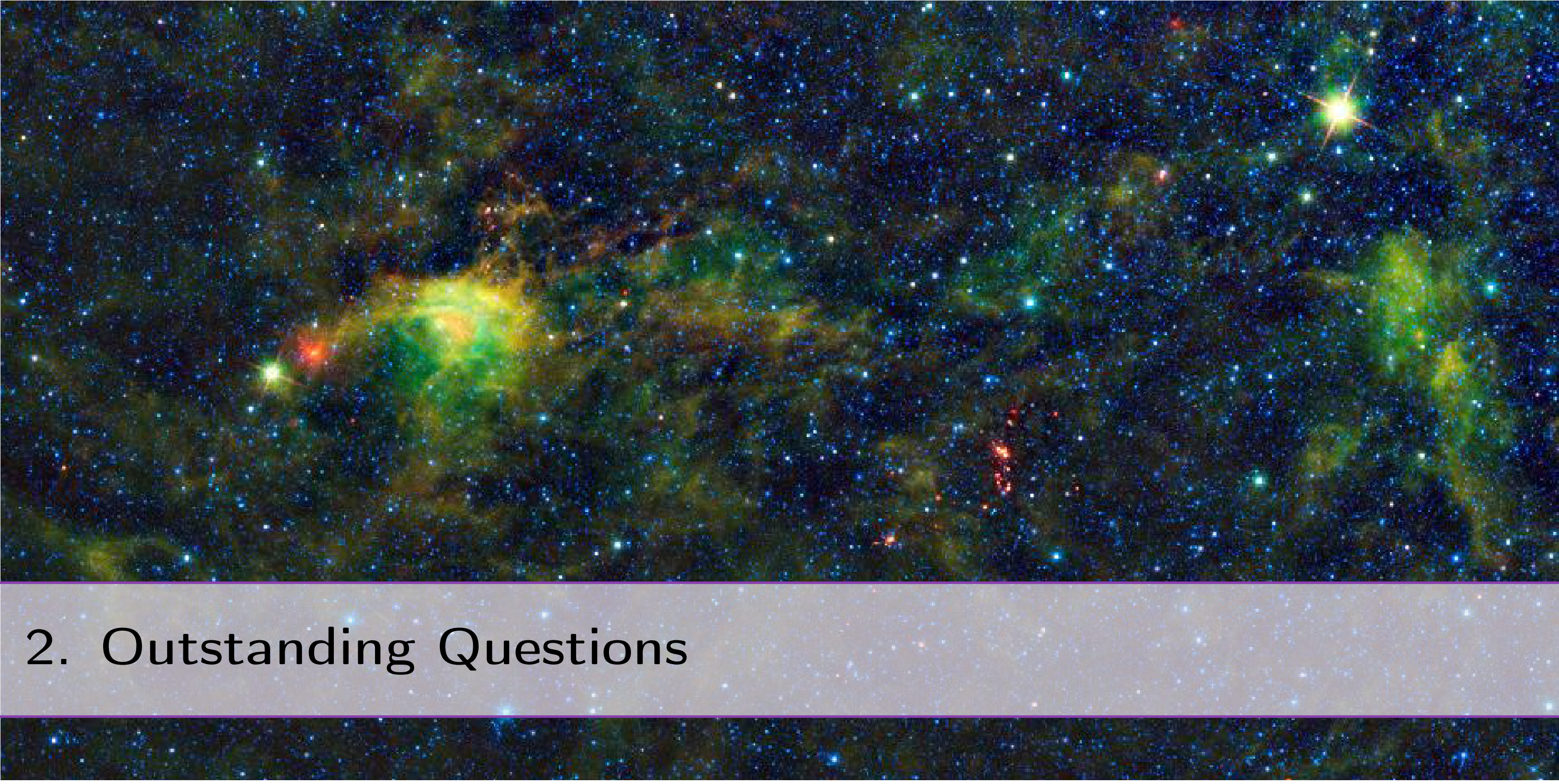
Simultaneously, great advances in the study of star formation in the Milky Way and nearby galaxies have been, and will continue to be made with Spitzer, Herschel, the Galaxy Evolution Explorer (GALEX), the Swift Gamma-Ray Burst Mission, SOFIA, the Hubble Space Telescope (HST), and ALMA, as well as other new facilities. Extensive surveys of the Milky Way and nearby galaxies using tracers of atomic gas (HI), well-shielded molecular gas (CO), and dust (far-infrared continuum emission), have become available. These surveys revealed several scaling relationships that provide strong tests for numerical simulations. The recent advances in detector technology have allowed the study of the local universe using far-infrared (FIR) spectral lines such as [CII], [OI], [OIII], [NII], etc. It has been shown that these tracers are excellent diagnostic tools for determining the physical and chemical state of the gas, thus tracing the impact of the radiative feedback of newly formed stars in its progenitor molecular gas. It has also been shown that these spectral lines, in particular [CII], can be used to trace different phases of the interstellar medium in galaxies that could not be traced before with the traditional HI and CO observations. Additionally, optical studies in nearby galaxies, which are able to resolve individual stars, have been used to study stellar populations, the origin of the stellar initial mass function (IMF), star formation timescales, the energetics of feedback into the interstellar medium, and supernovae populations. These new multi-wavelength observations have the potential to reveal new properties of galaxies that could provide strong constraints on numerical simulations.

The most recent technological developments for detecting FIR spectral line and dust continuum emission are associated with the development of the instruments on board of the Herschel Space Observatory (the Photoconductor Array Camera and Spectrometer (PACS), the Spectral and Photometric Imaging Receiver (SPIRE), and the Heterodyne Instrument for the Far-Infrared (HIFI)). These instruments allowed the observations of FIR lines in portions of a large sample of nearby galaxies. These include projects such as KINGFISH (Key Insights on Nearby Galaxies: A Far-Infrared Survey with Herschel, *Kennicutt et al.*, 2011) and SHINING (Survey with Herschel of the ISM in Nearby INfrared Galaxies, *Fischer et al.*, 2010) for normal galaxies and DGS (Dwarf Galaxy Survey, *Cormier et al.*, 2015) for low-metallicity dwarf galaxies. In the Milky Way, a great advance in understanding the emission of the [CII] 158 μm line has been obtained with the Galactic Observations of Terahertz C+ (GOT C+) project (*Pineda et al.*, 2013; *Langer et al.*, 2010), which surveyed the Galactic plane. Individual star forming regions have been studied in detail in the Warm and Dense ISM project (WADI, *Ossenkopf et al.*, 2011). These projects have revealed that FIR spectral lines can be used as diagnostics of the physical conditions of the gas in galaxies, to trace star formation activity, and to trace gas that could not be previously traced with HI and CO observations. This set of new observations is highly synergic with the rich data set currently available on HI, CO, and dust continuum emission in the local universe.

Several optical studies of stellar populations in nearby galaxies using the Hubble Space Telescope have provided important insights on the properties of stellar populations in nearby galaxies. Examples of such projects are the "Panchromatic Hubble Andromeda Treasury," devoted to study M31's star forming disk, covering a range from the ultraviolet through the near infrared, and the ACS Nearby Galaxy Survey Treasury (ANGST, *Dalcanton et al.*, 2008), which is a systematic survey to establish a legacy of uniform multi-color photometry of resolved stars for a volume-limited sample of nearby galaxies. These studies provide important observational constraints that help to break degeneracies in which physical mechanisms are responsible for the observed properties of galaxies, stars, and supermassive black holes.

Our greatest hope for progress lies in combining these next-generation observations with the latest advances in models of realistic physical processes in galaxy and star formation. This has the potential to generate fundamentally new ideas about the type of observations which will constrain the current "sub-grid" (i.e., unknown) physics in the models, and thus to resolve these deep cosmological mysteries. This, in turn, opened our discussion of the most optimal technologies and mission concepts that can have the maximal impact on our understanding of galaxy formation and evolution.

In the following sections we will report a set outstanding questions that our group identified, whose answers we believe will have the largest impact in advancing our understanding of galaxy evolution. From these questions we propose a set of advancement paths that will help us find their answers. We conclude with a set of specific, actionable recommendations whose implementation will be critical for the future progress of the study of star formation and galaxy evolution.



2. Outstanding Questions

The first task of the workshop was to identify specific aspects in the field of star formation and galaxy evolution where observational and theoretical progress is critically needed. These outstanding questions include how the interstellar medium evolves to form star-forming clouds, how star formation proceeds within them, what external influences can affect the process of star formation, and how the feedback from newly formed stars regulates the process of star formation and thus influences how galaxies evolve. The group determined that progress on these questions would represent a significant advance in our understanding of star formation (SF) and its connection to galaxy evolution and the growth of structure in the universe.

2.1 Towards a full understanding of the life cycle of the ISM: What forms and destroys GMCs? Can we account for the properties of different ISM phases?

Interstellar gas cycles from molecular gas to atomic and ionized gas, and then back to molecular gas. At the same time, giant molecular cloud complexes can be broken up into smaller units in the inter-arm regions, only to be reassembled in the next spiral arm. To understand this we must diagnose all components of this multiphase interstellar medium, including spiral arms and inter-arm regions in the Milky Way and nearby galaxies (Figure 2.1).

While atomic and molecular gas have been traditionally studied with HI 21cm and CO J=1-0 observations, velocity-resolved observations in the Milky Way of the gas that is in the transition between atomic and molecular that is traced by the [CII] 158μm line have only recently been made possible with the HIFI instrument on the Herschel Space Observatory (*Langer et. al.*, 2010; *Pineda et al.*, 2013). The [CII] line traces different phases of the interstellar medium (ISM),

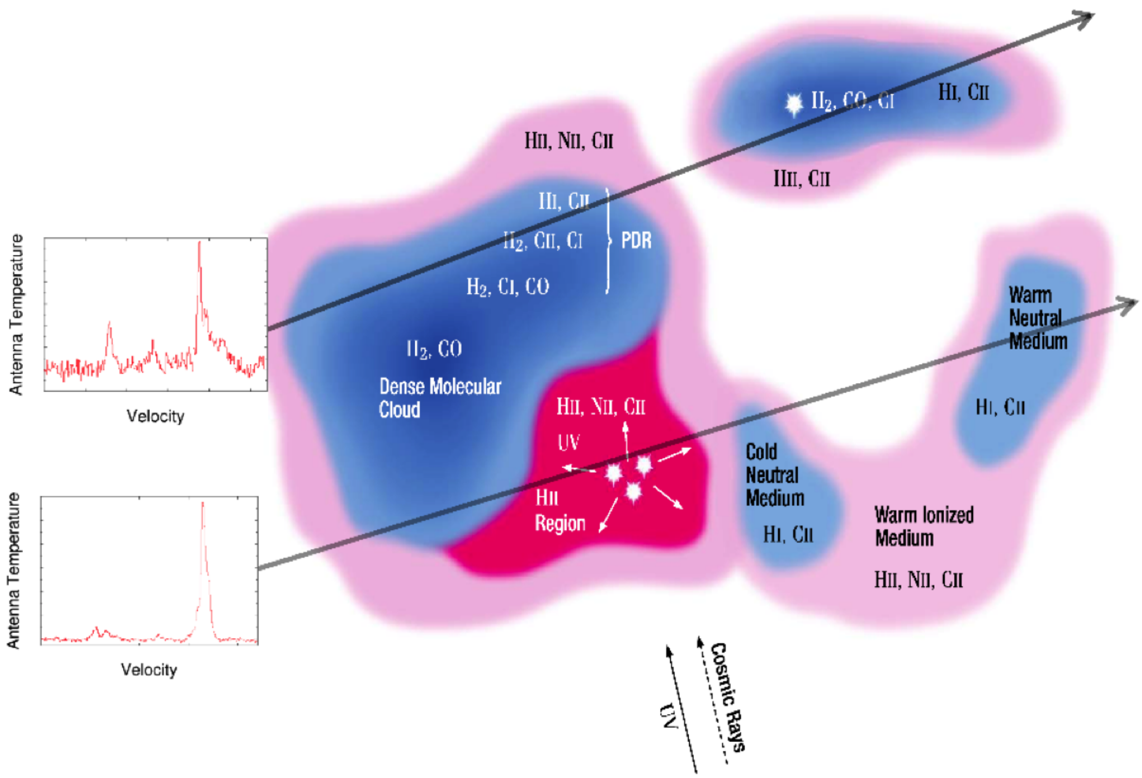


Figure 2.1: Schematic representation of physical conditions, primary emitting species, and morphology of a star-forming complex. [CII] arises from all regions in which hydrogen is atomic (HI) as well as molecular (H_2). Regions within clouds move with slightly different velocities, allowing us to distinguish them spectrally even if they lie along the same line of sight. Different lines of sight (indicated above and accompanied by GOT C+ [CII] spectra) manifest the different phases of the ISM through distinct velocity features in the spectra. High velocity resolution imaging of interstellar clouds will delineate the 3-D spatial distribution of these components and their interrelationship.

including the diffuse ionized medium, warm and cold atomic clouds, clouds in transition from atomic to molecular, and dense and warm photon dominated regions (PDRs). In particular, the [CII] line is a tracer of the CO-dark H_2 gas (*Grenier et al.*, 2005; *Wolfire et al.*, 2010; *Langer et al.*, 2010), which is molecular gas that is not traced by CO but by [CII]. This gas is the likely precursor of dense molecular gas that will eventually form stars. The CO-dark H_2 gas represents $\sim$30% of the molecular mass of the Milky Way and this fraction increases with galactocentric distance, which is an effect of the metallicity gradient of the galaxy (*Pineda et al.*, 2013). Quantifying the CO-dark H_2 gas component in galaxies provides an independent method to determine the true hydrogen mass of galaxies. It is therefore highly desirable to identify and characterize CO-dark H_2 clouds over a wide range of environmental conditions. Large scale mapping of the [CII] line in the Milky Way and nearby galaxies together with maps of the molecular and atomic phases from

HI and CO will therefore provide a complete picture of all the phases of the ISM and significantly improve our understanding of the life cycle of the interstellar medium.

