Warm H$_2$ as a probe of massive accretion and feedback through shocks and turbulence across cosmic time

Thematic Areas:
- Planetary Systems
- Star and Planet Formation
- Formation and Evolution of Compact Objects
- Cosmology and Fundamental Physics
- Stars and Stellar Evolution
- Resolved Stellar Populations and their Environments
- Galaxy Evolution
- Multi-Messenger Astronomy and Astrophysics

Principal Author:
Name: Philip N. Appleton
Institution: Caltech
Email: apple@ipac.caltech.edu
Phone: (+1) 626 780 6358

Co-authors:
Lee Armus: Caltech; Francois Boulanger: Observatoire de Paris; Matt Bradford: JPL; Jonathan Braine: University of Bordeaux; Volker Bromm: University of Texas at Austin; Peter Capak: Caltech; Michelle Cluver: Swinburne University of Technology; Asantha Cooray: UC Irvine; Tanio Diaz-Santos: Universidad Diego Portales; Eiichi Egami: University of Arizona; Bjorn Emonts, NRAO, Challottesville; Pierre Guillard: IAP-Paris; George Helou: Caltech; Lauranne Lanz: Dartmouth College; Susanne Madden: CEA, Saclay; Anne Medling: University of Toledo; Ewan O’Sullivan: Center for Astrophysics | Harvard-Smithsonian; Patrick Ogle: STScI; Alexandra Pope: University of Massachusetts; Guillaume Pineau des Forêts: IAS, Orsay; J. Michael Shull: U. of Colorado; John-David Smith: University of Toledo; Aditya Togi: University of St. Marys; C. Kevin Xu: National Astronomical Observatories, Beijing

Galaxy formation depends on a complex interplay between gravitational collapse, gas accretion, merging, and feedback processes. Yet, after many decades of investigation, these concepts are poorly understood. This paper presents the argument that warm H$_2$ can be used as a tool to unlock some of these mysteries. Turbulence, shocks and outflows, driven by star formation, AGN activity or inflows, may prevent the rapid buildup of star formation in galaxies. Central to our understanding of how gas is converted into stars is the process by which gas can dissipate its mechanical energy through turbulence and shocks in order to cool. H$_2$ lines provide direct quantitative measurements of kinetic energy dissipation in molecular gas in galaxies throughout the Universe. Based on the detection of very powerful H$_2$ lines from $z = 2$ galaxies and proto-clusters at the detection limits of Spitzer, we are confident that future far-IR and UV H$_2$ observations will provide a wealth of new information and insight into galaxy evolution to high-$z$. Finally, at the very earliest epoch of star and galaxy formation, warm H$_2$ may also provide a unique glimpse of molecular gas collapse at $7 < z < 12$ in massive dark matter (DM) halos on their way to forming the very first galaxies. Such measurements are beyond the reach of existing and planned observatories.
The importance of direct H$_2$ detection of warm gas throughout the universe. H$_2$ is an important coolant of shocked gas. Pure rotational (quadrupole) transitions radiate in the rest-frame mid-IR, and were routinely detected in the spectra of nearby galaxies by ISO and Spitzer. Unlike indirect measurements of molecular hydrogen mass through uncertain conversion from trace molecules such as CO, the direct detection of emission lines from molecular hydrogen (H$_2$) is easily achieved with cryogenically-cooled space telescopes when the gas is heated above $\sim$100K, or fluorescently excited (Shull & Beckwith 1982). Alternatively, H$_2$ can be detected in absorption in the far-UV through the Lyman and Werner-bands (e.g., Rachford et al. 2009).

When star formation dominates galaxies, photo-electrons, ejected from small grains and PAH molecules by UV photons, drive the heating of the neutral gas (Fig. 1a). However, non-radiative forms of excitation of the H$_2$ are also present, such as gas heated in shocks and turbulence in subsets of galaxies, collisional systems or groups (e.g., Fig. 1b; Appleton et al. 2006; Roussel et al. 2007; Cluver et al. 2010; Peterson et al. 2012; Cluver et al. 2013; Stierwalt et al. 2014), radio galaxies (Ogle et al. 2007, 2010; Guillard et al. 2012a), and cluster galaxies (Egami et al. 2006; Ogle et al. 2012; Sivanandam et al. 2014). Most show unusually strong H$_2$ line-luminosities, high H$_2$/FIR and H$_2$/PAH ratios, often in regions devoid of obvious star formation. Cosmic rays have also been suggested as another source of heating in some galaxy cluster sources (Ferland et al. 2008), as well as X-rays produced by accreting super-massive black holes (Active Galactic Nuclei - AGN). Observations and modeling of a well studied nearby example (Fig. 2a) suggest that low-velocity magnetohydrodynamic (MHD) shocks are responsible for the high power in the lowest-lying, most luminous, rotational lines (Guillard et al. 2009; Lesaffre et al. 2013; Appleton et al. 2017) which encode information about the most massive H$_2$ component.

