FIRST STAR SIGNATURE IN INFRARED BACKGROUND ANISOTROPIES

ASANTHA COORAY,1 JAMES J. BOCK,1,2 BRIAN KEATIN,1 ANDREW E. LANGE,1 AND T. MATSUMOTO3

Received 2003 August 20; accepted 2004 January 23

ABSTRACT

Recent cosmic microwave background anisotropy results from the Wilkinson Microwave Anisotropy Probe suggest that the universe was reionized at a redshift around 20 with an optical depth for Thomson scattering of $0.17 \pm 0.04$. Such an early reionization could arise through the ionizing radiation emitted by metal-free Population III stars at redshifts of 10 and higher. We discuss infrared background (IRB) surface brightness spatial fluctuations from such a generation of early star formation. We show that the spatial clustering of these stars at tens of arcminute scales generates a contribution to the angular power spectrum of the IRB anisotropies at the same angular scales. This excess can be potentially detected when resolved foreground galaxies out to a redshift of a few are removed from the clustering analysis. We do not expect faint galaxies at redshifts of $\sim 3$, with magnitudes less than 20 in the $K$ band, to be a source of strong confusion, since the fractional contribution to the IRB from these galaxies is at a level less than a few percent, while the expected contribution from first stars can be 50% or more. Additionally, assuming a Population III stellar spectrum, we suggest that the clustering excess related to the first generation of stars can be separated from brightness fluctuations resulting from other foreground sources and galaxies using multifrequency observations in the wavelength range of $\sim 1–5 \mu m$. In addition to identifying the IR clustering associated with low-redshift galaxy population, the multifrequency data are essential to account for certain foreground contaminants such as zodiacal light, which, if varying spatially over degree scales, can be a significant source of confusion for the proposed study. Using various instruments, we study the extent to which spatial fluctuations of the IRB can be studied in the near future.

Subject headings: cosmology: theory — diffuse radiation — galaxies: formation — infrared: galaxies — large-scale structure of universe

On-line material: color figures

1. INTRODUCTION

The Wilkinson Microwave Anisotropy Probe (WMAP) has now provided strong evidence for an optical depth for electron scattering of $0.17 \pm 0.04$ based on the large-scale polarization pattern related to rescattering of cosmic microwave background (CMB) photons (Kogut et al. 2003). If the reionization process is described as instantaneous and homogeneous, the measured optical depth implies a reionization redshift of $\sim 17 \pm 5$ in a spatially flat universe. Such a reionization redshift is higher than previously suggested by observational data involving the presence of a Gunn-Peterson trough in $z \sim 6$ quasars (Fan et al. 2002). The derived redshift for reionization is at the high end of expectations related to reionization scenarios based on the ionizing radiation from the first generation of star formation associated with Population III stars (e.g., Cen 2003; Fukugita & Kawasaki 2003; Venkatesan, Tumlinson, & Shull 2003; Wyithe & Loeb 2003). These stars are expected to be very massive with a top-heavy mass function (Bromm, Coppi, & Larson 1999, 2002; Abel, Bryan, & Norman 2000, 2002), and their detection is important to understand the astrophysics associated with the reionization process that followed the end of dark ages.

While the direct detection of an individual Population III star or a star cluster is beyond the capability of upcoming telescopes or instruments, the first star signature can be potentially detected via indirect methods such as associated line emission, e.g., He II recombination lines of the surrounding ionized halo (Oh, Haiman, & Rees 2001; Tumlinson, Giroux, & Shull 2001) or through Lyα emission lines (e.g., Tumlinson, Shull, & Venkatesan 2003). In Oh, Cooray, & Kamionkowski (2003), it was suggested that the heating of the surrounding interstellar medium by Population III supernovae and the subsequent transfer of energy to the CMB through Compton cooling lead to a substantial contribution to small angular scale CMB anisotropies, similar to the Sunyaev-Zel’dovich (SZ; Sunyaev & Zel’dovich 1980) effect. While an individual Population III supernova remains unresolved and undetectable, the large bias factors of halos containing these supernovae, with respect to the linear density field, lead to an excess clustering signature in CMB anisotropies at tens of arcminute scales when compared to the shot-noise contribution related to their finite density. The CMB anisotropies related to Population III supernovae can be separated from the dominant SZ fluctuations related to galaxy clusters since massive clusters that dominate the angular power spectrum at arcminute scale can be identified and removed in future higher angular resolution CMB data (A. Cooray et al. 2004, in preparation).

Here we suggest that a signature of stars themselves can be found in the infrared background (IRB). In particular, if Population III stars are found primarily at redshifts between 10 and 30, they are expected to contribute to the IRB at wavelengths between 1 and $5 \mu m$. Recent estimates now suggest that a large fraction of the IRB total intensity may be due to these stars (Santos, Bromm, & Kamionkowski 2002; Salvaterra & Ferrara 2003). A substantial IRB arises from
the Population III population as a result of not only the direct emission associated with these stars but also indirect processes that lead to free-free and Lyα emission from the ionized nebulae, or H ii regions, surrounding these stars. The fractional contribution to the absolute, and isotropic, background by the Population III population, however, cannot be easily determined observationally as a result of additional contributions from foreground galaxies, stars, and confusion from zodiacal light.

Instead of the background intensity, or the monopole, we suggest that the presence of first stars can be established through spatial fluctuations in the surface brightness of the IRB. For this purpose, we make use of the angular power spectrum of IRB anisotropies and show that the power spectrum is expected to contain an excess clustering signature at tens of arcminute scales related to the Population III population. While Population III stars may dominate tens of arcminute-scale fluctuations at short IR wavelengths (between 1 and 3 μm), any unresolved galaxies at redshifts of 3 and higher are expected to dominate IRB anisotropy at wavelengths greater than 4 μm over all angular scales of interest. Thus, we expect that the Population III clustering signature can be separated from the angular power spectrum due to foreground galaxies and other sources in multifrequency near-IR (NIR) images as a result of its unique spatial and spectral signature. Furthermore, when the angular power spectrum due to Population III stars is estimated, one can use that information, for example, to measure the formation rate of the first generation of stars in the universe at redshifts between 10 and 30.

Note that spatial fluctuations in the IRB have already been detected in several ground-, suborbital-, and space-based experiments: Xu et al. (2002) at 4 μm using a rocket-borne experiment; at J, H, and K band with Two Micron All Sky Survey (2MASS) data by Kashlinsky et al. (2002; see also Odenwald et al. 2003); Infrared Telescope in Space (IRTS) data by Matsumoto et al. (2000); and at large scales with the Diffuse Infrared Background Experiment (DIRBE) on the Cosmic Background Explorer (COBE) by Kashlinsky, Mather, & Odenwald (1996a; see also Kashlinsky et al. 1996b). Absolute background measurements generally indicate an excess above that expected from galaxies alone (Kashlinsky & Odenwald 2000; Madau & Pozzetti 2000; Cambresy et al. 2001; Wright & Johnson 2001). Additionally, IRTS data also indicate a clustering excess at 100′′ scales between 1.4 and 2.1 μm, which remains unexplained by the known properties of low-redshift galaxies (Matsumoto et al. 2000, 2001). Expectations for an overall excess in anisotropy include both faint galaxies at high redshifts (Jimenez & Kashlinsky 1999) and a Population III component related to the first generation of stars at redshifts greater than 10 (e.g., Magliocchetti, Salvaterra, & Ferrara 2003). While simple estimates suggest that the IRB total intensity can be easily explained with Population III stars alone (Santos et al. 2002; Salvaterra & Ferrara 2003), it is useful to also consider if these stars can be used to explain anisotropy fluctuations already detected or, if not, the expected level of spatial fluctuations from these stars.

The current observational results related to IRB fluctuations are, unfortunately, limited to either small angular scales (such as ~1″–30″ measurements by 2MASS) or large angular scales (>0.5′′ with COBE DIRBE), with limited measurements at arcminute to tens of arcminute scales. While 2MASS measurements by Kashlinsky et al. (2002) are useful to understand the presence of a Population III contribution, at such small angular scales, both foreground sources and Population III stars produce power-law–like contributions to the angular power spectrum and cannot easily be separated (Magliocchetti et al. 2003). Additionally, the Population III contribution can also be misidentified easily as a source of noise, since, at these small angular scales, one expects a shot-noise–type power spectrum as a result of the finite number density of these stars.