Turbulence may play an important role in the formation and evolution of interstellar clouds. In a standard scenario where cold atomic (CNM) clouds are formed from warm atomic (WNM) gas by thermal instability, we can picture the role of turbulence in two ways. First, large scale instabilities, density waves, and supernovae drive compressional motions that increase the thermal pressure and trigger the thermal instability. Alternatively, regions undergoing thermal instability may generate turbulence, and convert the CNM into a complex network of pancakes and filaments. Because of this dynamic nature of both the triggering and evolution of the thermal instability, departures from thermal pressure equilibrium may be widespread in the ISM, and the notion of a dynamic multiphase ISM has been proposed. Only a careful study of both the spatial structure and kinematics of diffuse gas in transition between phases can tell us the role of turbulence and dynamic pressure in the life-cycle of the ISM.

The formation of GMCs is a prerequisite for massive star formation, yet the process has not yet been directly observed. Four mechanisms have been proposed to consolidate gas into GMC complexes: (1) self-gravitating instabilities within the diffuse gas component (often in a spiral arm where the density is highest and the Jeans time is shortest (e.g., *Ostriker & Kim*, 2004); (2) random collisional agglomeration of clouds (*Kwan & Valdes*, 1987); (3) accumulation of material within high-pressure environments such as shells and rings generated by OB associations (*McCray & Kafatos*, 1987); and (4) compression in the randomly converging parts of a turbulent medium (*Hennebelle & Perault*, 2000). We can attempt to distinguish among these processes and consider new cloud formation schemes by:

- Accounting for all the molecular hydrogen mass (the H_2/C+ CO-dark H_2 clouds as well as the H_2/CO clouds) when computing global measures of the interstellar medium.
- Clearly identifying CNM clouds via the density sensitivity of [CII] compared to HI 21 cm.
- Constructing spatial and kinematic comparisons with sufficient resolution, spatial coverage and dynamic range to discriminate among the above 4 scenarios.

Analogous observations of the various phases of the ISM should be made on nearby galaxies, e.g., M33, M31, and M51. The advantage of such measurements are an unambiguous interpretation of the locations of these components within the galaxies, but at the cost of much reduced spatial resolution. This can only be partially compensated by use of future facilities with increased sensitivity and spatial resolution. Theoretical modeling the evolution of the lifecycle of the ISM will be of crucial importance.

2.2 In what "mode" do stars form? What triggers their formation? What are the relevant time and spatial scales?

There are a number of major open questions about how stars form. Star formation is highly localized, i.e., "clustered," within giant molecular clouds (GMCs). Why is this? Why do certain regions of GMCs, typically referred to as "star-forming clumps" undergo relatively active star formation while most regions of these molecular clouds are essentially inert? Empirically, there have been claims there is a threshold for star-forming gas, requiring column densities equivalent to $A_V > \sim 8$ mag. Such threshold effects can be explained if magnetic fields are playing an important role in supporting cloud material against gravitational collapse, since the ionization fraction (due to FUV photoionization of dust) remains high enough in low A_V regions to keep the mostly neutral gas coupled to the B-field (*McKee*, 1989). The question of "What sets the star formation rate in galaxies?" may then reduce to the question of "What creates magnetically supercritical (high A_V) clumps within GMCs?" (rather than "What creates GMCs?"). Are star-forming clumps created by some external triggering process (e.g., cloud collisions or stellar feedback), or are they created by some internal instabilities in the GMC (e.g., global contraction of the cloud or local turbulence)?

How close are the "initial conditions" of star formation to virial and pressure equilibrium? We can define these initial conditions for star-forming clumps as the state of the gas clump when the first star starts to form (and this can be generalized as a function of scale around this star). For individual stars, we can define this as the conditions of the localized gas around the star when it starts to form a near-stellar-density protostellar object (or a first hydrostatic core). Here we would especially want to consider scales that enclose the mass shells that will eventually be incorporated into the star. Some models of star formation invoke initial conditions of individual star formation that are "cores" that are in quasi-virial equilibrium and near pressure equilibrium with their surrounding, clumpy, environment. For low-mass star formation, such "Core Accretion" models have been summarized by *Shu, Adams & Lizano* (1987). Generalization of such models for massive cores have been presented by, e.g., *McLaughlin & Pudritz* (1997) and *McKee & Tan* (2002, 2003). Alternatively, "Competitive Accretion" models have been presented by *Bonnell et al.* (2001) and *Wang et al.* (2010), mostly based on behavior seen in some numerical simulations, where the gas that eventually accretes to a massive star is drawn in from a wider (clump) scale amidst a forming cluster of lower-mass stars that are competing for the mass. The initial conditions of massive stars are quite different in these models. In core accretion, there is a massive, starless core that feeds a significant fraction of its mass, via a central disk, to a single massive star or a small N multiple. In competitive accretion, there is no such massive starless core, but there can be a clump that fragments, putting most of its mass into a stellar cluster and a small fraction into the massive star near its center. Protostellar mergers have also been seen in some simulations of competitive accretion, and this mechanism has been invoked as being potentially important for building up the mass of the most massive stars in very dense protocluster environments (*Bonnell et al.*, 1998; *Bally & Zinnecker*, 2005).

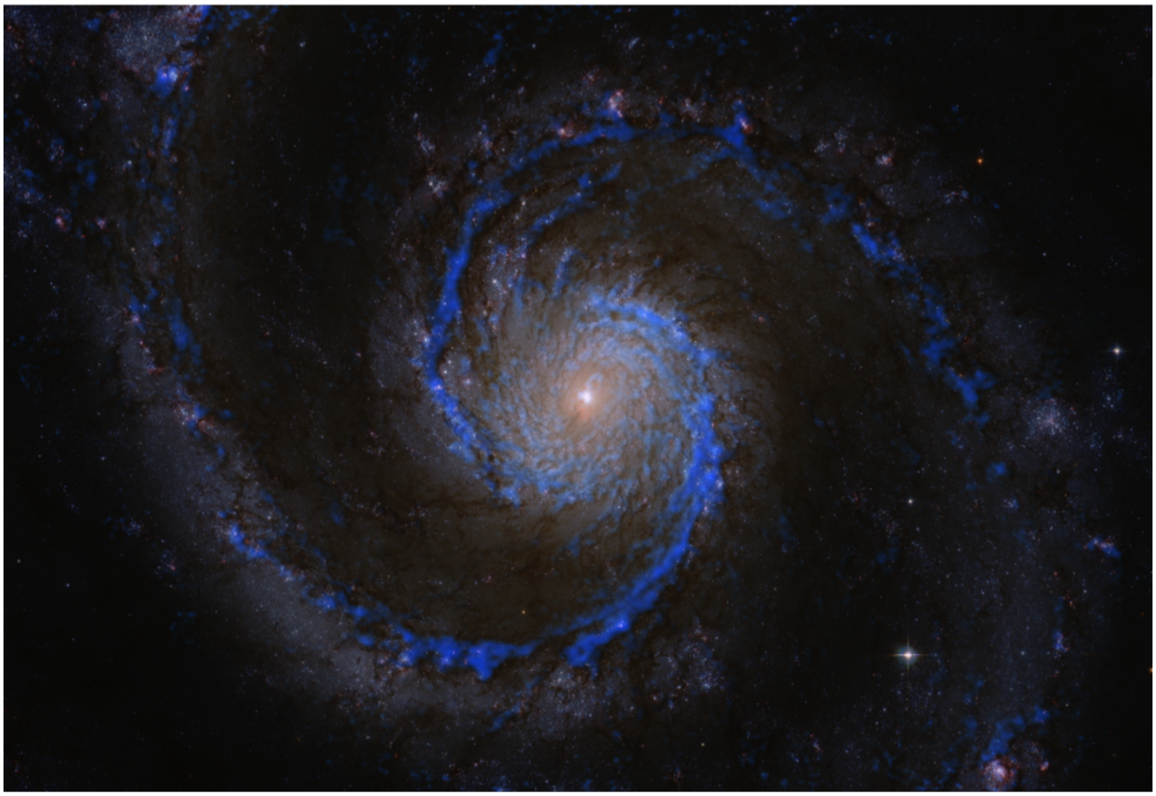


Figure 2.2: Grand design spiral galaxy M51 observed in optical emission with the Hubble Space Telescope (*T. Rector; U. Alaska*) and in CO with the Plateau de Blure (PdBI) and Institut de Radioastronomie Millimétrique (IRAM) 30m telescopes (*Schinnerer et al.*, 2013). Studying the distribution of the various constituents of the interstellar medium in galaxies like M51 can provide critical information on the properties of clouds over a wide range of environmental conditions.

What is the timescale of star formation, both of individual stars and of star clusters? Especially, what is the timescale relative to the dynamical time (or free-fall time) of the gas core or clump? For star clusters, rapid formation models (in about 1-2 free-fall times) have been proposed by *Elmegreen* (2000, 2006) and *Hartmann & Burkert* (2007), while slow models have been proposed by *Tan, Krumholz & McKee* (2006) and *Nakamura & Li* (2007), in which quasi-virial equilibrium is established in the clump after a dynamical timescale. Slow models require a mechanism to prevent global collapse of the clump, which could be by maintaining turbulence by protostellar outflows (or massive star feedback of ionization, winds and radiation pressure in massive clusters) or by retaining magnetic support. One way to try and answer this question is to look at the star formation histories of individual star clusters, i.e., ask the question, "Is there evidence for internal age spreads?" Such work has been carried out in the Orion Nebula Cluster by *Da Rio et* *al.* (2014), who find that about 6 free-fall times have been needed to form 90% of the stellar population (as evaluated at the half-mass radius of the cluster).

What processes then set the initial mass function of stars and their multiplicity properties (especially the proportion of single and binary stars)? To answer such a question we will also need to better understand how protostellar disks can form in the face of apparently catastrophic magnetic braking (*Li et al.*, 2014) and the nature of protostellar outflows (e.g., disk winds and X-winds). Such issues relate to the long standing problems of how angular momentum and magnetic flux are shed from the gas that ends up becoming a star.

A general theory for star formation that could answer the above questions should also be able to predict how the outcome (e.g., IMF and multiplicity) would vary in different galactic and cosmic environments.

2.3 Can external influences affect star formation?

There are two primary modes in which clouds are assumed to interact with their surroundings. In the first case, which is thought to apply generally within the disks of normal, star-forming galaxies observed in the local universe, clouds are argued to be decoupled from their environments, given their highly supersonic internal gas motions (M$\sim$10 M_{Sun}) and large density contrast with the surrounding medium (implying an effective turbulent pressure $\sim$100 times that of the external intercloud ISM). The universality of cloud properties inferred by comparing the observed scaling relations of cloud populations in the solar neighborhood and nearby galaxies in the Local Group supports this view (e.g., *Bolatto et al.*, 2008); all clouds appear the same, independent of location. In this scenario, the process of star formation should be regulated primarily by internal factors, like turbulence and local feedback from star formation, although the environments of such clouds may nevertheless contain agents (shear, feedback) that influence their formation and evolution.