JWST will provide a remarkable view of some of the brightest IR H$_2$ lines in the nearby ($z < 2$) universe, but its wavelength coverage is limited. However, the next generation of wide-band Far-IR (FIR) and Far-UV (FUV) telescopes will open a new window into the turbulent universe to cosmological distances, including the virialized gas that cools to form early galaxies.

Figure 1: Spitzer spectra showing the contrast between galaxies excited by star formation processes compared with shocks (very strong H$_2$ lines)

H$_2$ excitation diagrams are a powerful way of exploring the distribution of level populations in H$_2$ molecules in galaxies (Rigopoulou et al. 2002; Roussel et al. 2007) and are especially powerful when many H$_2$ transitions are measured together. In galaxies, multiple temperature components are often present, resulting in a distribution of curved points in the excitation diagrams. This is
most naturally explained by a power law distribution of temperatures (Togi & Smith 2016). In systems suspected of shock or turbulent heating, a large fraction (perhaps all) of the H\(_2\) is warm (T > 200 K) and exhibits low values of power-law index (Fig. 2b) and can be fit with shock models to derive important properties (Fig. 2c,d), such as mass flow into the shock and the total energy dissipation. Observations of the intergroup gas in Stephan’s Quintet reveal that the bulk of the energy deposited on large scales is involved in a turbulent cascade and is dissipated mainly through H\(_2\) line emission (Guillard et al. 2012b; Appleton et al. 2013, 2017), with a second component from Ly\(\alpha\) emission (Guillard et al. in prep). The cooling time of the gas through these lines is so short that kinetic energy must be injected faster than it can dissipate through line cooling. Strong mechanical forcing creates supersonic turbulence in the H\(_2\) gas (Guillard et al. 2009) and leads to powerful H\(_2\) line emission as the dominant cooling channel (Appleton et al. 2013, 2017). In these systems, rotational H\(_2\) line emission provides a direct measure of the turbulent dissipation rate.

H\(_2\) as a Tracer of Gas Accretion in Galaxies, Clusters and ProtoClusters:

The details of how gas accretes onto galaxies, and how feedback from star formation and AGNs might redistribute, heat and stir-up the gas, is still one of the unsolved problems in galaxy evolution. In the standard “hot mode” ΛCDM picture, gas accumulating in DM halos above a critical mass (M > 10\(^{12}\) M\(_\odot\)) is shock-heated at the virial radius (Rees & Ostriker 1977), but then cools, eventually raining down isotropically to form a slowly rotating galaxy. Shocked molecular hydrogen lines would be detected just inside the main virial shock as the gas passes through the warm stage. Star formation and feedback from AGNs will also drive further heating. Alternatively, at lower halo masses, gas can accumulate at the center through “cold flows” of diffuse, T = 10\(^4\) K gas (Fardal et al. 2001; Dekel & Birnboim 2006). The gas can flow into the central forming galaxy without developing strong shocks, creating high angular momentum proto-disks, with spiral-like filamentary morphologies (Danovich et al. 2015; Stewart et al. 2017). In such cases, rapid cooling by shocks is likely to occur close to the center of the galaxy, and major cooling to form stars may be more centrally concentrated. Feedback from AGN (e.g., radio jets) can also complicate the picture (Russell et al. 2014; Faucher-Giguère et al. 2016; Cornuault et al. 2018).

Exploring the Warm H\(_2\) properties of Galaxies across Mass Scales and Redshift:

There are already strong hints that the the H\(_2\) rotational lines will be powerful in galaxies at high z. Fiolet et al. (2010) discovered unusually strong 0-0S(3)9.6µm emission in the stacked Spitzer spectra of 16 z~2 ULIRGs, suggesting that prominent rotational lines will be common in IR luminous galaxies at z ~2-3. Even more exciting was the discovery of tremendous luminosity (3 × 10\(^{10}\)L\(_\odot\)) in the single observed H\(_2\)9.66µm line detected in center of the massive Spiderweb...
 proto-cluster at $z = 2.15$ (Ogle et al. 2012). Unfortunately, Spitzer was unable to measure more than one line because of limited spectral coverage and sensitivity, and so a complete picture of the extent, luminosity and excitation properties of the warm gas awaits future FIR observations. The core of the Spiderweb protocluster contains a vast extended molecular reservoir emitting in CO and [CI] (Emonts et al., 2016, 2018b), with H$_2$O observations implying that shock heating and localized gas cooling occurs along the powerful radio jet (Gullberg et al. 2016; Fig. 3c). It is thus possible that a significant fraction of the H$_2$ luminosity comes from a combination of accretion power and/or gas heated in the radio jet. Radio jets can deposit significant power which emerges as H$_2$ emission (Ogle et al. 2010; Lanz et al. 2015, 2016).