In order to understand the clustering nature of the Population III population, here we study the angular power spectrum of IRB fluctuations, as a function of wavelength, and introduce the use of power spectra at each of these wavelengths and the cross spectra between different wavelengths as a basis to study the star formation history related to Population III objects. Our calculations extend those of Magliocchetti et al. (2003), by considering the angular power spectrum of Population III stars over a wide range of angular scales instead of arcsecond scales studied there with respect to 2MASS measurements. We argue that a well-planned observational program, concentrating on the clustering at arcminute to tens of arcminute scales or more in the wavelength range of 1–5 μm, is clearly needed to understand the presence of Population III stars in IRB data. Since Population III stars are expected at redshifts below 30, we find that the upcoming and planned wide-field IR and near-IR missions from space, such as the Spitzer Space Telescope5 and the proposed WISE mission, are unlikely to play a major role since the Population III signature is reduced for λ ≥ 3 μm. The upcoming Astro-F2 mission has imaging capabilities at arcsecond resolution in the K, L, and M bands between ~1.7 and 5 μm with its Near-Infrared Camera (NIRC) and at a higher sensitivity at short wavelengths (Watarai et al. 2000; Pearson et al. 2001). While observations with Astro-F are only restricted to the 10′ × 10′ field of view, these observations are still useful to understand the presence of Population III sources and their contribution to IRB spatial fluctuations at angular scales of a few arcminutes.

In order to understand the spatial fluctuations due to Population III sources at tens of arcminute scales and above and also to probe clustering at arcminute scales at wavelengths below 2 μm, we also consider a multiwavelength wide-field survey using a rocket-borne experiment. This experiment is planned to image the sky in multiple bands in the wavelength range between 1 and 5 μm, with the primary goal of exploring the presence of Population III stars in IRB spatial fluctuations. While we present a general discussion, we focus on the ability of these experiments to extract Population III information through spatial clustering and to remove confusing point sources related to both galactic stars and foreground galaxies.

The paper is organized as follows. In the next section we calculate the power spectrum of IRB anisotropies for both Population III stars and foreground sources. To describe the emission from Population III stars, we make use of the stellar spectra calculated by Santos et al. (2002). These calculations also include nebular emission associated with both free-free and Lyα radiation associated from the ionized H ii region surrounding individual Population III stars. Note that although we consider Population III sources as the source or ionization, our calculations do not necessarily depend on this assumption. The WMAP results indicate evidence for a highly efficient source of star formation at high redshift and could be in the form of Population II to Population III stars. While we make use of a Population III spectrum to illustrate the clustering in spatial fluctuations in the IRB, we also expect similar spatial

4 See http://sirtf.caltech.edu.
5 See http://www.ir.isas.ac.jp/ASTRO-F.
fluctuations (although with a different amplitude) if the reionization is related to a different stellar spectrum, say, associated with mildly metal-enriched stars, instead of pure metal-free stars. In such a case, one can simply replace the correct stellar spectrum instead of the Population III one used here and repeat our calculations. When illustrating our calculations, we take cosmological parameters from the currently favored ΛCDM cosmology consistent with recent WMAP results (Spergel et al. 2003), including a normalization for the matter power spectrum, at scales of 8 h⁻¹ Mpc, of σ₈ = 0.9.

2. ANGULAR POWER SPECTRUM OF IRB FLUCTUATIONS

Following standard approaches in the literature, we calculate the angular power spectrum of IRB surface brightness fluctuations due to both foreground sources at redshifts between 0 and 7 and Population III stars at redshifts between 10 and 30; except for the lower end of the Population III distribution (which we simply take here as the transition from Population III to Population II stars, discussed further in § 2.1), our calculations are independent of the end points of the redshift ranges considered.

In general, the clustering contribution to the angular power spectrum resulting from foreground galaxies is mostly a power law, which is now well explained with techniques such as the halo model (e.g., Cooray & Sheth 2002). Population III stars, or any other source that traces linear clustering at redshifts of order 15 or so, on the other hand, have an angular power spectrum that peaks at tens of arcminute scales. Such a spectrum is expected since high-redshift sources are expected to be strongly biased with respect to the linear dark matter density field and the linear fluctuation power spectrum contains a peak, or a turnover, at the scale corresponding to the matter-radiation equality. A detection of such a peak in the power spectrum of any tracer source of linear density field would be significant since it will help establish our basic understanding of clustering evolution. Moving to smaller scales corresponding to a few tens of arcseconds and below, one expects the background Population III stars to show a shot-noise-type spectrum associated with the finite number density of these sources on the sky. In the case of Population III stars, the transition scale from linear to shot-noise-type clustering is highly model dependent, although with observational data, one can use a measured transition as an additional constraint on the astrophysics of the Population III population. Given that shot-noise contributions result from both detector noise and other confusions, such as galactic foregrounds, we do not expect a precise measurement of the shot-noise power associated with Population III stars to be feasible. For similar reasons, it is unlikely that arcsecond-scale IRB clustering information is useful for a separation of the Population III contribution from foreground galaxies.

Our calculational approach to modeling the angular power spectrum of the IRB is similar to the one that was used in Knox et al. (2001) to understand the clustering in the far-infrared background at wavelengths of a few hundred microns or more due to dusty starburst galaxies at redshifts of ~3. The contribution to the IRB intensity, at a given wavelength and toward a direction Ω, can be written as a product of the mean IRB emissivity and its fluctuation

\[
I_\lambda(\Omega) = \int_0^{z_{\text{mir}}} \frac{dz}{dz} \frac{d^2r}{H(z')} \left[ 1 + \frac{\delta j_\lambda(r(z))}{j_\lambda(z)} \right],
\]

where \( \delta j_\lambda(z) \) is the mean emissivity per comoving unit volume at wavelength \( \lambda \) as a function of redshift \( z \) and \( r \) is the conformal distance or look-back time, from the observer, given by

\[
r(z) = \int_z^{\infty} \frac{dz'}{H(z')}.
\]

where the expansion rate for adiabatic cold dark matter (CDM) cosmological models with a cosmological constant is

\[
H^2 = H_0^2 \left[ \Omega_m (1 + z)^3 + \Omega_\lambda (1 + z)^2 + \Omega_k \right].
\]

Here \( H_0 \) can be written as the inverse Hubble distance today \( cH_0^{-1} = 2997.9 \) h⁻¹ Mpc. We follow the conventions that in units of the critical density \( 3H_0^2/8\pi G \), the contribution of each component is denoted \( \Omega_i \), \( i = c \) for the CDM, \( b \) for the baryons, and \( \Lambda \) for the cosmological constant. We also define the auxiliary quantities \( \Omega_m = \Omega_c + \Omega_b \) and \( \Omega_k = 1 - \sum_i \Omega_i \), which represent the matter density and the contribution of spatial curvature to the expansion rate, respectively.