In the second case, which is thought to apply to clouds embedded in very gas rich environments (such as starbursts, galaxies at higher redshift, and galaxy centers), the frequency of cloud-cloud collisions is expected to be high and the surface density of the surrounding medium becomes comparable to the cloud surface density itself, raising the external pressure, thus coupling the cloud to its environment, and shortening the timescale for clouds to collapse. This has been argued to lead to a mode of star formation which is much more efficient than in the first case described above (*Krumholz, Dekel & McKee*, 2012).

These two scenarios (and two SF modes) are usually thought of as being distinct, but state-of-the art observations are beginning to suggest that they may more realistically represent two ends of a spectrum, even in normal star-forming galaxies. Now that cloud populations can be surveyed across a greater variety of environments, and in particular within proto-typical star-forming disks like the grand design spiral galaxy M51 (Figure 2.2), a much larger range in cloud properties is inferred than observed in the lower mass galaxies surveyed in the Local Group or in the solar neighborhood (e.g., *Hughes et al.*, 2013; *Colombo et al.*, 2014; *Leroy et al.*, 2015). The properties of clouds appear to vary with the galactic environment, challenging the standard "universal cloud" picture, and exhibiting characteristics of both scenarios. Indeed, many massive star-forming disks contain special features like spiral arms where locally very high gas densities are reached, thus providing environments more similar to those characteristic of the second scenario, in which clouds are tightly coupled to the surrounding medium.

Whereas especially strong influences would be required to alter cloud properties in a scenario in which clouds are completely decoupled from their surroundings, in a picture in which there is a spectrum of coupling, such influences need not be so extreme. Any process that affects the organization and structure of the ISM generally is a potential influence on clouds, their stability, and their ability to collapse and form stars. This includes changes in ISM density, pressure, and temperature that accompany passage through galactic spiral arms, or any generic non-axsiymmetric stellar structure. These common disk features represent perturbations to the underlying galaxy gravitational potential, and thus a change in gas dynamics, but they may also entail a change in ISM conditions due to, e.g., a locally modified interstellar radiation field (galactic structures are commonly characterized by enhanced numbers of old stars as well as concentrations of young energetic OB stars newly formed out of compressed gas) or an enhancement in the interactions (collisions) between clouds as the gas density is increased.

In this light, the assumption that molecular clouds are universal, self-gravitating objects may need to be relaxed and updated to allow for dynamically evolving structures subject to the local galactic gravitational potential (in addition to the potential due to just their own gravity). This naturally leads to a modified view of cloud stability and collapse, as the Jeans mass in the presence of an additional (bar or spiral) potential perturbation is changed (*Jog*, 2013), as is the ISM pressure, given the resulting complex pattern of gas flows (*Meidt et al.*, 2013). It also allows for cloud destruction via processes like star formation feedback, shear and/or other dynamical influences in a timescale which is shorter than the characteristic timescale for the cycling of molecular hydrogen (i.e., to the more diffuse HI and HII phases).

All of these influences should be recognizable by studying the mass spectrum of clouds, which contains the record of cloud formation and evolution, and by linking changes in cloud properties with local ISM conditions. Surveys of molecular gas at better than 40pc resolution from throughout a variety of galactic environments will be key to reliably capture cloud scales and confirm the influence of external factors on clouds and star formation.

2.4 Does the Initial Mass Function (IMF) vary? If yes, why? Where are the best places to look for this variation?

Knowledge of the stellar initial mass function (IMF) and its possible dependence on environment or star formation conditions is key for models of star and galaxy formation and evolution, as well as to our knowledge about the total stellar content of the universe. The IMF is often assumed to be universal, with a low-mass cutoff of $\sim$0.1 M_{Sun}, with some variations in the low and high mass cutoffs and the mass scale where the IMF may turn over. There is no undisputed observational evidence for deviations from a universal IMF. Simulations of Population III star formation currently predict many different IMFs. Empirically, probing the low end of the IMF requires modeling observations across all environments where there is active star formation, and is therefore challenging. In this section we address the question of IMF variation and its implications.

The most reliable method for measuring the IMF is direct star counts. This can be accomplished in the field for stars that have parallax measurements (*Kroupa, Tout, & Gilmore*, 1993). It can also be accomplished for stellar populations at a known, uniform distance, such as globular clusters (*Paresce & de Marchi*, 2000). In external galaxies, the IMF may be measured spectroscopically. One technique is to model the integrated light of absorption lines sensitive to surface gravity (*Conroy & van Dokkum*, 2012). Another technique is to model the kinematics of spatially unresolved stars (ATLAS^{3D}, *Cappellari et al.*, 2012). Nearly all studies of direct star counts in and around the Milky Way have found a universal IMF (*Bastian, Covey, & Meyer*, 2010), but many extragalactic studies have found an IMF that depends strongly on the mass, mass-to-light ratio, or star formation history of the galaxy. However, the conclusions of the extragalactic studies are more controversial than the local results.

Much of the controversy surrounding the extragalactic results revolves around the difficulty in interpreting integrated light spectra. Gravity-sensitive stellar absorption lines are also sensitive to chemical abundance. For example, the Na doublet at 8190 Å is much stronger in dwarfs than in giants. A strong Na doublet might indicate that the IMF is bottom-heavy. On the other hand, it might instead indicate a stronger-than-expected Na abundance. So far, this technique is promising, but more sophisticated models are necessary to eliminate some of the potentially plausible explanations for line strengths other than the IMF. Dynamical interpretations of integrated light spectra, such as those from ATLAS^{3D}, have also been subject to scrutiny. *Clauwens, Shaye, & Franx* (2015) have alleged that selection biases and systematic errors masquerade as dynamically observed variations in the IMF.

In order to find any evidence for a change in the IMF, we suggest measuring it in stars that formed in extreme conditions. A globular cluster is one such place. The first generation stars in clusters formed with "normal" abundances, but the second generation formed out of the AGB (asymptotic giant branch) ejecta from the first generation. As a result, the chemical abundance pattern of the second generation is highly unusual compared to stars formed anywhere else. In particular, the second generation stars have lower oxygen, higher sodium, lower magnesium, and higher aluminum than the first generation. The changes in cooling and opacity might affect the fragmentation of the second generation gas cloud and hence affect the IMF. We propose HST observations of a globular cluster in Strömgren filters, which can split the first and second generations on the main sequence. Deep enough observations can reach down to 0.1 M_{Sun}, permitting a direct-star-count measurement of the IMF in both generations simultaneously. This is a simple, imminently possible experiment to find indisputable evidence of a non-universal IMF.

2.5 What are the physical mechanisms that generate galaxy outflows? Are local examples similar to high-z? How well do models match the data in local and high-z galaxies?

Superwinds are known to exist in local starburst galaxies ranging in bolometric power from below 10^{11} to well over 10^{12} L_{Sun} (e.g., *Heckman et al.*, 1990; *Lehnert & Heckman*, 1996; *Heckman et al.*, 2000; *Rupke et al.*, 2005; *Veilleux et al.*, 2005). *Heckman* (2002) demonstrated that such winds are present in any galaxy with a star-formation surface density that exceeds $\sim$0.1 M_{Sun} yr^{-1} kpc^{-2}. Local superwinds are thought to be driven by the combined action of multiple supernovae, which heat a large volume of gas in the central starburst to high temperatures (10^6–10^7 K) where cooling is inefficient, powering an outflow which escapes along the largest pressure gradient—typically the minor axis of the galaxy. In addition to the starburst-powered, mechanically driven winds, radiation pressure and active galactic nuclei (AGN) are thought to also play an important role in some galaxies. Outflows also seem to be ubiquitous in normal star-forming galaxies at z$\sim$2–4, as evidenced by blueshifted interstellar rest-frame UV absorption features (e.g., *Shapley et al.*, 2003; *Steidel et al.*, 2010; *Shapley et al.*, 2011; *Jones et al.*, 2012).

While cosmological simulations including outflows and feedback have been quite successful in reproducing the bulk properties of ensembles of galaxies (e.g., star formation rate (SFR), total stellar mass vs. redshift), the physics of "feedback" is typically parameterized through simple semi-analytic prescriptions rather than first-principle physical laws. Therefore, all the physics of winds is encapsulated into prescriptions whose link to the underlying physical processes can be tenuous and ambiguous. As pointed out recently by *Heckman et al.*, (2015), it would be good to test whether or not more realistic models that employ some simple empirically-based scaling relations can reproduce the range in and correlations between the outflow velocities, outflow rates, and mass loading factors seen in local galaxies, and what these models might then imply for high-z galaxies where the gas fractions, dynamics of the disks, dust/gas ratios, etc., are very different.

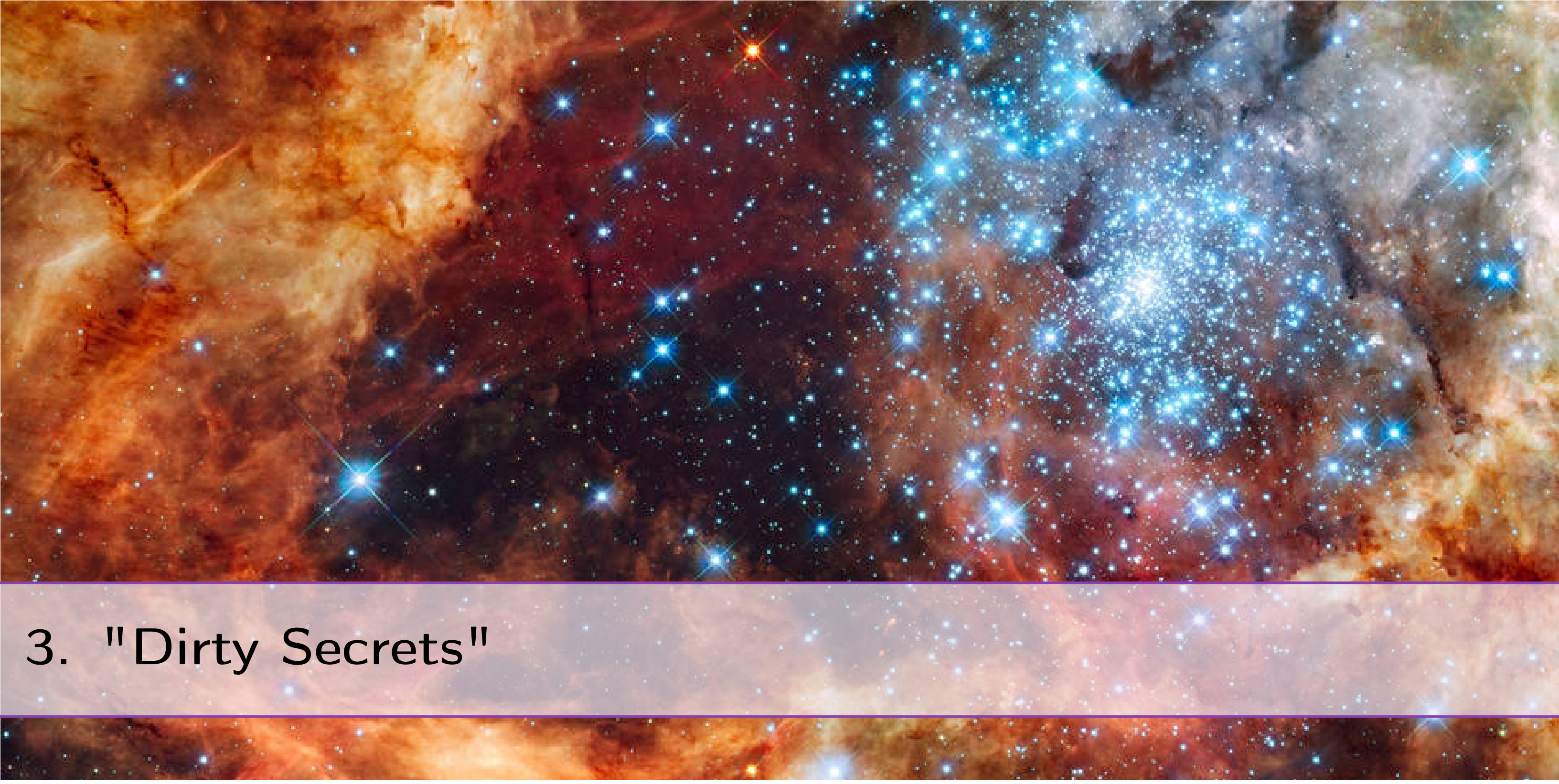
2.6 What influences inflows (gas accretion), and how it relates to SF? Do outflows impact the rate of gas accretion and the physical conditions of the accreted gas?