A key to exploring the H$_2$ excitation properties over a wide range of redshifts in the FIR is broad spectral coverage allowing the capture of many transitions in the same spectrum, enabling estimates of the gas temperatures and masses. The often powerful H$_2$ 0-0S(3) and S(1) lines are unobservable above a redshift of 1.7 by JWST/MIRI.

Future deep, large-volume, surveys of the Universe with sensitive FIR and FUV spectrometers in the next generation of space telescopes will enable the following science goals:

- **Characterizing the full H$_2$ excitation properties (temperatures, warm H$_2$ masses) of large samples of galaxies spanning critical halo masses in the range $10^{11}$ to $10^{13}$ M$_\odot$.** Using well developed modeling tools for photodissociation regions and shocks, combined with observations of other species, such as HD, [CI], CO, and far-IR [OI] and [CII] emission, it will be possible to explore those cases where turbulence and shocks dominate over star formation. Trends with mass, environment, redshift (cold mode accretion should become more common at higher z as halos become less massive), and metallicity will be explored. Since almost nothing is known about warm H$_2$ above $z = 0.5$, this will open up completely unexplored territory.

- **Spatial mapping of warm H$_2$ on scales commensurate with the virial radius in multiple H$_2$ rotational lines for galaxies at $z = 2-3$, at the peak in the cosmic star formation rate energy density.** Mapping the full spectral properties of the lowest rotational H$_2$ transitions inside the virial radii of galaxies at $2 < z < 3$ will provide a strong constraint on the most massive, cooler, gas reservoirs. They will span a range of halo masses where gas infall transitions from hot mode to cold flow accretion. At $z = 3$, for a galaxy with a halo mass of few $\times 10^{12}$ M$_\odot$, the typical virial radius at which infalling gas would shock and undergo rapid cooling in the hot mode accretion model is $\sim$100-140 kpc (12-18”). The 0-0S(1)17$\mu$m line (shifted to 70$\mu$m) is likely to be the strongest line in the typical $T = 150-300$ K range. Future FIR telescopes with apertures > 3 m could resolve
gas collapsing from the virial radius to the center. Lower halo mass galaxies with “cool flows” will likely exhibit different warm H$_2$ morphologies than the hot-mode examples, since the gas can smoothly fall directly to the center before undergoing significant compression. Multi-beam mapping in many rotational H$_2$ lines simultaneously (as envisaged in e.g. the OST concept) will be incredibly powerful for examining the sources of heating and multi-waveband morphologies in galaxies both near and far.

- **Identifying the role of AGN-radio mode feedback in redistributing the gas both spatially and into different phases.** Warm H$_2$ measurements, NIR/FIR AGN line diagnostics and IR SEDs, when combined with higher resolution radio continuum (SKA, ngVLA) and other molecular species (ALMA), will help us disentangle the complexity of AGN interactions with their surroundings (U et al. 2019). AGN jets are known locally to heat the gas and suppress star formation (e.g. Ogle et al. 2012; Ogle 2019 in prep.) where the warm H$_2$ dominates the mass in the outflows (unlike CO and HCN which are often just tracers). H$_2$ diagnostics may also be helpful in measuring the large-scale reservoirs of gas heated by the AGN. Such gas may be a critical link to an AGN feedback cycle, which is central to structure formation and galaxy evolution models. Extensive filaments of ionized and molecular gas at the center of lower-z cool-core clusters and groups likely fuel the AGN. Based on Spitzer observations, such clusters are known as powerful warm H$_2$ emitters. However, the cooling mechanism in the filaments is poorly understood, and it is not certain what triggers the gas to precipitate out of the hot phase, nor how it cools so rapidly. H$_2$ detection will be an excellent new window into the AGN heating and cooling cycle as observations progress to higher and higher redshift, where the balance between different kinds of AGN feedback (radiatively efficient versus jet-dominated) may shift.