The absolute IRB has now been studied both observationally and theoretically (see, e.g., Hauser & Dwek 2001 for a recent review). Here we focus on the spatial fluctuations of the background, \( \delta I_\lambda(\Omega) \), and consider the angular power spectrum of the IRB, which is simply the Legendre transform of the two-point correlation function, \( C(\theta) \):

\[
C_l = 2\pi \int_0^\pi d\theta \cos \theta C(\theta) J_l(\theta).
\]

Denoting the Fourier transform of \( \delta I_\lambda(\Omega) \) as \( \delta I_\lambda(\mathbf{k}) \), one can define the angular power spectrum of the IRB at wavelengths \( \lambda \) and \( \lambda' \), in the flat-sky approximation, as

\[
(\delta I_\lambda(\mathbf{k}) \delta I_\lambda(\mathbf{k}')) = (2\pi)^2 b_\lambda (\mathbf{k}) P_{ss}(\mathbf{k}, \mathbf{k}').
\]

In order to calculate the spatial fluctuations related to the emissivity, we assume \( \delta j_\lambda(r(z)) \) and \( j_\lambda(z) \) trace fluctuations in the source density field, \( \delta s = \delta \rho_s / \bar{\rho}_s \), such that, in Fourier space,

\[
\frac{\delta j_\lambda(k, z)}{j_\lambda(z)} = \frac{\delta s(k, z)}{j_\lambda(z)}.
\]

The source density field fluctuations are defined by the three-dimensional power spectrum, which we define as

\[
(\delta s(k, z) \delta s(k', z)) = (2\pi)^3 b_D (k + k') P_{ss}(k, z).
\]

We calculate this power spectrum of the source distribution, related to both Population III stars and foreground galaxies, based on the halo approach (see Cooray & Sheth 2002 for a recent review). In general, the total power spectrum can be written as a combination of one- and two-halo terms with

\[
P_{ss}(k) = P_{ss}^1(k) + P_{ss}^2(k),
\]

\[
P_{ss}^1 = \int dm n(m) b(m) \langle N_s(m) \rangle \frac{u(k|m)}{\bar{n}_s} \left( \frac{N_s - 1}{\bar{n}_s} \right) u(k|m),
\]

\[
P_{ss}^2 = \int dm n(m) \left( \frac{N_s - 1}{\bar{n}_s} \right) u(k|m),
\]

where \( u(k|m) \) is the density profile (e.g., NFW profile of Navarro, Frenk, & White 1996) in Fourier space normalized
appropriately with mass, \(n(m)\) is the mass function (e.g., PS mass function of Press & Schechter 1974), and \(b(m)\) is the halo bias (e.g., Mo & White 1996; Mo, Jing, & White 1997). The source distribution within halos is encoded by \(\langle N_s(m)\rangle\), the mean source occupation number or the average number of individual sources of interest in each dark matter halo of mass \(m\), and \(\langle N_s(N_s - 1)(m)\rangle\), the second moment of the source distribution. In the case of a Poisson-type distribution of sources, \(\langle N_s(N_s - 1)(m)\rangle = \langle N_s(m)\rangle^2\). In equation (8), \(P^\text{lin}(k)\) is the linear power spectrum of the density field and we use the transfer function of Eisenstein & Hu (1999) to describe the small-scale behavior of this power spectrum. Note that the mean density of sources is given by

\[
\bar{n}_s = \int dm \, n(m) \langle N_s(m)\rangle.
\]  

(9)

Note that this approach to describing clustering of sources that lead to the IRB is similar to the one used in Song et al. (2003) to describe clustering properties of far-infrared background sources.

Using the Limber approximation (Limber 1954), the angular power spectrum for a distribution of sources that trace a three-dimensional power spectrum \(P_{ss}(k)\), when projected on the sky, is given by

\[
C_l^{\gamma \gamma} = \int dz \frac{d^2}{dz^2} f(z) \bar{J}_l(z) \bar{J}_l(z) P_{ss}(k = l/d_A, z),
\]

(10)

where the comoving angular diameter distance is

\[
d_A = H_0^{-1} \Omega_K^{1/2} \sinh\left(H_0 \Omega_K^{1/2}/r\right).
\]

(11)

Note that as \(\Omega_K \to 0, d_A \to r\). At \(l = 50\), the Limber approximation is valid to within \(\sim 5\%\) of the exact calculation that involves radial integrations over the spherical Bessel functions (see Knox et al. 2001 for the formulae that involve the exact calculation) and rapidly converges to the approximated calculation here as one moves to higher \(l\). At multipoles below \(l \sim 50\), the Limber approximation results in an overestimate of power at the level of 10\%, and we do not consider this to be important for the present discussion given the large uncertainties, which lead to several orders of magnitude change in \(C_l\), associated with the description of Population III sources and discussed in § 2.1.

In addition to the clustering signal, at small angular scales, the finite density of sources leads to a shot-noise-type power spectrum in the IRB spatial fluctuations. This shot noise can be estimated through number counts, \(dn/dS\), of the contributing sources, as a function of flux \(S\), and can be written as

\[
C_l^{\text{shot}} = \int_0^{S_{\text{cut}}} S^2 \frac{dn}{dS} dS,
\]

(12)

where \(S_{\text{cut}}\) is the flux cutoff value related to the removal of resolved sources.

We now describe how \(\bar{J}_{\gamma}(z)\) and ingredients related to \(P_{ss}(k)\) can be obtained from simple analytical methods making use of various approaches in the literature.

2.1. Population III Stars

To calculate \(\bar{J}_{\gamma}(z)\), we follow the calculation in Santos et al. (2002) and assume that the mass of stars formed per halo is

\[
m_s = \eta \Omega_b / \Omega_m m
\]

when \(m \geq m_{\text{min}}\) and zero otherwise, where \(\eta\) is the star formation efficiency, which we take here as a free parameter. Furthermore, we assume that Population III stars trace the star formation history at high redshifts and, for calculational purposes, make use of a model that includes molecular hydrogen cooling at a temperature of 400 K. The star formation rate is calculated as

\[
\psi(z) = \eta \frac{\Omega_b}{\Omega_m} \frac{d}{dt} \int_{m_{\text{min}}}^{\infty} dm \, m \, \frac{dn}{dm},
\]

(13)

where \(m_{\text{min}}\) is given in equation (4) of Santos et al. (2002). The star formation rate is plotted in Figure 7, and we return to this later in the context of its estimation from Population III clustering data in the IRB spatial fluctuations.

Given the star formation rate, we can write

\[
\bar{J}_{\gamma}(z) = \frac{1}{4\pi} F_{\gamma} \langle t_{\text{Pop III}} \rangle \psi(z),
\]

(14)

where \(\langle t_{\text{Pop III}} \rangle = 10^6\) yr is the mean lifetime of a Population III star and \(F_{\gamma}\) is the stellar flux as a function of wavelength, including nebular emission, of a Population III star (Fig. 1). Again, we make use of calculations by Santos et al. (2002) on the Population III spectrum and consider two extreme cases with \(f_{\text{esc}} = 0\) and 1, where \(f_{\text{esc}}\) is the escape fraction of ionizing radiation to general intergalactic medium from the nebula. The spectra are reproduced in Figure 1 for a star with mass 300 \(M_\odot\). While the stellar spectrum can easily be described by a simple blackbody, the nebular emission, related to Ly\(\alpha\) radiation and free-free emission, involves a detailed calculation and is model dependent on \(f_{\text{esc}}\). In the case of \(f_{\text{esc}} = 1\), one finds more emission at shorter wavelengths, and we will see later that this leads to an increase in the clustering amplitude at shortest wavelengths when compared to the case with \(f_{\text{esc}} = 0\).

To calculate the three-dimensional power spectrum of sources, we make use of the same assumptions as above,
including the fact that the Population III occupation number is simply determined by the halo mass such that \( \langle N_\star(m) \rangle \propto m \) when \( m \geq m_{\min} \) and zero otherwise. Since the number of stars formed is linearly proportional to the halo mass, the Population III clustering is expected to simply trace that of dark matter, but with a bias factor determined by the halo masses in which stars are found. We also assume a second moment, which is given by the square of the mean, as in the case of a Poisson distribution. At large scales, when \( u(k|m) \to 1 \), the simple dependence on the halo mass leads to the well-known mass-averaged bias as

\[
\langle b^2 \rangle = \frac{\int_{M_{\min}}^{\infty} dm \, m \, b(m) n(m)}{\int_{M_{\min}}^{\infty} dm \, n(m)},
\]

where \( b(m) \) is the halo bias with respect to the density field.

Since our model has a large number of uncertainties, in terms of the redshift distribution of Population III sources, their spectra, and the biasing factor of halos containing these stars, we consider a range of models to establish both optimistic (model A) and pessimistic (model B) estimates of the angular power spectrum. This is similar to the approach considered in Oh et al. (2003) to estimate the Population III supernova contribution to CMB anisotropy fluctuations. Our optimistic estimate involves (1) highly biased sources with a cutoff mass for Population III containing dark matter halos that corresponds to a temperature of 10^7 K, (2) a low end for the Population III redshift distribution at a redshift of 10, and (3) a star formation efficiency of 100\% \( (\eta = 1) \) such that all baryons in these halos convert to Population III stars. Our pessimistic estimate involves (1) sources that have low bias factors (with a minimum temperature set at 5000 K), (2) the low end of the redshift distribution of Population III sources set at 15, and (3) a lower star formation efficiency with \( \eta = 0.1 \). In general, the supernovae related to Population III stars of model A, following Oh et al. (2003), can generate the excess CMB anisotropy fluctuations detected by small angular scale CMB experiments. In each of these two models, because of the uncertainty related to the Population III spectrum, we also vary \( f_{\text{esc}} \) between 0 and 1 to consider variations allowed for the amplitude of the angular power spectrum.