The "conventional wisdom"—indeed, almost orthodoxy—in cosmology and galaxy formation is that, despite our tremendous uncertainty in modeling star formation and galactic outflows, inflow onto galaxies (particularly at low masses and high redshifts) is well-understood. Because gravity acts top-down, we only need to follow large-scale structure formation, gravity, and radiative cooling to accurately predict inflow rates. In other words, the inflow rate from a "feedback-free" simulation is the same as in one with feedback, in most models. In galaxies with halo masses $<10^{12}$ M_{Sun} (where under feedback-free conditions, the cooling time is shorter than the dynamical time), this leads to the simple, convenient prediction that the accretion rate onto the galaxy is simply the (well-modeled) growth rate of the dark matter halo times the universal baryon fraction.

However, this is not necessarily the case, and this theoretical bias may severely limit our understanding of some observations. Taken at face value, such high accretion rates lead to great difficulty in understanding the masses of low-mass galaxies: a dwarf with halo mass $\sim10^{10}$ M_{Sun} would have to "blow out" $\sim$1000x as much mass as it turns into stars, assuming the universal baryon fraction reaches the galaxy. This has proven essentially impossible in most theoretical models, without invoking unphysical energetics from supernovae and other sources of stellar feedback. Moreover, even if it were possible (assuming every unit mass accreted simply "throws out" the appropriate mass, tuned to give the "correct" final mass), this leads to contradictions with the observed stellar mass-metallicity relation.

These observational tensions are alleviated if the inflows onto galaxies in the real universe are not as efficient as commonly supposed. Physically, this could happen for a variety of reasons. Photoionization suppresses accretion onto the smallest halos by raising the intergalactic medium (IGM) temperature to 10^4 degrees (circular velocity <15 km/s). Later photoionization of helium, or a harder local photo-ionizing field (from nearby sources) could raise this to $\sim$3–5 $\cdot$ 10^4 degrees, which could lead to halos as massive as 10^{10} M_{Sun} being affected. More dramatically (and at all halo masses), outflow could directly interact with inflows. There are various possibilities: direct radiation pressure from the galaxy could "hold up" or arrest inflows around the virial radius (if the stellar luminosity is comparable to the Eddington limit for optically thin, dusty gas at the virial radius). Outflows in cool or cold gas could accomplish the same effect. Essentially, the spherically symmetric case is similar in either situation; it's just a question of the outward momentum flux ($dP/dt = L/c$ for single-scattering radiation pressure, or $dP/dt = \dot{M} \cdot v_{wind}$ for a cold outflowing shell) balancing the force exerted by gravity at that radius ($G \cdot M_{total} \times M_{accreting}/R^2$). Hot outflows, or shocked outflowing material, could reach high temperatures and establish a quasi-equilibrium "hot halo" (pressure-supported atmosphere of hot gas in the dark matter halo). It is commonly assumed that this only occurs in massive ($>10^{12}$ M_{Sun}) halos; but that supposes the only source of shocks is virial shocks. In fact, any shock which can raise the internal gas temperature of the halo to $\sim 10^6$ degrees (any shock with velocities $>\sim$200 km/s) formally sets up a "hot halo" (although this can later be "overwhelmed" if enough cool material is accreted). This could then dramatically suppress inflows. Large-scale magnetic fields may also play a role, although they are generally predicted to be sub-dominant. Theoretical models increasingly include the resolution and physics needed to ask whether or not these processes occur, and whether they significantly alter the predictions for inflow. The question simply has not been well-studied, owing to the dominance of the assumption that inflow is understood. However, it is clear that the conclusion from any simulation will depend on the (uncertain) details of feedback.

Therefore, observational probes of feedback are needed. One possibility would be imaging of key spectral lines optimized for low surface brightness emission, presumably with 1-3 arcsec pixels to detect this very extended outflow and halo gas.



3. "Dirty Secrets"

Since the goal of this workshop is to bring together observations and theory, it is a good idea for theorists to understand where observers often hide factors of a few or more, and vice-versa. This is an attempt to list the most obvious culprits on both sides.

Observers

- Accuracy of measures of the SFR in different types of galaxies and environment: extinction, built in assumptions about the IMF, age, etc.
- Accuracy of measures of the molecular and atomic gas mass in different types of galaxies and environments. Scale factors between flux and mass under varying physical conditions.
- "Complete, representative" samples: hidden selection effects in small and even large samples. This is obvious, and one that authors try and address most of the time, but it is still always good to understand how observational limitations affect interpretations.
- Measurements of the "total" flux, "total" mass, etc.: just your basic aperture effects across wavelengths. This is a big one for comparing local galaxies to high-z samples, less so for models as long as care is taken when comparing properties extracted from the models to real data.

Theory

- Some cooling functions assume fixed equilibrium chemistry (doesn't evolve in time, although non-equilibrium chemistry module does exist)
- Some cooling functions work down to only certain temperature
- Cooling turned off in some crude stellar feedback recipes
- We pick only the nicest images and movies to show to others/observers

4. Proposed Paths for Advancement

After identifying the most outstanding questions in the field of star formation and galaxy evolution, the group focused on the steps that should be taken to answer these questions. These paths for advancement include resolved surveys of tracers of the different ISM phases in nearby clouds and external galaxies that could help the theoretical interpretation of unresolved observations, the development of new computational tools for accurately modeling the influence of radiation in interstellar gas, the development of a better diagnostic of metallicity in high redshift galaxies, and a better comparison between models and observations of outflows.

4.1 Characterize the internal structure of local clouds and calibrate integrated diagnostics for extragalactic use

Characterizing the internal structure of GMCs in external galaxies (D>1 Mpc) by directly resolving the high density structures within them using dense gas tracers (e.g., HCN) is currently unfeasible and will be unfeasible for decades to come. Resolving the 0.5 pc scales necessary to do so requires interferometric observations with very long baselines ($\sim$10 km) which translates into small filling factors of the UV plane and therefore bad surface brightness sensitivity. Given this limitation, the most promising way of exploring the internal structure of extragalactic GMCs is through low spatial resolution (10–100 pc) spectroscopic studies (Figure 4.1).

The velocity profiles of emission lines tracing gas above different critical densities and the relative intensities of these transitions (line ratios) encode information regarding the internal density and kinematical structure of clouds. Extracting and interpreting this information requires a mapping between these integrated spectroscopic observables and the distributions that describe them, as well as quantification of the internal structure of the clouds (e.g., the internal gas density and

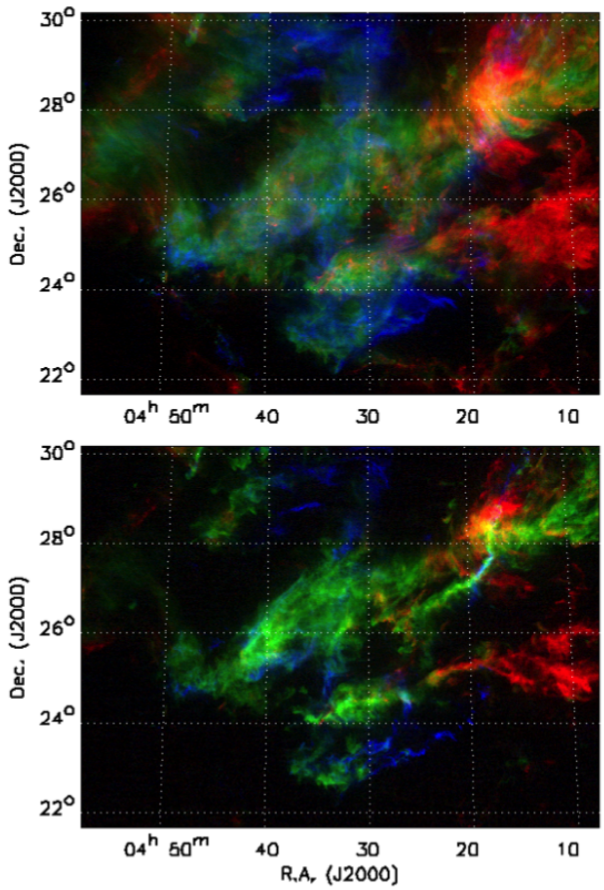


Figure 4.1: The FCRAO survey of the Taurus Molecular cloud (*Goldsmith et al.*, 2008) in CO and ^{13}CO is an example of the wealth of information that can be provided by a detailed survey of nearby molecular clouds. Detailed surveys of nearby clouds in many different ISM tracers will be used to characterize clouds in external galaxies where the level of detail archived in the Milky Way will be hard to achieve.

velocity PDFs). Such "mapping" can be obtained via two different avenues. First, numerical and analytical theoretical models of GMCs can in principle predict how these integrated observables change as a function of the internal properties of the clouds. Second, these observables can be calibrated empirically using observations of local GMCs in the Milky Way.

The theoretical approach can be informative and useful for identifying the connections between the microphysics of gas and the large scale observable properties of clouds, but the systematic uncertainties associated with sub-grid physics, post-processing radiative transfer, and physical processes not included in the models (e.g., magnetic fields) would make it difficult for theory to produce quantitative predictions. Empirically calibrating these large scale observables using nearby Galactic regions appears as the most promising way to move forward in this regard. We propose a coordinated observational effort aimed at calibrating and applying these spectroscopic observables in order to quantify the internal structure of GMCs in the ISM of external galaxies. Such an effort should push forward three specific observational programs:

- High spatial resolution observations of the tracers necessary to determine the distributions of interest that characterize the internal structure of a sample of nearby GMCs. This includes determining the density PDF, gas velocity distribution, thermal structure, presence and intensity of shocks, and star formation activity.
- Low spatial resolution spectroscopic observations of the same sample of nearby GMCs to emulate the equivalent observations that could be conducted for these objects in external galaxies and at high-z.
- Spectroscopic observations of GMCs in a limited sample of nearby galaxies at $\sim$50 pc resolution, at $\sim$1 kpc resolution in a large sample of galaxies in the local universe, and integrated measurements for distant high-z galaxies. By applying the local calibrations to these data we can explore how the internal structure of clouds depends on the environment and other parameters like metallicity, level of ISM turbulence on galactic scales, etc.

Caveats, decisions, and things to keep in mind:

- It is not clear whether this program would benefit from using as much existing data as possible while re-processing it uniformly, or if it would be better to take all the data for this program in a uniform manner. It would make sense to build on datasets that are already available, but the ability to overcome potential sources of systematic error associated with mixing datasets from different programs must be evaluated.
- It is important to keep in mind that locally calibrated relations between observables and the internal properties of GMCs might not be applicable in a straightforward manner to clouds in other radiative, chemical and dynamical environments. It is important to understand the limitations of these calibrations, and how we can correct or modify them so we can use them over a wide range of environments. This is an aspect in which theory can provide important insights.
- It would be interesting to adopt a "ladder" scheme in which line ratios and velocity profiles can be used for moderately distant sources, but line ratios alone for more distant sources, for which the S/N will not be sufficient to measure the velocity profiles.
- The calibrations must be done at different spatial resolutions and levels of sensitivity. It would be very useful to create a software tool that could allow anyone to access this legacy dataset and redo the calibration given the sensitivity and resolution of their data.

4.2 Make 1pc-scale, 1 km/s map of extragalactic GMC

Detailed studies of clouds in the MW will supply an indispensable baseline for linking internal cloud structure to conditions in the local environment. But the view that can be constructed with the multi-scale, multi-wavelength observations described above will likely be optimized for only a subset of clouds and thus remain rather limited to probing a narrow range in physical conditions; distance ambiguities and the challenges of linking objects across wavelengths complicate assembly of the relevant information throughout a more varied cloud population.

It is thus imperative that we also continue to build up an extragalactic view of clouds that allows us to sample a variety of environments. This ideally pushes at the overlap of Galactic and existing extragalactic studies, namely on scales of a few parsecs. With this resolution, clouds can be reliably resolved, allowing for disambiguation of massive clouds into smaller objects and providing a view of internal cloud kinematics, both of which are, so far, rarely achieved in the extragalactic context but essential for constraining the processes responsible for cloud formation, evolution and destruction and the scales over which they act.