- **Detecting H$_2$ in cold flows through UV absorption** Although we expect to trace cold molecular filaments as far as possible in emission in the molecule hydrogen lines and other mid-IR atomic lines, the detection of Lyman and Werner-band H$_2$ in absorption along the line of sight to AGN has previously been demonstrated (Savage et al. 1977; Tumlinson et al. 2002; Rachford et al. 2009), and will be feasible to $z = 3$ with the next generation of large FUV/UVO telescopes (e.g., LUVOIR). This will enable very sensitive measurements of a possible molecular component to cold filaments along sight lines that pass near large mass concentrations. Although cold flows are expected to be mainly ionized by background UV radiation, merging substructures, flow instabilities and shocks may create pockets of higher density gas that might lead to in situ star formation within the cold flows. Such stars will populate the outer virialized galaxy halos with time. Higher density clumps are common features of numerical models of the cosmic web (Fig. 3a).

- **Searching for H$_2$ in massive DM halos at $z > 7$.** In very low-metallicity environments, gas cooling in massive galaxies will be dominated by Ly$\alpha$ and H$_2$ pure-rotational lines. Since Ly$\alpha$ emission can be easily extinguished via absorption and scattering in the increasingly neutral intergalactic medium (IGM) at high $z$, H$_2$ emission lines will serve as a powerful probe for such first-generation galaxies. Estimating H$_2$ line luminosities in early galaxies has been a focus of theoretical studies for many years (Omukai & Kitayama 2003; Santoro & Shull 2006; Mizusawa et al. 2005; Gong et al. 2013; Liu et al. 2019). Such predictions contain many uncertainties, such as the importance of self-shielding effects from dissociative UV radiation from the first stars (Wise & Abel 2008), the growth (with or without grains) and destruction of H$_2$ in shocks, as well as other feedback effects. Magnetic fields may also play role in reducing compression via C-shocks, lowering the gas temperature, and boosting rotational H$_2$ emission (see Lesaffre et al. 2013).
Most predictions suggest\(^1\) that the direct detection of H\(_2\) line luminosities for realistic halos would be challenging, but not impossible, for future cold, large IR telescopes like OST, especially with gravitational lensing. For example, utilizing strong lensing clusters to provide lensing amplifications factors of 6-10 for H\(_2\) point-like cores, we have identified a range of DM halo masses (Fig. 4a) that are common enough at \(z = 8\) to fall within the lensing volume of a small sample of bright lensing clusters. In such cases, lensing can potentially provide a strong enough H\(_2\) signal to be detectable with currently envisaged large cold FIR sensitive telescopes (Fig. 4b). Finally, we emphasize that the proposed deep FIR spectroscopic observations will also naturally address the onset of heavy elements. Small amounts of metal enrichment can shift the cooling from H\(_2\) to metal fine-structure lines (e.g., [Si II]34.8\(\mu\)m, [Fe II]25.99/35.35\(\mu\)m; [CII]158\(\mu\)m; Santoro & Shull 2006), many of which are also observable in the FIR at high-\(z\).

\[^1\)\]except for very rare DM halo masses in excess of \(10^{12} \text{ M}_\odot\) (Liu et al. 2019)

**In summary:** The direct detection of H\(_2\) in the Universe is vital for the following reasons:

- H\(_2\) is an important coolant of shocked gas, including the virialized gas that cools to form early galaxies. To fully capitalize on this potential, we need to harness both UV absorption from cool H\(_2\), and rotational emission from warm H\(_2\) for which models have already proven successful locally. This exciting science cannot be done with currently planned telescopes.

- The excitation conditions of H\(_2\) are strikingly variable in the local Universe, encoding information on the accretion modality, heating conditions, AGN-driven feedback, and star-forming molecular reservoirs that is robust against the changing metal and dust abundance in the early Universe. Full exploitation of the information contained in the multiple H\(_2\) cooling lines requires a very broad, sensitive, spectroscopic FIR wavelength coverage.

- H\(_2\) emission/absorption may be the only way to directly probe the gas cooling and feeding the most massive metal-free dark matter halos and to assess the molecular reservoirs inside dust- and metal-free star forming regions at the earliest epochs. This goal is within reach for lensed regions for the next generation of cryogenic FIR telescopes like OST.

---

**Figure 4:** (Left) Dark matter mass function for \(\Lambda\)CDM showing the target mass range that provides (at \(z = 8\)) approximately 1 halo per 133 Mpc\(^3\). This is \(-\) the co-moving \(z = 8\) sample volume if the lensing caustics of a single rich lensing cluster is observed. (Right) Predictions for 6 and 10 x lensed (colored lines) and un-lensed (grey lines) for H\(_2\) line emission as a function of redshift for a halo of mass \(6.9 \times 10^{10} \text{ M}_\odot\), extrapolated from the models of Liu et al. (2019) and Omukai & Kitayama (2003). Also shown is the 5\(\sigma\) sensitivity threshold for the Origins Space Telescope concept. Detection may be possible around \(z = 8\) with lensing.
References