For Population III sources, note that the source power spectrum, in the angular scales of interest, is fully described by the linear clustering power spectrum scaled by a bias factor. The one-halo term, related to the nonlinear part of the power spectrum, is not significant as a result of the fact that strong biasing of halos containing Population III sources leads to a dominant two-halo term. This is also consistent with the rapid reduction in nonlinear clustering one expects at redshifts of order 10 and higher when compared to today. In order to study the extent to which nonlinearities may be significant, one can introduce the fully nonlinear dark matter power spectrum based on numerical fits to dark matter simulations (e.g., Peacock & Dodds 1996). Since sources are not expected to trace dark matter, especially in the nonlinear regime of dark matter clustering, an alternative approach is to calculate the nonlinear correction to source clustering based on the one-halo term related to the halo approach. This requires knowledge related to the second moment of the Population III occupation number. Here, as an approximation, we make use of the Poisson model and assume that the second moment is simply given by the square of the first moment.

In addition to the nonlinear clustering, at these small angular scales, what determines the power spectrum is the shot noise associated with the finite density of Population III sources. To calculate the shot-noise power spectrum reliably, we need a detailed model describing number counts of Population III sources, including the exact mass function of Population III stars within each of the dark matter halos. For the clustering calculation, this information is not required since, in the end, the clustering amplitude depends only on the total baryonic mass within each halo converted to stars, and not how this mass gets divided into stars. Additionally, the shot-noise contribution is more uncertain than the clustered component since the shot noise depends on the rarest Population III sources, which also happen to be the brightest. Similar to our approach for the clustering signal, we also make both optimistic and pessimistic estimates of the shot-noise contribution, but for simplicity, we only show the shot noise related to the \( f_{\text{esc}} = 0 \) spectrum (see Fig. 3); in cases in which the clustering signal is high as a result of higher biasing of rarer halos, we also find a higher shot-noise contribution since such rare halos have a low projected surface density. Note that our highest shot-noise estimates, corresponding to the most optimistic estimate on the Population III clustering contribution, can be excluded by fluctuation analyses of already existing deep K-band imaging data. Thus, the recent 2MASS clustering measurement at small angular scales, discussed with respect to Figure 4, cannot be explained with Population III shot noise alone.

2.2. Foreground Galaxies

To describe the spectrum of an individual galaxy, we make use of the spectrum of M82 (e.g., Silva et al. 1998, where a model fit is presented to explain the observed spectral energy distribution over a wide range of wavelengths from far-infrared to UV light). While the use of such a spectrum to describe all galaxies is likely to be an overestimate, we consider it here since this provides a conservative upper limit on the extent to which foreground galaxies can confuse clustering studies of the first star population using IRB data. To describe the redshift distribution of foreground galaxies, we make use of the normalized star formation rate between redshifts of 0 and 7 but introduce an exponential cutoff at redshifts above 4 such that the redshift distribution turns over and converges essentially to zero at a value of 7. Note that the observational data related to galaxy counts, at redshifts greater than 3, are highly uncertain, although preliminary studies indicate that the fractional contribution to the IRB from galaxies at redshifts greater than 3 (and K-band magnitudes fainter than \( \sim 23 \)) may not be significant (Totani et al. 2001a). Thus, we expect our modeling to provide only an estimate of the galaxy contribution to the IRB. While a lack of massive halos at redshifts greater than 3 reduces the presence of a large number of IR-bright galaxies considerably, the presence of unusually red galaxies in certain K-band imaging data (e.g., Totani et al. 2001b; Daddi et al. 2003) suggests that the situation may be more complex than suggested by numerical calculations. In addition to the suggested distribution, we also considered alternative redshift distributions for foreground galaxies but found consistent results as long as the distribution is not sharply peaked toward the low-redshift end, such as below \( z \) of 1; such models are clearly inconsistent with K-band galaxy counts among other data, and we do not consider such models further.
The clustering of foreground galaxies is also calculated following the halo-based approach with a halo occupation number based on fits to semianalytic numerical simulation data by Sheth & Diaferio (2001):

\[
\langle N_g \rangle = \begin{cases} 
N_0 & 10^{11} M_\odot h^{-1} \leq m \leq M_B, \\
N_0 (m/M_B)^\alpha & m > M_B,
\end{cases}
\]  

(16)

where we set \( M_B = 4 \times 10^{12} M_\odot h^{-1} \) and take \( \alpha \) and \( N_0 \) as free parameters. In our fiducial description, these two parameters take numerical values of 0.8 and 0.7, respectively. In addition to the mean, we also account for departures from the mean in the one-halo term with a detailed model for the second moment, instead of the simple Poisson description used above. Following Scoccimarro et al. (2001), we make use of the binomial distribution, matched to numerical data, to obtain a convenient approximation. The second moment is then

\[
\langle N_g (N_g - 1) \rangle^{1/2} = \beta(m) \langle N_g \rangle,
\]  

(17)

where \( \beta(m) = \log (m/10^{11}) h^{-1} M_\odot \) for \( m < 10^{13} h^{-1} M_\odot \) and \( \beta(m) = 1 \) thereafter. Note that in equation (8) the simplest approach is to set \( p = 2 \) when calculating \( P_{dm}(k) \). In halos that contain only a single galaxy, however, we assume that this galaxy sits at the center and take \( p = 2 \) when \( \langle N_g (N_g - 1) \rangle > \) greater than unity and \( p = 1 \) otherwise. These descriptions for galaxy clustering lead to a power-law–like three-dimensional spectrum over a wide range of physical scales and also explain why the clustering of galaxies shows no excess, as in dark matter clustering, between the transition from linear to nonlinear clustering (see Cooray & Sheth 2002 for details).

While the use of a low-redshift determined occupation number, motivated by semianalytical modeling of “blue” galaxies, to describe clustering properties of IR galaxies may seem problematic, we suggest that this description is adequate for purposes of the present discussion. The main problem related to our description comes from the fact that, based on observations, one naively expects the minimum mass of an IR galaxy at redshifts of \( \sim 3 \) to form in halos with a mass at least 2 orders of magnitude above the minimum mass we have indicated in equation (16). For example, the existing clustering data of IR galaxies at high redshifts (e.g., Daddi et al. 2003) indicate strong clustering of K-band–selected galaxies at a level similar to or greater than the clustering strength of the Lyman break galaxy population at redshifts \( \sim 3 \) and suggest that the minimum mass must be set at a level of \( \sim 10^{13} M_\odot \). In the case of Lyman break galaxies, the halo occupation number is investigated in detail by Bullock, Wechsler, & Somerville (2002). Contrary to naive expectations, however, these authors find an occupation number with a minimum mass at the level we have indicated in equation (16), although there is a large range of allowed values (see, e.g., Fig. 3 of Bullock et al. 2002). The difference can be understood based on the fact that the minimum mass, as indicated by the halo occupation number, is not necessarily the one related to the mass above which halos are found to contain one observed object, but rather the minimum mass in which one galaxy may form. While analytically halos with mass below the minimum expected and above the minimum allowed by the halo occupation number form a fraction of a galaxy in each halo, in reality this is meant to indicate the fact that only a fraction of such halos form galaxies occasionally. On the other hand, all halos with mass above the expected minimum mass, which in this case is 2 orders of magnitude higher than the inferred minimum, are expected to host one or more galaxies. While the overall dilution of the clustering strength by a sample of low-mass halos, where galaxies are occasionally found, is not significant, these halos are important to describe the clustering properties of the high-redshift population (see Bullock et al. 2002 for a further discussion).