This optimized view will provide the leverage to resolve lingering issues about the nature of star formation and its relation to the properties of GMCs. Specifically, a fundamental but unanswered question remains: is star formation "bimodal"? By this we mean to ask, "Are there multiple physical processes that regulate the properties of GMCs and their ability to collapse to form stars and star clusters? These different physical processes would appear as greatly increased rates of conversion of gas into stars per unit mass of gas or as variations in the resulting stellar IMF. The former is clearly apparent from the varying ratios of $<M(H_2)$ / SFR$>$ = timescale for cycling dense interstellar matter into stars. In the Milky Way, this timescale is $\sim$1 billion years; in the ultraluminous infrared galaxies, it is $<$100 million years. A similar variation would be evident when comparing the H_2 gas in the inter-arm areas and spiral arms of normal galaxies.

To really understand these variations, a comprehensive multi-wavelength imaging survey of a nearby spiral galaxy could become the Rosetta stone for understanding star formation processes (birth and feedback), the interplay of the interstellar gas with galactic structure, and the coupling of clouds to their local environment. This legacy project would be a resource for detailed comparison with theoretical simulations of ISM evolution in galactic disks.

The primary observations would be imaging with ALMA at 0.02–0.1 arcsec resolution of a selected grand design spiral galaxy at $\sim$10 Mpc distance, yielding physical scales of 1–5 pc. Ideally, the galaxy should be selected in an equatorial area to enable imaging with both Southern and Northern telescopes, such as ALMA and E-VLA, and have prior coverage with far-infrared facilities such as Spitzer and Herschel. The ALMA component should include both low density molecular tracers (CO and ^{13}CO 1-0 and 2-1) and high excitation probes (e.g., HCN) and the Rayleigh-Jeans dust continuum. A kinematic resolution of $\sim$1 km/s will be key to resolving streaming motions associated with galactic shear, spiral arm perturbations, cloud-cloud collisions and star formation feedback processes (i.e., radiation pressure, HII region expansion and ionization), all of which are potential sources for star formation "bimodality" influencing cloud formation and evolution. The required high spatial and kinematic resolution and high signal to noise ratios will necessitate very large program status with ALMA—perhaps 200–400 hours—but the utility of this comprehensive Legacy project would far surpass that of lower quality partial datasets. This Legacy project will also attract the highest quality multi-wavelength observations from other facilities, e.g., optical and near infrared imaging from HST and eventually JWST in both the continuum and key spectral lines. Particularly important would be the optical emission lines characterizing the ionized gas and the near infrared lines of H_2 probing shock excited H_2 at 2000K.

A vital component of this public Legacy project would be an associated suite of high resolution numerical hydrodynamic simulations including accurate thermodynamics and radiative transfer for continuum and lines and stellar feedback. In terms of simulations of Photon Dominated Regions (PDRs) on pc-scales, much work has gone into 1D astrochemical codes and techniques, but there have been only few attempts in 3D (*Röllig et al.*, 2007; *Glover et al.*, 2010; *Bisbas et al.*, 2012). A crucial advancement to be made is therefore a code that simultaneously treats astrochemistry and radiation in ionized, PDR, and molecular zones of the ISM.

4.3 Develop on-the-fly radiative transfer algorithms for simulation

Stars and galaxies radiate in a wide range of the electromagnetic spectrum. The radiation affects their surroundings and changes their own paths of evolution. Understanding the effects of radiation is thus of vital importance in virtually every corner of astrophysical phenomena, including formation of stars and galaxies, transformation of stellar atmosphere, ionization of interstellar and intergalactic gas, and reionization of the universe. At the end of the day, radiation is why we can "observe" many of the astronomical objects in the universe. While tracing radiation is important in many areas of astronomy and has been extensively studied analytically, treating it in a four-dimensional spacetime in a reliable fashion is anything but trivial. One also needs to consider hydrodynamic and thermal responses to the radiation field, which complicates the problem even further. Only in the past decade or so have researchers finally made headway in developing multi-dimensional radiation hydrodynamics algorithms in realistic astrophysical applications.

We find that the need and opportunity have never been greater for a comprehensive radiative transfer physics algorithm that can be easily incorporated into an existing hydrodynamics simulation code. Many disciplines and communities within astrophysics will greatly benefit from this endeavor. This is especially true now when there is a plethora of opportunities to compare realistic radiation hydrodynamics simulations with high-resolution observations, and as the parallel supercomputing resources ever grow in computational powers.

Therefore, we propose that we put together a community effort to develop an adaptive, on-the-fly, massively-parallelized radiative transfer module that can track continuum emission, line emissions, and momentum transfer. The module will consider both point sources and diffuse emission, and will need to adeptly switch between optically thick and thin regimes. We recommend that such a radiative transfer module be framed to harness the recent advancements in processor architectures, such as nVidia's Graphics Processing Unit (GPU) or Intel's Many Integrated Core (MIC), by making the radiative transfer calculation efficiently subcycle on these cores.

We also suggest that, for the stated effort, we actively seek and utilize existing radiative transfer or ray tracing technologies in other fields of studies (e.g., the video game industry or the film industry). For this purpose, we advocate that we organize joint efforts and workshops, so that experts from diverse backgrounds and experiences (astronomers, physicists, mathematicians, computational engineers, visualization experts, etc.) can bring together fresh ideas for developing an advanced radiative transfer algorithm for astrophysical uses.

4.4 Develop stellar metallicity diagnostics for high-z systems

The overall chemical abundances of a galaxy, as well as the full distribution of metallicities of its stars and gas, contain important information about the history of star formation and gas flows within that galaxy. While gas-phase metallicity tracers have been used for over a decade at high-z, stellar metallicity tracers are much less commonly explored. One of the principal reasons is that gas-phase metallicity tracers rely on bright emission lines, while many stellar features are seen in absorption and therefore require significantly higher S/N spectra.

Recently, the effectiveness of commonly used strong-line metallicity calibrations of gas phase metallicity such as N2 $\equiv$ [N II] λ6583/Hα, O3N2 $\equiv$ ([O III]λ5007/Hβ)/([N II] λ6584/Hα), and R23 $\equiv$ ([O II] + [O III])/Hβhas been called into question. New near-IR observations of the rest-optical emission lines from high-redshift star-forming regions have conclusively demonstrated that the abundance pattern and/or the ionization and excitation conditions within high-z star-forming galaxies are markedly different from those within low-z star-forming regions (*Masters et al.*, 2014; *Steidel et al.*, 2014; *Shapley et al.*, 2015). The difference in these emission lines—the same as are used to determine gas-phase metallicities at high redshift—suggests a new metallicity calibration at high-redshift may be required. Such a calibration necessitates independent constraints on the metallicity of high-redshift systems. One such check would be the stellar metallicity of the young stars (O-type) in these galaxies whose metallicities would presumably be similar to that of the gas from which they recently formed.

We are therefore interested in having stellar metallicity diagnostics both for young stars whose metallicity should be similar to that of the H II regions as well as older stellar populations (in quiescent systems where we can't measure a gas chemical abundance anyway).

For young stars, it may be possible to use UV photospheric absorption lines as well as the P-Cygni features from Wolf-Rayet stars in order to constrain the chemical abundance. Possible features are discussed in the context of rest-UV observations of the lensed galaxy MS 1512-cB58 in *Pettini et al.* (2000) and more recently for CASSOWARY 20 in *James et al.* (2014). Photospheric absorption lines S V 1502Å, O IV 1343Å, C III 1428Å may be of use. However, there is some evidence that these lines are not particularly sensitive to metallicity (*Rix et al.*, 2004). The stellar wind lines of C IV 1546Å, 1550Å and Si IV 1393Å, 1402Å (*James et al.*, 2014; *Leitherer et al.*, 2011) and a blend of photospheric Fe III lines between 1900-2050Å have been used with some success (*Rix et al.*, 2004; *James et al.*, 2015).

It is worthwhile to examine the usefulness of these lines as metallicity diagnostics in single local O stars before they are implemented at high-z. Spectra from the Cosmic Origins Spectrograph (COS) or the Space Telescope Imaging Spectrograph (STIS) of sufficient quality likely exist. Such lines are already detected in stacks of hundreds of high-z UV-color selected galaxies (*Shapley et al.*, 2003). In particular, using a subset of galaxies with constraints on the temperature of the ionizing spectrum from nebular spectroscopy (e.g., *Steidel et al.*, 2014) will be favored in order to mitigate the effect of the temperature on the depth of these lines (the age-metallicity degeneracy).

4.5 Compare outflow observations with simulations

Detailed measurements of gas in the circumgalactic medium (CGM) now exists both at low redshift (*Tumlinson et al.*, 2011; *Werk et al.*, 2013; *Johnson et al.*, 2015) and at high-z (*Rudie et al.*, 2012). In particular, high-redshift observations of the gas around galaxies known to have strong outflows provide a unique opportunity to study the physical properties of gas in winds and the degree of impact they are likely to have on their parent galaxy during the epoch when cosmic flows are most important. CGM observations, because they probe the physical properties of this gas, can yield specific insights into the physical mechanisms responsible for outflows, as different mechanisms should heat and accelerate the gas in different ways and may also have unique chemical signatures. While detailed observations of the column density, kinematics, covering fractions, and internal kinetic energy within the high-z CGM (all as a function of distance from galaxies) now exist (e.g., *Rudie et al.*, 2012), current comparisons with simulations use comparatively blunt tools and typically only compare to regions within 1–2 virial radii (*Faucher-Giguère & Kereš*, 2011; *Shen et al.*, 2012; *Fumagalli et al.*, 2014; *Faucher-Giguère et al.*, 2014; *Suresh et al.*, 2015), despite evidence that the sphere of influence of individual galaxies extends to larger distances. More specific comparisons should be done and will likely result in important advances in our understanding of the main drivers of feedback in high-redshift galaxies during the cosmic peak of star formation, one of the key missing pieces in our theory of galaxy formation and evolution.

5. Recommendations

After identifying the paths for advancement, the group discussed the tools needed to accomplish these proposed paths. These tools include better ways to share the results of cosmological simulations with the astronomical community, so that theory can be more readily tested against data, advancements in the computational tools used in simulations, new-generation telescopes that will allow us to obtain the required critical observations, and new ways to allow the scale of investigations required to significantly advance our understanding of the evolution of galaxies.

5.1 The Simulation Repository & Telescope: Enabling worldwide, direct, and detailed comparisons of observations and theory

Numerical simulations are an indispensable tool in all fields of theoretical astrophysics, and simulation data should be considered the theoretical equivalent of observational data—it is the only way we can "observe" a model universe. Despite this equivalency, facilities for archiving and providing access to simulation data lag far behind the equivalent facilities for observational data. Typically, simulation data are only shared by personal request, and even for a specific class of simulations (e.g., cosmological simulations), data formats vary greatly. Consequently, in most cases, simulation data are simply not used by astrophysicists outside of the immediate group that produced those simulations.

The status quo is problematic for multiple reasons. For example, comparisons among different simulations and with observations are hindered by reliance on manual sharing of data on a case-by-case basis. Moreover, even when best practices are followed (i.e., the specific source code version and parameters used are retained), simulations are inherently not reproducible because the results can vary depending on the architecture of the computer, number of cores used, etc. Consequently, simulation data must be retained to ensure the possibility of verifying and reproducing results. Furthermore, simulations are typically vastly under-utilized, despite their significant cost (in terms of person-hours, computational resources, and energy); a central repository would enable others to extract additional value from simulation data that would otherwise be discarded.

Some simulation databases that are intended to disseminate simulation results widely to the community in a user-friendly manner do exist. Examples in the field of galaxy formation include the Millennium Simulation SQL-queryable database (*Lemson and the Virgo Consortium*, 2006[1]), GalMer (*Chilingarian et al.*, 2010[2]), the Millennium Run Observatory (*Overzier et al.*, 2013[3]), the Theoretical Astrophysical Observatory (*Bernyk et al.*, 2014[4]), and the Illustris Galaxy Observatory (*Torrey et al.*, 2015[5]) and simulation database (*Nelson et al.*, 2015). However, existing databases are generally specific to the simulations of single collaborations, and they do not share a common data format or interface. They are often maintained by a single or a small number of astrophysicists, and because these individuals typically perform database maintenance and support in their "spare time" (i.e., they are faculty, postdocs, or students that have a range of responsibilities—not solely database support), support can be limited. If said individuals change institutions or leave the field, the database may not be maintained further.