We note that our predictions related to IR galaxy clustering at redshifts around 2 are consistent with Daddi et al. (2003) measurements and find this to be adequate enough for the present paper. This is mainly due to the fact that we are mainly concerned with clustering properties of brightness galaxies at redshifts between 1 and 2, or in the case of K band, at magnitudes below 20 or so. As we discuss below, this is within the range of magnitudes for which upcoming experiments will be able to remove resolved sources from a clustering analysis. We are less concerned about the clustering properties of faint high-redshift galaxies, at \( z \sim 3 \), for one significant reason. The suggested measurement involves a clustering analysis of brightness fluctuations and not source counts, as in traditional studies of galaxy clustering at low redshifts. Thus, the clustering power spectrum from each redshift range, or bin, is weighted by the contribution to the absolute IRB produced by galaxies in that range or bin. Assuming fluctuations in the brightness trace source counts, we can roughly write this as \( C_l \sim \tilde{P}_{wl} \), where \( w_l \) is the angular power spectrum of source clustering. The fractional contribution to IRB from high redshifts is substantially less than that produced by nearby bright galaxies out to a \( z \) of 1 (e.g., Cambresy et al. 2001). In the K band, the fractional contribution to the IRB from galaxies with magnitudes fainter than \( \sim 20 \), which are expected to be at redshifts greater than 2, is at most at the level of 10%, if not less. On the other hand, the fractional contribution from Population III sources can be at the level of 50%, or more, especially under extreme models discussed in the literature (e.g., Salvaterra & Ferrara 2003). In relative terms, using the fractional contribution to the background, one can safely conclude that Population III clustering will dominate brightness fluctuations, when compared to that of faint galaxies at redshifts of around 3, by at least a factor of several tens or more since \( C_{F_{\text{Pop III}}}/(kz>20) \sim (I_{\text{Pop III}}/kz>20)^2 \). The extent to which this ratio can be known exactly depends on the relative strength of source clustering at redshifts of around 3 when compared to that at a redshift around 15. Although some subsamples of faint galaxies, such as the red galaxy \( J–K > 1.7 \) population of Daddi et al. (2003) at redshift of 3, are strongly clustered, we do not necessarily expect this clustering strength to be significantly higher than Population III sources such that \( C_{F_{\text{Pop III}}}/(kz>20) \). However, we note that this conclusion could substantially change if the \( z \sim 3 \) population is to be a significant contributor to the IRB and its fractional contribution is substantially higher than the 10% or so level implied in Cambresy et al. (2001).

The shot-noise power spectrum related to galaxies, as well as galactic stars, is more certain as a result of the availability of the number counts in certain wavelength ranges between 1 and 5 \( \mu m \). In Figure 2 (right panel) we plot the average number counts in the K band compiled from a variety of sources in the literature following Saracco et al. (2001) and Cambresy et al. (2001 and references therein). We use these number counts to establish the confusion noise in various experiments, summarized in the left panel of Figure 2, at near-IR wavelengths to understand the extent to which these
foreground sources dominate clustering studies through their shot-noise power spectrum. In both plots of Figure 2 we also illustrate the magnitude limit to which resolved sources can be removed at the 5 σ instrument noise level in various past and upcoming wide-field imaging data that are adequate enough for a clustering analysis of any first star contribution to the IRB.

2.3. Angular Power Spectra

In Figure 3 (top panels) we show the angular power spectra of galaxies and Population III sources as a function of wavelength at 1, 2, and 4 μm. These wavelengths are loosely chosen to represent the clustering behavior of sources in the IR regime. In the case of galaxies, the angular power spectrum is nearly a power law over a wide range of angular scales. The excess clustering at degree scales is related to the linear power spectrum that projects at low redshifts (z ~ 1). When interpreting this galaxy power spectrum, one should consider this as a representative case in which foreground galaxies, out to a redshift of ~1, are removed from the data such that nearby low-redshift galaxies are absent. In our calculations we allow for such a removal by simply cutting off the redshift distribution of galaxies at redshifts below 1. Note that our approach to describing clustering of galaxies at IR wavelengths is different from previous approaches. Instead of extending the measured power spectrum at low redshifts, such as from the Automated Plate Machine (APM) survey, to higher redshifts, or using numerically calibrated fitting functions for the non-linear power spectrum, such as from Peacock & Dodds (1996), which are valid only for dark matter, we have made use of the halo approach to calculate the clustering of galaxies directly.

Our approach to describing clustering is also independent of techniques that make use of galaxy number counts at IR wavelengths. While there is now detailed information related to galaxy counts, the exact redshift distribution of IR contributing galaxies still remains somewhat uncertain. As a simple approach, here we have made use of a distribution based on the normalized star formation rate. As long as this distribution is not dominated by galaxies at the high end of this redshift distribution, galaxy clustering follows a power law. Although we have not considered it here, our halo-based approach can be extended, under certain assumptions, to model number counts of IR galaxies similar to approaches considered in studies such as by Jimenez & Kashlinsky (1999).

The Population III stars, since they are present at redshifts greater than 10 and are highly biased with respect to the density field, trace the linear regime of clustering at tens of arcminute scales when projected on the sky today. As shown in the top panels of Figure 3, the Population III stars thus show the expected signature of an excess clustering, above the shot-noise power spectrum, at angular scales where the linear power spectrum has a peak associated with the matter-radiation equality. The difference between the linear clustering of galaxies at low redshifts and at angular scales of few degrees or more and Population III sources at high redshifts and at tens of arcminute scales can be understood based on the redshift evolution of the linear density field power spectrum alone. As discussed in Cooray et al. (2001), this evolution can in fact be used to constrain certain cosmological parameters although we do not pursue such possibilities here.

In Figure 3 we also show Population III source clustering angular power spectra for flux spectra involving $f_{esc} = 0$ and 1. We consider these two values as a potential bound on the uncertainties related to the Population III spectrum. While the clustering excess at a short wavelength like 1 μm is small in the case of $f_{esc} = 0$, with $f_{esc} = 1$ we find a significant contribution at the level comparable to 2 μm. This can be understood based on the Population III spectra we show in Figure 1 and results from the fact that when $f_{esc} = 1$, the short-wavelength cutoff is decreased to a wavelength value lower.
than for \( f_{esc} = 0 \). Additionally, as a result of the decrease in free-free emission with \( f_{esc} = 1 \), one finds a lower clustering amplitude at the high end of the IR wavelengths when compared to the case with \( f_{esc} = 0 \). While we have considered these two extreme cases, one expects \( f_{esc} \) to be of order a few tens of percent at most, such that a more realistic scenario will be more closely related to \( f_{esc} = 0 \) than \( f_{esc} = 1 \).

In addition to the spectrum, other uncertainties in physics of the Population III population lead to a highly indeterminate overall normalization for the angular power spectrum. These uncertainties, which are more significant than \( f_{esc} \) alone, are related to the Population III redshift distribution and biasing with respect to the linear density field. We bound these uncertainties and consider two estimates at the high and low end of the amplitude estimates based on the redshift distribution and biasing factor of dark matter halos, containing Population III sources, with respect to the linear density field. For each of these models, we also plot spectra for the two values of \( f_{esc} \), with solid lines showing \( f_{esc} = 0 \) and faint lines showing \( f_{esc} = 1 \), to illustrate wavelength dependence on the clustering signal. In addition to the clustering part (two-halo part of the power spectrum), we also estimate the shot-noise contribution related to the finite density of Population III stars. Note that this shot-noise contribution related to Population III stars is highly uncertain as a result of reasons discussed in the text, and we show the upper and lower ranges of allowed estimates, each in case of models A and B, respectively, in the top two panels. As shown in the bottom left panel, the foreground galaxies have an almost power-law–type clustering spectrum; here we only consider galaxies with redshifts greater than 1 under the assumption that the nearby foreground population can be resolved and removed from the data. Note that we do not show the galaxy shot-noise spectrum here, but we refer the reader to Fig. 4 for details. We compare galaxy and Population III clustering using model A and \( f_{esc} = 0 \) in the bottom right panel. As shown, background Population III stars are expected to show an excess clustering at tens of arcminute scales as a result of the strong biasing of dark matter halos that contain these stars at redshifts of order 15. In the bottom right panel, for comparison with thick long-dashed lines, we also show the total (one-halo plus two-halo) galaxy power spectrum at 2 \( \mu \text{m} \) when all galaxies (top curve) and galaxies at redshifts greater than 2 (bottom curve) are not removed from the imaging data. The extent to which Population III clustering can be studied depends on the ability of an experiment to resolve galaxies at redshifts out to at most 2 and remove them from the clustering analysis. [See the electronic edition of the Journal for a color version of this figure.]
order 20; varying the projection from redshifts 10 to 50 or so does not lead to a large change in this angular scale since the angular diameter distance does not change significantly in this range of redshifts.