Still, these examples as test cases have demonstrated the extraordinary increase in value that comes with making simulation data public and easily accessible to observers. The specific examples above have had a disproportionately large impact on the field, with thousands of citations in observational papers. Meanwhile, simulations of comparable resolution and expense that test alternative physical models and take about same amount of time to produce, but that are not made public have orders-of-magnitude smaller mean impact factors. Almost universally across astrophysics theory, groups recognize the value of such a database. However, the vast majority of groups simply lack the resources to permanently support their own "home grown" system. And even if such resources were provided by the same NSF and NASA grants that fund the simulations themselves, a proliferation of independent databases with different standards would soon render the results just as inaccessible to observers as the current reliance on personal, manual data-sharing.

To address the aforementioned disadvantages of current simulation repositories, we recommend the creation of a centralized archive and database for simulation data that is analogous to those that exist for observational data (e.g., the Infrared Science Archive (IRSA) and the NASA/IPAC Extragalactic Database (NED)). The archive would be curated by one or more dedicated support astronomers. Simulation data would be stored in common formats designed to match those of

[1] http://www.mpa-garching.mpg.de/millennium/
[2] http://galmer.obspm.fr/
[3] http://galformod.mpa-garching.mpg.de/mrobs/
[4] https://tao.asvo.org.au/tao/
[5] http://www.illustris-project.org/galaxy_obs/

existing observational data, and the simulation parameters would be required to be included as metadata when uploading simulation data to the archive. The type of data stored would depend on the simulation. For example, semi-analytical models of galaxy formation might include galaxy catalogs in which each object has a position, various masses, star formation rate, magnitude in various bands, etc. Cosmological simulations could include not only mock catalogs but also raw particle or cell data such that the details of individual objects (e.g., three-dimensional mass profiles) could be analyzed. If a simulation provides mock observables such as images and spectral energy distributions, these would also be included. Crucially, data in the simulation repository would be linked with the relevant observational database(s), thereby facilitating comparisons between observations and simulations. Tools to facilitate such comparisons would also be included. For example, one such tool could take an observed galaxy spectrum as input and then return the best match from all simulated galaxy spectra stored in the database. A huge wealth of tools would be immediately available to search, categorize, and compare the simulation data, because these tools have already been developed for observations. By simply converting the simulated data into mock observables, the observers are able to perform any operation on it that they would on real observations. Additional functionality would be added in response to community interest.

We propose starting with a pilot project that would be limited to a small set of simulations, including large-volume cosmological simulations, zoom-in simulations, and semi-analytical models (very different theoretical data sets, but all critical for modeling galaxy and star formation). Initially, a single representative dataset of each type of simulation will be incorporated into the existing observational archives. The pilot will involve developing a uniform data format and functionality for both uploading and accessing simulation data, and a set of simple scripts for conversion between the relevant formats. How the simulation data can be linked with observational archives will be clearly demonstrated. Tools to do this (casting the simulation data in "observational units," mock radiative transfer, and more) have already been developed; they simply need to be applied. Once the pilot archive is functional, its utility will be demonstrated via a few test cases in which an observer, specifically one who was not previously familiar with the simulations, uses the archive to perform analyses of the simulation data (treating it like real observations) and comparisons with observations. Successful completion of this pilot project will facilitate securing funding for the full-scale archive, demonstrating the viability of this project and its utility.

The potential value and impact of the proposed simulation archive cannot be overstated. At the most basic level, it would serve as a long-term archive for simulation data, thereby ensuring that simulation-based studies are reproducible. Because code parameters would be included along with the data, transparency would be drastically increased. Moreover, the simulations made available via the archive would likely have a much greater impact, in terms of both usage and citations, than they would otherwise. For example, the Millennium Simulation (*Springel et al.*, 2005[6]) has

[6] http://www.mpa-garching.mpg.de/millennium/

been used for at least 711 papers and continues to be used (despite being since superseded by many simulations), and the paper has been cited over 2100 times. As has been the case with the Millennium Simulation, via the proposed archive, theorists from small institutions with modest computational resources will be able to analyze state-of-the-art simulations that would otherwise be inaccessible to them. This is simply not possible in the status quo.

The archive would also be of great value for observers, who could use it to make direct comparisons with simulations without relying on "in-house" theorists. Observers could make comparisons in terms of both raw physical quantities (e.g., stellar masses of galaxies) and direct observables predicted from the simulations, such as mock images (e.g., *Torrey et al.*, 2015) and spectral energy distributions (e.g., *Lanz et al.*, 2014) of galaxies. Moreover, because of the uniform data format, there may be visualization synergies, i.e., visualization tools developed for the simulation data, such as yt (*Turk et al.*, 2011[7]), could be applied to observational data to fly through datasets, make movies, etc.

The public would also benefit considerably. The archive would enable NASA, the NSF, and other bodies to require the recipients of large CPU time allocations (i.e., tens to hundreds of millions of CPU hours) and grants to make their results publicly available in a useable format. It is already a requirement of simulation-centric funding proposals that some data storage plan be provided; unfortunately in theory, the data storage is often only accessible by the model authors. Thus, the modest investment necessary to create and support the proposed archive would significantly increase the return-on-investment of funds and CPU time awarded to theorists. Moreover, the archive would have significant outreach potential. For example, data from the archive could be incorporated into university courses, used for high school science projects, and explored by amateur astronomers.

5.2 Dedicated, optimized facilities for high-dynamic-range models

The current trend in supercomputer development is increasing the number of CPU and GPU cores while providing less memory per core. This scheme is very beneficial to easily parallelizable problems; however, star formation physics necessarily includes processes acting over a large dynamic range that are mutually affecting one another. This has two major consequences:

- The problem is inherently resistant to parallelization. Because large spatial/time/mass scales in the problem are causally influenced by structures which represent only a tiny, small-scale fraction of the problem (e.g., a single stellar explosion), they must await their calculation before "going forward," and performance cannot scale efficiently with indefinitely more cores.

[7] http://yt-project.org/

- A large amount of memory is required to implement more detailed physics. This is not just a resolution issue; more detailed subgrid physics models require significantly more memory, as additional physical equations and properties must be followed.

We have to emphasize that star formation is just one of myriad scientific problems with huge dynamical ranges that are adversely affected by this trend. Condensed matter, geophysics, weather science, turbulence, and many other simulations all suffer from the same problems. There are also commercial and national security applications with similar issues (e.g., global image analysis, data mining, etc.).

To address this issue, what the field deeply needs is a computing facility with fewer but faster cores (since parallelization options are limited) and more memory per core (current average is 1–2 GB/core, whereas the optimal number for scientific purposes for most of the above applications would be 4–8 GB/core). The memory issue is serious even with current supercomputers, as numerical simulations often have to request several times the actually required number of cores just to get enough memory to run. If current hardware trends continue, this gap will only grow. It will soon become extremely inefficient to run simulations of such problems (requiring an order-of-magnitude more cores for memory reasons as can actually be used), holding back research in these fields and limiting the feasibility of computing time requests.

As a possible solution, we propose the creation of computing facilities dedicated to large-dynamic-range problems. We want to emphasize that these facilities would be used not only by the star formation community, but by a wide range of other fields. Such a facility would have a tremendous impact as it would be the only device available to run efficient simulations for large-dynamic-range topics. Some computing facilities include limited "large-memory" nodes, but these are token efforts or highly-specialized setups for problems which require terabytes of memory per core, and as such represent a tiny fraction (much less than a percent, in almost all cases) of the computing power of the machine. But the concept could easily be expanded to the scales needed to enable transformative new science.

We think it is important to recognize how inefficient modern supercomputers are at such problems, so we propose the creation of a new, separate metric which—unlike the current ones—does not assume infinitely parallelizable problems. Such a metric would clearly show which computers are best suited to tackle large-dynamic-range problems.

5.3 Large UV-optical space telescope

High mass stars have an outsized role in star formation and galaxy evolution. Wherever star formation occurs, high mass stars ($M>15\ M_{Sun}$) dominate the luminosity and chemical output. In their catastrophic deaths as supernovae, they feedback energy into the ISM and can drive spectacular winds. This "feedback" has become a key component of current models, which has been claimed to solve a number of outstanding problems in galaxy evolution, such as the inability of simulations to produce bulge-less disk galaxies and dark matter halo profiles with isothermal cores as observed in nature. It is thus critical to understand the processes that affect high mass formation and evolution.

Key elements in the observational attack on the problem of massive star formation are the imaging and spectroscopic observations of the resolved stellar population—studies and counts of individual stars. We must understand the mass distributions of their initial formation, as well as the fraction of stars that form in binary or multiple systems, which may alter their evolutionary paths. We must also understand how the high mass stellar IMF and multiplicity may be affected by gravitational stability, pressure, and metallicity of the larger scale environments in which natal molecular clouds form. Driving the attack is our quest to identify the physics that describe the conditions leading to high mass star formation, and can characterize the stars that result from that process.

High mass stars are hot, and their spectral energy distributions peak in the vacuum ultraviolet. Thus their characterization requires observations at those wavelengths. Observations of individual stars also require exquisite resolution at those same wavelengths. HST has enabled observations of resolved high mass stars within low-density environments to $\sim$10 Mpc of the Milky Way, but cannot access the more extreme environments found only in galaxies at larger distances. It is these environments which may more closely resemble conditions in the high redshift universe, and allow us to test the dependence of star formation over a large range of gravitational stabilities, pressures, and metallicities.

A large space-based observatory ($\sim$12 m) with UV capabilities would not only enable access to such systems and support the next generation of studies of massive star formation; it would also enable investigation of a full range of crucial questions of cosmic origins, particularly if optical and NIR capabilities are also available.

5.4 Large far-IR space telescope

The optical and infrared regions of the electromagnetic spectrum represent roughly equal quantities in the energy output of the universe. Optical wavelengths are dominated by starlight, whereas the infrared is largely reprocessed emission from dust-enshrouded star forming regions and active galactic nuclei. At higher redshifts, galaxies are expected to be increasingly gaseous and dusty. Therefore a complete picture of the mass and energy budget of galaxies with redshift requires spatially-matched observations at infrared and optical wavelengths. Although we have an enormous amount of data at sub-arcsecond resolution in the optical and increasingly in the near-IR (with the Wide Field Camera 3 (WFC3) on HST and soon with JWST), far-infrared data at high spatial

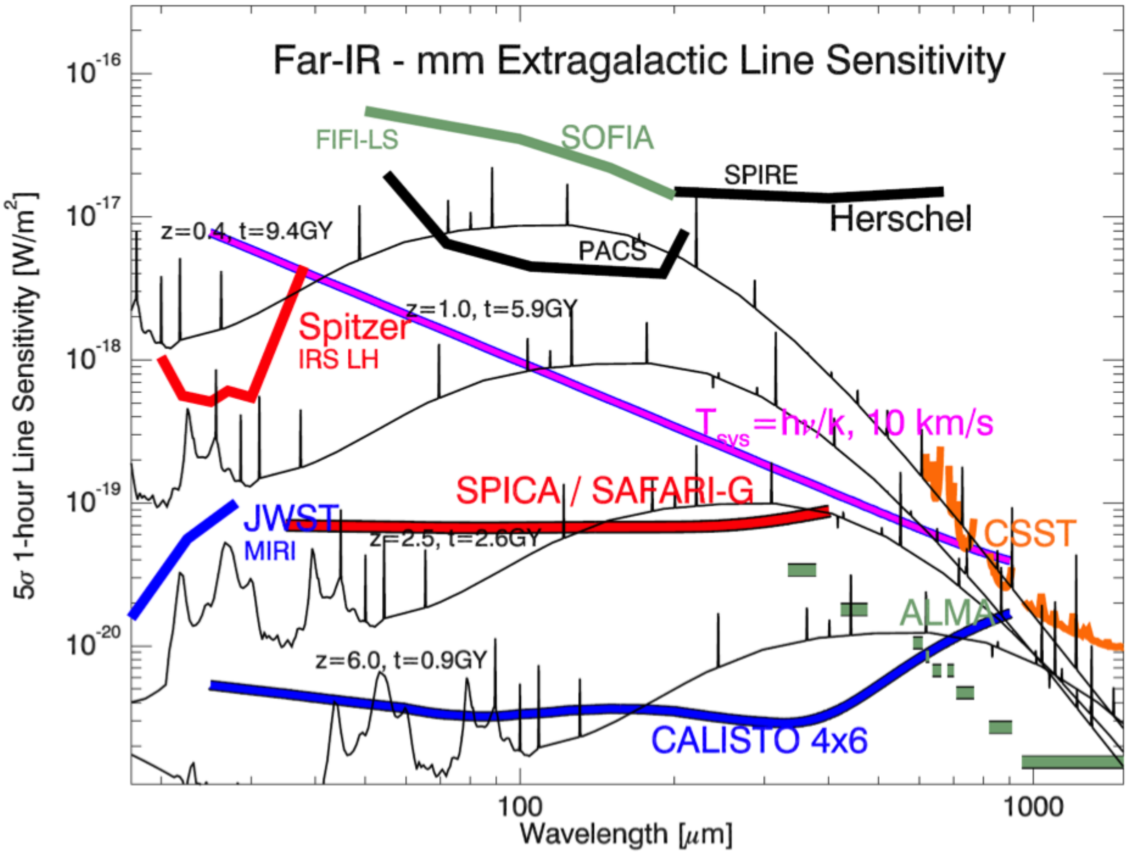


Figure 5.1: Sensitivity estimation for line emission for different observatories covering the far-infrared spectral regime. Far-infrared mission concepts like the Cryogenic Aperture Large Infrared Space Telescope Observatory (CALISTO) will allow us survey spectral lines in a large number of high-redshift galaxies. Courtesy of Matt Bradford.

resolution are absent (although ALMA offers a new and unique window in the millimeter and submillimeter range). The peak of the SED in the IR is between 100-300 microns—a wavelength range recently explored by the Herschel Space Observatory, but unavailable from the ground due to the lack of atmospheric transmission.