In addition to a clustered signal, the finite density leads to a Poisson noise. These shot-noise power spectra for Population III sources, as a function of the wavelength, are also shown in Figure 3 (top panels). As we discussed earlier, the shot noise is more uncertain than the clustering power spectrum since the shot noise depends strongly on the rare and bright events while the clustering amplitude, as well as the cumulative background, does not. While the shot noise can easily be confused with the shot-noise power due to faint unresolved galaxies, among other foregrounds, the clustering excess, at tens of arcminute scales, however, provides a potentially interesting signature of Population III that can be observed.

As shown in the bottom right panel of Figure 3, the excess Population III clustering signature lies above the angular power spectrum of galaxies when wavelengths are in the range of \(\sim 1-2 \mu m\) and when nearby galaxies are removed from the data. Allowing for the presence of nearby galaxies leads to a slightly higher clustering signal for galaxies, which we have shown by a long-dashed line in the bottom right panel of Figure 3, and at a wavelength of 2 \(\mu m\). The removal of bright nearby galaxies leads to a decrease in the nonlinear clustering, but not necessarily the large angular scale clustering related to the two-halo part of the power spectrum since the large angular scale clustering is determined by the linear power spectrum scaled by the bias factor. For galaxy models considered here, the average bias factor is not strongly sensitive to the removal of nearby sources. In general, at wavelengths above a few microns, the galaxy clustering fully dominates IRB anisotropies, although one can improve the detection of Population III stars even at these wavelengths if most resolved galaxies are removed. For comparison, in Figure 3, the lower thick long-dashed line shows the expected clustering in the IRB at 2 \(\mu m\) from galaxies at redshifts greater than 2, suggesting that a removal down to this redshift range (or magnitudes at the level of 25 in \(K\) band) is essential to extract the clustering signal associated with the first star population. In general, to perform such a removal and to detect the Population III clustering signal at tens of arcminute scales, an experiment with subarcminute-scale angular resolution, but covering several tens of square degrees, is needed.

In Figure 4 we highlight the main aspects of Population III clustering. For reference, we also show the 2MASS clustering measurement in the \(K\) band at small angular scales (Kashlinsky et al. 2002) as a reference. Note that the Population III signature is present at tens of arcminute scales with an excess above the level expected from galaxies (assuming that foreground galaxies, out to redshifts \(\sim 1-2\), can be resolved and removed from the analysis). In addition to 2MASS, DIRBE (Kashlinsky et al. 1996a) and IRTS (Matsumoto et al. 2000) measured Population III clustering at large angular scales and detected excess fluctuations (discussed with respect to Fig. 6 below). In addition to the linear clustering of Population III sources, we also show the correction to the power spectrum associated with nonlinearities. While nonlinearities are expected to increase the power by an order of magnitude at very small scales \((l \sim 10^8-10^6)\), a comparison to Figure 3 reveals that at such small angular scales the clustering spectrum related to Population III sources is more likely to be dominated by their shot noise. At these same angular scales, the shot noise related to unresolved foreground galaxies also becomes important such that the individual contributions related to galaxies and first stars are not easily separable.

Based on our description of the Population III population and its clustering aspects, we find that the clustering amplitude resulting from emission by first stars is not fully responsible for the clustering signal detected in the 2MASS data at arcsecond angular scales. The 2MASS angular power spectrum is more likely due to a combination of the nonlinear clustering of unresolved galaxies and the shot noise associated with such sources. The number counts of galaxies in the \(K\) band now extend down to a magnitude of 24, and if any shot noise existed at the 2MASS level, its fluctuation amplitude from
pixel to pixel, as a result of the Poisson behavior, would correspond to a magnitude of ~22–23 in 1" pixels. The deep counts will easily be sensitive to such variations, and based on such an argument, we suggest that the Population III source shot-noise or clustering contribution must be below the large shaded region on the top right corner of Figure 4. This region essentially corresponds to 5 σ fluctuations down to a K-band magnitude of 24 in a 10" pixel. On the other hand, using measured galaxy counts alone, a significant shot noise is expected from unresolved sources that are below the instrumental detection limit but above the confusion limit (see, e.g., Lagache, Dole, & Puget 2003 for a description); we show this contribution, together with the shot noise from sources below the confusion limit, as a dot-dashed line in Figure 4. It is likely that a combination of this shot noise and the nonlinear clustering of the unresolved sources contributes to the 2MASS measurements, although a substantial contribution from nonlinear clustering of Population III sources may be expected in certain descriptions of the Population III population (Magliocchetti et al. 2003).

Since various possibilities may exist to explain 2MASS measurements (in terms of either galaxies, Population III sources, or in combination), we suggest that further investigation of small-scale clustering behavior of the IBR8 may not be adequate to separate these contributions. While one can extend a survey such as 2MASS to large angular scales by tiling subsequent fields, this can cause problems when combining adjacent fields to measure large-scale clustering beyond that of a single field of view (see discussions in Odenwald et al. 2003). The alternative approach is to image a wider area (of tens of square degrees) directly at slightly lower resolution and perform a clustering analysis, either after removing resolved stars and galaxies directly in such an image or after accounting for such sources based on higher resolution data of smaller fields, to see if there is any evidence for an excess as indicated by Figure 4 related to Population III sources. The best combination to search for the Population III signal appears to be a wide-field camera with fidelity on degree scales combined with a deep survey, such as Astro-F, to minimize foreground confusions.

The detection of excess Population III source clustering will be limited to the extent that foreground stars and galaxies can be removed. In surveys where the source removal is not significant, one expects a higher shot-noise contribution from galactic stars and foreground galaxies, and this shot-noise level could potentially be above the level of expected clustering of galaxies. To model the expected shot noise, we follow calculations by Lagache et al. (2003), and we show the case for a wide-field rocket-borne experiment (see Table 1 for details) in Figure 4. The instrumental noise power spectrum $C_l^\text{noise}$ related to each experiment is shown by a dashed line and is given by

$$C_l^\text{noise} = 4\pi f_{\text{sky}} \sigma_{\text{pix}}^2 f^2 (\Delta\theta)^2,$$

where $\sigma_{\text{pix}}$ is the noise per pixel (see Table 2), $N_{\text{pix}}$ is the number of pixels, and $\Delta\theta$ is the pixel scale or the scale of additional smoothing, if the latter is employed. As written, the noise contribution to the measured power spectra at each wavelength is described through standard CMB analysis approaches (Knox 1995) given the filtered Gaussian beam and the noise of the whole focal plane array. With the noise power spectrum, one can write the error, or standard deviation, for a measurement of the angular power spectrum at each multipole as

$$kC_l = \sqrt{\frac{2}{f_{\text{sky}}(2l+1)}} (C_l + C_l^\text{noise}),$$

where the term related to $C_l$ accounts for the cosmic variance associated with the finite sky coverage. When the power spectrum measurements are binned in the multipole space, the error associated with the binned power spectrum measurement is reduced further by a factor of $(\Delta l)^{1/2}$, where $\Delta l$ is the width of the bin. While we have not shown these errors in Figure 4, as a result of uncertain $C_l$ of Population III stars, the extent to which upcoming instruments probe the power spectrum related to $f_{\text{sky}}$ is shown by a combination of thick and thin lines when plotting $C_l^\text{noise}$.

For instrumental parameter values in Table 2 related to the rocket-based imaging data, one can remove point sources, at the 3 σ confidence level above the expected instrumental noise and confusion noise from sources below the resolution limit, down to a flux level of ~2 × 10−7 nW m−2. This corresponds to a K-band magnitude of 16.8. Using K-band counts for stars and galaxies, we expect a shot noise at the level of $C_l \sim 5 \times 10^{-8} \text{nW}^2 \text{m}^{-4} \text{sr}^{-1}$ as a result of unresolved and unsubtracted sources. We show this contribution by a dashed line labeled "SN" in Figure 4. In the case of a wide-field rocket-borne experiment, one expects a potential detection of the Population III population, at multipoles of order 103, if these stars follow our models at the optimistic end.

At the low end of our models, the Population III clustering amplitude is lower and the detection will strongly depend on the extent to which galaxies, with K-band magnitudes at the level of 20 and fainter, can be resolved and removed from the

### Table 1

**Instrumental Characteristics: Rocket-borne Experiment**

| Band | $\lambda$ (µm) | $\Delta l/\lambda$ | $\delta F$ (µJy (5 σ)) | $\sigma_{\text{pix}}$ (nW m−2 sr−1) |
|------|----------------|-----------------|--------------------------|--------------------------|
| J    | 1.25           | 0.24            | 530                      | 7.5                      |
| H    | 1.65           | 0.17            | 325                      | 6.5                      |
| K    | 2.2            | 0.16            | 205                      | 4.4                      |
| L    | 3.5            | 0.26            | 134                      | 1.8                      |

*Note:* The sensitivity, $\delta F$, and noise per pixel, $\sigma_{\text{pix}}$, assume a useful integration time of 200 s. The pixel size $\Delta l$ is 15". With pixel array formats of 1024 × 1024, observations cover a field of view of 4.27.

### Table 2

**Instrumental Characteristics: Astro-F**

| Band | $\lambda$ (µm) | $\Delta \theta$ (arcsec) | $\delta F$ (µJy (5 σ)) | $\sigma_{\text{pix}}$ (nW m−2 sr−1) |
|------|----------------|--------------------------|--------------------------|--------------------------|
| K    | 1.8–2.7        | 1.46                     | 1.3                      | 14.5                     |
| L    | 2.7–3.7        | 1.46                     | 1.2                      | 15.6                     |
| M    | 3.7–5.1        | 1.46                     | 1.9                      | 17.6                     |

*Note:* The sensitivity, $\delta F$, and noise per pixel, $\sigma_{\text{pix}}$, are based on 500 s of integration per pointing. The NIRCam of the Astro-F mission contains arrays of 512 × 412 pixels with an effective area of 102 arcmin2 per pointing. See Watarai et al. 2000 for more details.
data. While reaching such a low magnitude level for resolved sources is beyond the capability of the rocket-based experiment, as a result of restrictions in the integration time among others, such a study is clearly possible with Astro-F. At the 5σ confidence level above instrumental noise, Astro-F allows identification and removal of foreground sources down to a magnitude limit of 21.3 in the K band or 2.2 × 10^{-9} nW m^{-2} in 1.54 pixels. Using the same counts as before, we now find a shot-noise contribution at the level of $C_l \sim 2.0 \times 10^{-9} \text{nW}^2 \text{m}^{-4} \text{sr}^{-1}$; while the extent to which Population III star clustering can be detected with Astro-F depends on the clustering nature of unresolved galaxies below its detection limit, we find that Astro-F is useful for the study of Population III stars since it allows a large range in the clustering amplitude to be probed at arcminute scales as a result of the limited field of view, 10′ × 10′. Such an angular power spectrum measurement is useful to understand the extent to which clustering of Population III sources contributes to spatial fluctuations of the IRB given that we predict the presence of an excess at angular scales close to 10′ and above. In addition to the limited field of view, another drawback with Astro-F is the limited wavelength coverage restricted to K band and longer. To properly establish the clustering of Population III sources, one will require a study of the frequency spectrum related to the clustered component, and lack of observations below K band may complicate such an analysis. Furthermore, if coverage at shorter wavelengths existed, one can use such data for an extraction of the Population III contribution, based on its expected spectrum, such that the confusion between Population III sources and foreground galaxies is further reduced.

While observations in a single band, e.g., K band, are useful to understand Population III clustering, this can be better achieved with multiwavelength data as different channels can be combined to study clustering in each band and between bands. In Figure 5, as an illustration, we show the correlation coefficient related cross wavelength power spectra. This correlation coefficient is defined as

$$C(l_i, l_j) = \frac{C_{l_i}^{\lambda_i, \lambda_j}}{\sqrt{C_{l_i}^{\lambda_i, \lambda_i} C_{l_j}^{\lambda_j, \lambda_j}},}$$

under the assumption that the Population III sources follow our most optimistic assumptions (model A) and that the foreground galaxies are removed out to a redshift of 1. As shown, the total correlation coefficient ranges near unity at large angular scales, decreases at tens of arcminute scales, and increases back to almost unity at small angular scales. If the Population III contribution were to be nonexistent, the correlation coefficient would be more uniform while the decrease in the correlation at arcminute scales is associated with the Population III population. While an individual angular power spectrum measured at each wavelength was expected to show excess clustering due to Population III stars, the correlation coefficient from cross power spectra is expected to show a decrease at the same angular scales.

In addition to separating foreground galaxies and background first star contribution to spatial clustering in the IRB, multiwavelength observations are desirable to understand additional contaminants, among which zodiacal light is important. While spatial clustering properties of zodiacal light, or dust particles responsible for this emission especially at degree angular scales, are unknown, one expects a distinct spectral dependence to its emission given the size distribution of dust particles. With observations at several or more frequencies in the IR band, one can make use of this expected spectral shape of the zodiacal emission to partly remove its contribution and to perform a clustering study related to the zodiacal emission separately. This is similar to what has been done in estimating the total IRB in imaging data such as with DIRBE (e.g., Kashlinsky et al. 1996b). With imaging data such as from a rocket, another approach is to make use of the temporal variations related to the zodiacal emission. Since the zodiacal light is associated with micron-size particles near the Earth’s orbit, observations spread over the Earth’s orbit that span a few months or more are expected to view a different dust content toward the same background sky area at different observational epochs. This could give additional information to remove the zodiacal light, although we emphasize that the best option to remove this contaminant is to perform multiwavelength imaging over the IR band.

As a general consistency check on our models, instead of calculating the angular power spectrum, we also calculated the mean or the absolute background to be compared with observed results. Note that our Population III redshift distribution follows that of Santos et al. (2002), and since we use the same stellar spectrum as they calculated, the Population III star model produces the absolute background consistent with Figure 7 of their paper, to the extent that star formation efficiency is taken here as a free parameter, and by varying the minimum redshift at which Population III sources are found to 7 from 10; in our fiducial description of the Population III population, we use a somewhat higher minimum redshift for Population III sources and the IRB related to this model results in an overall decrease in the Population III contribution at wavelengths below ~1.5 μm. Our use of a higher minimum redshift for Population III sources is motivated by arguments related to the expected rapid transition from Population III to Population II sources, although we note that, based on
calculations for Population II star spectra by Tumlinson & Shull (2000), Population II stars will continue to make a substantial contribution to the IRB and background fluctuations at lower wavelengths. In the case of foreground galaxies, our model produces less than $\sim$40% of the background at 2 $\mu$m; this is consistent with, although potentially higher than, estimates based on galaxy number counts (Cambrésy et al. 2001). Note that a significant fraction of this contribution comes from sources at redshifts below $\sim$2; as expected from studies involving source counts, we expect galaxies out to redshift $\sim$2 to be the dominant confusion for a study such as this, and the extent to which such galaxies can be removed from the data will determine the ability to extract Population III clustering excess in the surface brightness fluctuations.