Herschel made transformative observations in the far-infrared/submillimeter range between ∼50 and 500 microns wavelength, but with its 3.5-m aperture, the angular resolution was in the range ∼8"–20". To reach sub-arcsecond resolution, compare the dust-enshrouded star formation with unobscured star formation activity, and better understand star formation activity on ∼kpc scales in galaxies of various types, we need a large far-IR space telescope. The recently completed NASA Roadmap also advocates for a large Far-IR Surveyor which would, in addition to imaging the dust continuum, map important ISM cooling lines such as OIII, NII and CII, as well as offer the promise of tracing water and other important molecules. ALMA can only do this at high redshift (z>2) but if we want to explore the detailed assembly and evolution of galaxies over the last 10 billion of cosmic time, we need a large far-IR telescope in space (Figure 5.1).

A 20-m diameter telescope at 100 microns wavelength would offer a beam size of 2.5", and a collecting area more than 30x greater than that of Herschel. For observations of the local universe, moderately high spectral resolution is required. The development of heterodyne focal plane arrays as well as of direct detectors means that the specifications of the next-generation far-infrared space telescope need to be worked out to optimize its ability to deliver critical information about star formation and the interstellar medium of galaxies. An important consideration is to what extent the telescope optics need to be cooled, which enhances sensitivity for low/moderate spectral resolution observations, but which is not of significant benefit for high spectral resolution (heterodyne) observations. Various missions of this type have been proposed, including the Japanese Space Infrared Telescope for Cosmology and Astrophysics (SPICA), the Russian Millimetron, and the US Single Aperture Far Infrared/Cryogenic Aperture Large Infrared Space Telescope Observatory (SAFIR/CALISTO) missions. Collaboration is clearly an important aspect for advancing this field of astronomical research.

5.5 Large mm/sub-mm surveyor

In recent years, great strides have been made in fields including galaxy evolution, structure formation, and star formation from large extragalactic surveys at optical and near-IR wavelengths, such as the Sloan Digital Sky Survey (SDSS), the UKIRT Infrared Deep Sy Survey (UKIDSS), etc. The exceptional success of these surveys has only been possible thanks to their coverage of galaxies with a wide range of stellar masses, redshifts, star-formation rates and environments. However, they all only trace stellar emission, i.e., the result of star formation. There are currently no equivalent large surveys of the drivers of star-formation, i.e., the gas and dust clouds in which young stars are born.

In fact, in order to understand star formation on galactic scales and to interpret small (GMC) scale star formation in the broader context of the buildup of stellar mass and galaxies in the universe, we need a similar understanding of the gas processes in galaxies. Such an understanding can only be attained by undertaking a census of the gas in galaxies. Furthermore, stars form in optically-thick dust-enshrouded clouds, resulting in their direct emission being absorbed by the dust and reemitted in the infrared. Therefore, to trace the embedded star-formation requires observation of the cold dust in which it is embedded.

Direct observations of the gas and dust in galaxies require observations of far-IR emission lines (e.g., [CII], [OI], [NII], etc.) and far-IR continuum. However, such observations have so far been limited to the brightest, most extreme systems, predominantly due to the technologically driven limitations of current facilities. In particular, the community now requires direct detections of the gas and dust in galaxies with a range of masses and ongoing star-formation rates from the local universe out to high redshifts. Such data can only be realized with a dedicated survey facility with a large field of view and high sensitivity.

Therefore, we now recommend the implementation of a wide-field sub-millimeter/millimeter survey facility, which will be able to survey a significant fraction of the sky[8] in the sub-millimeter/millimeter (i.e., cold dust) continuum at high resolution (∼1–5"), allowing direct matching with counterparts at other wavelengths. In addition to continuum instruments, the observatory should also be equipped with spectroscopic capabilities at these wavelengths, to survey cold gas tracers in large samples of galaxies across redshift, stellar mass, star-formation rate and environment. The data collected by such a facility would be directly analogous to those already collected by wide-field optical surveys of the stellar emission from galaxies (e.g., SDSS).

In order to achieve ∼arcsecond resolution at the shortest sub-millimeter bands available from the ground, a ∼50-m telescope with reasonable performance at ∼400 microns is required. A particular large sub-millimeter telescope concept—the Cerro Chajnantor Atacama Telescope (CCAT)—has been studied extensively in the past few years, although the project has a poor outlook at present. Technically, a 25-m telescope with high 350 micron efficiency having cameras with requisite ∼50,000 detectors is considered challenging but feasible for construction in the present decade. Advanced concepts for multi-object spectroscopy are in active development. A far-IR survey facility such as that proposed here may be feasible in the next decade. At millimeter wavelengths, a fully functional and equipped LMT (Large Millimeter Telescope) could provide dust and molecular gas masses within the present decade.

5.6 The next-generation high frequency radio telescope (extended VLA)

The best probes of many stages of the star formation process come from cm- and mm-wavelengths. Low excitation molecular lines (CO, HCN, HCO+, CS, CN, SiO, HNCO, etc.) trace cold gas at a variety of densities and physical states. Specifically, low excitation transitions of these molecules in the range ∼50 to 120 GHz are ideal to trace the kinematics, distribution, density, and chemical state of the cold gas that forms stars (e.g., see Figure 5.2). Similarly, the continuum diagnostics over this frequency range offer the chance to study the ionized gas, dust, cosmic rays, and magnetic fields in galaxies (including our own) in ways not possible at other wavelengths. For example, the free-free continuum offers debatably the most direct tracer of recent high mass star formation thanks to its linear dependence on ionizing photon production and total robustness to extinction. These tools have been recognized for decades; however, beyond a few bright lines

[8] We note that although ALMA and the space mission discussed in the previous section are both powerful far-IR facilities, they are not able to perform large surveys of the kind that we describe here, and they will therefore be unable to address science goals that require large statistical samples.

and pioneering detailed studies of individual bright sources, the full power of cm- and mm-wave observations remains largely untapped. The fundamental limitation has, so far, been the surface brightness sensitivity of previous- and current-generation cm- and mm-wave arrays. The emission from tracers that can provide accurate measurements of the gas and plasma properties in galaxies has simply been too faint to reliably use as a general purpose tool, which inevitably requires an interferometer with a tremendous amount of collecting area.

A "next-generation Very Large Array" (ngVLA) focused on the rich frequency range $\nu \sim$ 1–116 GHz has been proposed as a next major direction for the U.S. astronomical community, and will make transformational improvements in observations of these star formation diagnostics. The major advance with the ngVLA would be the enormous gain in sensitivity at many spatial scales due to an unprecedented collecting area operating in the $\nu \sim$ 1–116 GHz range. Additionally, with a new generation of extremely wide bandwidth receivers, heavily multiplexed spectral line and continuum observations will be possible, thereby yielding huge gains in science per hour. A wide bandwidth is needed for continuum sensitivity to measure the shape of the continuum spectral energy distribution, which is key to its interpretation. The diagnostic and astrochemical utility of spectral line observations is vastly improved by simultaneously observing a wide range of transitions; in fact, the comparative behavior of multiple lines and the continuum is the central measurement for this kind of science. Consequently, such a facility would be a groundbreaking tool to study the detailed astrophysics of star formation in the Milky Way and other galaxies. If optimized for high brightness sensitivity, the ngVLA would bring detailed microwave spectroscopy and modeling of the full radio spectral energy distribution into regular use as survey tools at resolution 0.1–1", thus following the star formation process down to the scales of individual giant molecular clouds and HII within heterogenous samples of nearby galaxies.

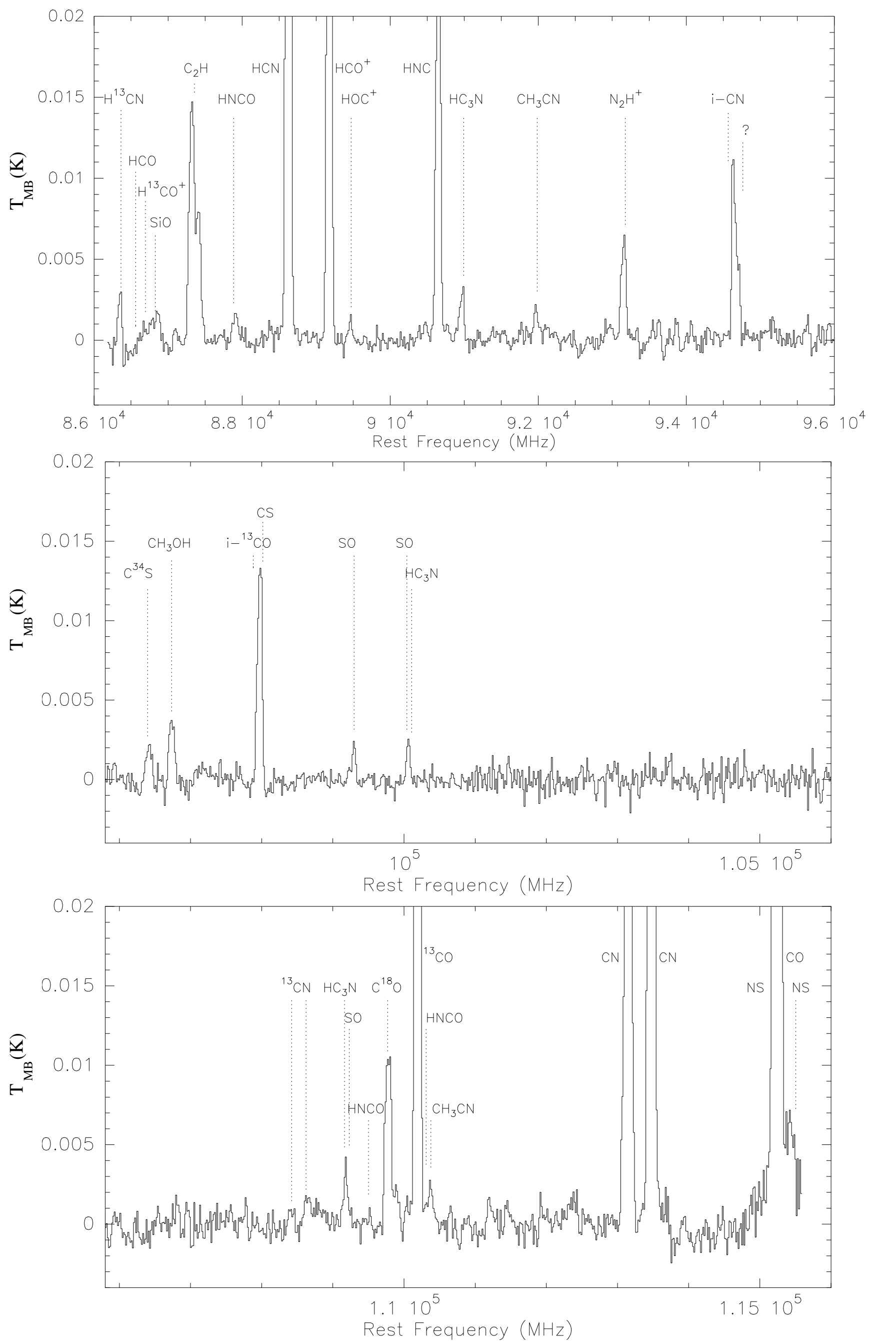
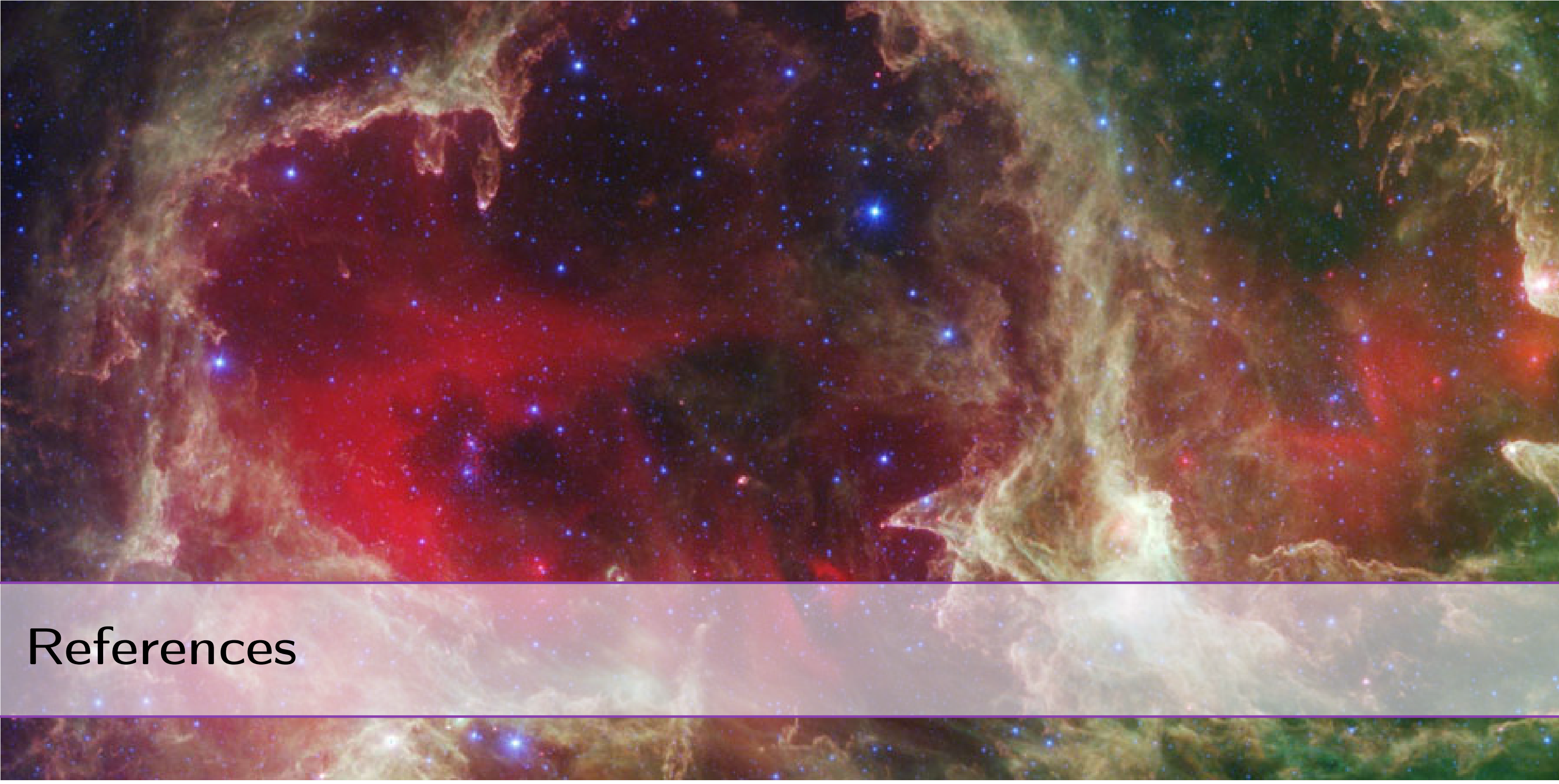


Figure 5.2: From *Aladro et al.* (2013): This spectral scan of the nucleus of NGC 1068 shows the rich spectrum over the range 86-115 GHz. A ngVLA could have the sensitivity to quickly image all of these transitions at $\approx 1''$ resolution across the entire area of a normal galaxy or Galactic star forming region. These lines trace shocks, gas density, excitation, ionization and the UV field, and offer the chance to harness the growing field of astrochemistry as a main tool to understand the physics of the interstellar medium, star formation, and galaxy evolution.



References

Bally, J., & Zinnecker, H. 2005, AJ, 129, 2281

Bastian, N., Covey, K. R., & Meyer, M. R. 2010, ARA&A, 48, 339

Bernyk et al. 2014, arXiv:1403.5270

Bisbas, T. G., Bell, T. A., Viti, S., Yates, J., & Barlow, M. J. 2012, MNRAS, 427, 2100

Bolatto, A.D., Leroy, A.K., Rosolowsky, E., Walter, F., & Blitz, L. 2008, ApJ, 686, 948

Bonnell, I.A., Bate, M.R., & Zinnecker, H. 1998, MNRAS, 298, 93

Bonnell, I.A., Bate, M.R., Clarke, C.J., & Pringle, J.E. 2001, MNRAS, 323, 785

Cappellari, M., McDermid, R. M., Alatalo, K., et al. 2012, Nature, 484, 485

Carilli, C.L., & Walter, F. 2013, ARA&A, 51, 105

Chilingarian, I. et al. 2010, A&A, 518, A61

Clauwens, B., Schaye, J., & Franx, M. 2015, MNRAS, 449, 4091

Colombo, D., Hughes, A., Schinnerer, E., et al. 2014, ApJ, 784, 3

Conroy, C., & van Dokkum, P. G. 2012, ApJ, 760, 71

Cormier, D., Madden, S.C., Lebouteiller, V., et al. 2015, A&A , 578, A53

Da Rio, N., Tan, J.C., & Jaehnig, K. 2014, ApJ, 795, 55

Dalcanton, J., Williams, B., & ANGST Collaboration 2008, Galaxies in the Local Volume, 115

Dokkum, P.G.v. 1999, Ph.D.Thesis,

Elmegreen, B. G. 2000, ApJ, 530, 277

Elmegreen, B. G. 2006, ApJ, 648, 572

Elmegreen, B.G., & Falgarone, E. 1996, ApJ, 471, 816

Faucher-Giguère, C.-A., & Kereš, D. 2011, MNRAS, 412, L118

Faucher-Giguère, C.A., Hopkins, P. F., Kereš, D., et al. 2015, MNRAS, 449, 987

Fischer, J., Sturm, E., González-Alfonso, E., et al. 2010, A&A, 518, L41

Fumagalli, M., Hennawi, J. F., Prochaska, J. X., et al. 2014, ApJ, 780, 74
Glover, S. C. O., Federrath, C., Mac Low, M.-M., & Klessen, R. S. 2010, MNRAS, 404, 2
Goldsmith, P.F., Heyer, M., Narayanan, G., et al. 2008, ApJ, 680, 428
Grenier, I.A. et al. 2005 Science, 307, 1292
Hartmann, L., & Burkert, A. 2007, ApJ, 654, 988
Heckman, T. M., Alexandroff, R.M., Borthakur, S., et al. 2015, ApJ, 809, 147
Heckman, T. M., Armus, L., & Miley, G. K. 1990, ApJS, 74, 833
Heckman, T. M., Lehnert, M. D., Strickland, D. K., & Armus, L. 2000, ApJS, 129, 493
Hennebelle, P., & Pérault, M. 2000, A&A, 359, 1124
Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, MNRAS, 445, 581
Hughes, A., Meidt, S.E., Schinnerer, E., et al. 2013, ApJ, 779, 44
Jog, C.J. 2013, MNRAS, 434, L56
Johnson, S. D., Chen, H.-W., & Mulchaey, J. S. 2015, MNRAS, 452, 2553
Jones, T., Stark, D. P., & Ellis, R.S. 2012, ApJ, 751, 51
Kennicutt, R. C., Calzetti, D., Aniano, G., et al. 2011, PASP, 123, 1347
Kroupa, P., Tout, C. A., & Gilmore, G. 1993, MNRAS, 262, 545
Krumholz, M.R., Dekel, A., & McKee, C.F. 2012, ApJ, 745, 69
Kwan, J & Valdes, F., 1987, ApJ, 315, 92.
Langer et al. 2010, A&A, 521, L17
Lanz, L. et al. 2014, ApJ, 785, 39
Lehnert, M. D., & Heckman, T. M. 1996, ApJ, 462, 651
Lemson, G. and the Virgo Consortium 2006, astro-ph/0608019
Leroy, A.K., Bolatto, A.D., Ostriker, E.C., et al. 2015, ApJ, 801, 25
Li, Z.-Y., Banerjee, R., Pudritz, R. E., et al. 2014, Protostars and Planets VI, 173
McCray, R. & Kafatos, M, , 1987, ApJ, 317, 190.
McKee, C. F. 1989, ApJ, 345, 782
McKee, C.F., & Tan, J.C. 2003, ApJ, 585, 850
McKee, C.F., & Tan, J.C. 2002, Nature, 416, 59
McLaughlin, D. E., & Pudritz, R. E. 1997, ApJ, 476, 750
Meidt, S.E., Schinnerer, E., Garcia-Burillo, S., et al. 2013, ApJ, 779, 45
Nakamura, F., & Li, Z.-Y. 2007, ApJ, 662, 395
Nelson, D. et al. 2015, arXiv:1504.00362
Ossenkopf, V., Röllig, M., Kramer, C., et al. 2011, EAS Publications Series, 52, 181
Ostriker, E. & Kim, W.-T. , 2004, in Proceedings of ASP Conference 317, p.248.
Overzier, R. et al. 2013, MNRAS, 428, 778
Paresce, F., & De Marchi, G. 2000, ApJ, 534, 870
Pettini, M., Steidel, C. C., Adelberger, K. L., Dickinson, M., & Giavalisco, M. 2000, ApJ, 528, 96
Pineda, J.L., Langer, W.D., Velusamy, T., & Goldsmith, P.F. 2013, A&A, 554, A103
Rix, S. A., Pettini, M., Leitherer, C., et al. 2004, ApJ, 615, 98
Röllig, M., Abel, N. P., Bell, T., et al. 2007, aap, 467, 187

Rudie, G. C., Steidel, C. C., Trainor, R. F., et al. 2012, ApJ, 750, 67
Rupke, D. S., Veilleux, S., & Sanders, D. B. 2005, ApJS, 160, 115
Schinnerer, E., Meidt, S.E., Pety, J., et al. 2013, ApJ, 779, 42
Shapley, A. E. 2011, ARA&A, 49, 525
Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, ApJ, 588, 65
Shen, S., Madau, P., Guedes, J., et al. 2013, ApJ, 765, 89
Shu, F. H., Adams, F.C., & Lizano, S. 1987, ARA&A, 25, 23
Springel, V. et al. 2005, Nature, 435, 629
Steidel, C. C., Erb, D. K., Shapley, A. E., et al. 2010, ApJ, 717, 289
Suresh, J., Bird, S., Vogelsberger, M., et al. 2015, MNRAS, 448, 895
Tan, J. C., Krumholz, M. R., & McKee, C. F. 2006, ApJL, 641, L121
Torrey et al. 2015, MNRAS, 447, 2753
Tumlinson, J., Werk, J. K., Thom, C., et al. 2011, ApJ, 733, 111
Turk, M. et al. 2011, ApJSS, 192, 9
Veilleux, S., Cecil, G., & Bland-Hawthorn, J. 2005, ARA&A, 43, 769
Wang, P., Li, Z.-Y., Abel, T., & Nakamura, F. 2010, ApJ, 709, 27
Werk, J. K., Prochaska, J. X., Thom, C., et al. 2013, ApJS, 204, 17
Wolfire, et al. 2010, ApJ, 716, 1191