In order to consider the extent to which our clustering predictions are consistent with prior results, we compare our angular power spectra based on predictions with COBE DIRBE measurements, as a function of wavelength, by Kashlinsky et al. (1996a). These authors presented a measurement of $C(\theta = 0)$, the correlation function at zero lag, or rms fluctuations at the beam scale, at several wavelengths and using top-hat filtered maps. This quantity can be calculated from our power spectra using the fact that

$$C(0) \equiv \langle \delta I^2 \rangle_\sigma = \int \frac{d^4l}{(2\pi)^4} C_{\delta I}^{ij} W^2(l \sigma),$$

where $W(\lambda) = 2J_1(\lambda)/\lambda$ is the top-hat window function and $\sigma = 0.46$ is the angular scale of filtering. We summarize our results in Figure 6, where we plot $[C(0)]^{1/2}$ for the three cases involving galaxies and Population III sources (with $f_{esc} = 0$ and 1) and a comparison to COBE DIRBE measurements at 1.25, 2.2, 3.5, and 4.9 $\mu$m. Note that, in general, our model predictions, even at the optimistic end for Population III sources, suggest at least an order of magnitude lower values for $[C(0)]^{1/2}$ than measured. The prediction based on galaxies alone is similar and is consistent with previous estimates (e.g., Kashlinsky et al. 1996b). While we have some freedom to fit the rest of the data by varying our model parameters, we have not attempted to perform such an analysis here. The Population III sources, if at the low end of our prediction, suggest roughly 3 orders of magnitude smaller contribution to clustering seen in DIRBE data.

For comparison, in Figure 6 we also show results related to the fluctuation analysis in IRTS data as a function of the wavelength in the IR regime. These fluctuation measurements are made at the IRTS beam scale of $8' \times 20'$, which probes slightly smaller angular scales than the DIRBE measurements shown in Figure 6. There is considerable agreement, however, between DIRBE measurements and IRTS. It is unclear if these excess fluctuations are due to the shot noise associated with unresolved galactic stars, to the large beams of these experiments, or to an additional extragalactic component. Note that these observations probed large angular scales, while a substantial contribution from the Population III population is expected at tens of arcminute scales and, thus, these data do not necessarily aid in understanding the extent to which first stars are an important source of spatial fluctuations in IR wavelengths. In future experiments attempting to study Population III clustering, sufficient angular resolution to measure and remove bright stars and galaxies is essential.

3. RECONSTRUCTING THE STAR FORMATION HISTORY

Since the Population III redshift distribution is expected to closely follow the star formation history at redshifts between 10 and 30 or so, the anisotropy power spectrum of the IRB can be used to extract information on the star formation history, under the assumption of certain aspects related to the source clustering and the spectrum. For this purpose, we make use of multifrequency measurements including the cross power spectra between frequencies. To understand the extent to which star formation history can be studied with IRB fluctuations, we make use of the Fisher information matrix

$$F_{ij} = \langle \frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \rangle_x,$$

whose inverse provides the optimistic covariance matrix for errors on the associated parameters (e.g., Tegmark, Taylor, & Heavens 1997). In equation (22), $L$ is the likelihood of observing data set $x$, in our case the angular power spectrum, given parameters $p_1, \ldots, p_n$ involved in describing these data. Respecting the Cramér-Rao inequality (Kendall & Stuart 1969), no unbiased method can measure the $i$th parameter with standard deviation less than $(F_{ii})^{-1/2}$ if all other parameters are known exactly and less than $(F^{-1})^{1/2}$ if other parameters are estimated from the data as well.

For the present case involving $C_i$ values of the IRB, we can write

$$F_{ij} = \sum_i \left( l + \frac{1}{2} \right) f_{sky} \text{Tr} \left( C_i^{-1} \frac{\partial C_i}{\partial p_i} C_i^{-1} \frac{\partial C_i}{\partial p_j} \right),$$

where $f_{sky}$ is the sky fraction and $C_i$ are the cross power spectra.

![Figure 6](image-url)
FIRST STAR SIGNATURE

where \( f_{\text{sky}} = \Omega / 4\pi \) is the fraction of sky area covered by the imaging observations. Here \( \partial C_i / \partial p_i \) denotes derivatives of the power spectrum matrix, \( C_i^{\text{ij}} \), with respect to the parameter \( p_i \) at each multipole and \( f_{\text{sky}} \) is the fractional sky coverage of the observations. This matrix can be written as

\[
C_i^{\text{ij}} = \sum_k C_k^{\text{ij}} + \delta_{ij} C_i^{\text{noise}},
\]

where index \( i \) labels all contributions to the power spectrum and \( C_i^{\text{noise}} \) is the instrumental noise contribution to the observations. Here we assume that the noise contribution between maps is uncorrelated. In terms of individual components of the power spectrum (i.e., signal), we include contributions from the Population III stars (which is considered to be the primary component of astrophysical interest) and the foreground galaxies. The clustering properties of zodiacal light are unknown and may be a major source of uncertainty in studies of early star formation with IRB fluctuations, but as a result of the unknown properties related to its spatial clustering, we do not consider its impact in this estimate. Moreover, the proposed multiwavelength data may allow us to reduce the confusion related to the zodiacal emission.

In terms of parameters of interest, we study how well \( j_i(z) \) can be studied with clustering information. We split this quantity into four redshift bins in the range of 10–30 and extract the star formation rate in those bins. Note that, in the strictest sense, what one extracts is \( \langle b(M) \rangle j_i(z) \), which includes Population III spectrum information and halo bias in addition to the star formation rate. Here, for simplicity, we assume that both the spectrum and halo bias are known although one can constrain a combined quantity instead, if the latter two parameters are assumed to be unknown. Note that, in general, the spectrum, or at least the shape of the spectrum with an arbitrary normalization, can be assumed to be known a priori.

We calculate the Fisher matrix to a maximum \( l \) of 5000 such that the clustering of Population III stars is in the linear regime (two-halo term) and is not shot noise dominated (as shown in Fig. 3, we do not expect the Population III shot-noise domination to be important until \( l \sim 10^4 \)). In addition to the astrophysical parameters related to the star formation rate, we also include cosmological parameters that define the linear matter power spectrum, namely the power spectrum normalization, tilt, \( \Omega_{m}h^2 \), and \( \Omega_{b}h^2 \). We set priors for these parameters as known from current WMAP results and its extension related to information from other large-scale structure surveys. This information, in general, leads to a power spectrum at accuracy at the level of a few percent and is not a source of major concern for these studies.

We summarize our results in Figure 7, where we plot the star formation rate of the Population III population and reconstructed errors on this rate based on clustering information for the parameters as tabulated in Table 2. We assume a total number of 512 pixels per band and consider two pointings with each pointing of 453 field of view. At redshifts of order 10, where most of the Population III contribution arises, the star formation rate is constructed with accuracies of order \( \sim 0.2 M_{\odot} \text{yr}^{-1} \text{Mpc}^{-3} \), while at higher redshifts, \( \sim 30 \), the clustering information is no longer sensitive to the star formation rate as a result of the decrease in fractional distribution of Population III stars at such high redshifts.

Note that we have assumed that Population III sources follow the optimistic model with a clustering angular power spectrum given by the upper curve shown in Figure 4. If clustering were to be lower, then the reconstructed star formation errors increase. In such a scenario, observations with Astro-F can be used to investigate the existence of Population III clustering at arcminute scales to a lower flux level with a better removal of the foreground galaxies that may confuse the detection of Population III sources.

4. SUMMARY

The recent results related to CMB anisotropies suggest that the universe was reionized at a redshift around 17 with an optical depth for electron scattering of \( 0.17 \pm 0.04 \). Such an early reionization could arise through the ionizing radiation emitted by metal-free Population III stars at redshifts of 15 and higher. We discuss the contribution to the IRB surface brightness anisotropies from such a generation of early star formation. We have shown that the spatial clustering of these stars at tens of arcminute scales can potentially generate a significant contribution to the angular power spectrum of the IRB.

Note that measurements of IRB spatial clustering already indicate both an overall excess (e.g., Kashlinsky et al. 1996b) and a specific excess at 100° scales (Matsumoto et al. 2000, 2001) when compared to other angular scales. The models based on a contribution from Population III stars indicate that one should expect a clustering excess at few tens of arcminutes related to the overall projection of the linear power spectrum at redshifts between 10 and 30. While the amplitude of the clustering excess is model dependent, the angular scale at which this clustering is expected is more precise given that it is simply a reflection of the projected linear power spectrum at redshifts of 10–30. The direct detection of the amplitude of clustering excess, as a function of wavelength, can be used as...
a measure of the Population III star formation rate, under the assumption of a spectrum and clustering bias, or the combined quantity involving the star formation rate and bias weighted by the source flux. As an initial attempt in this direction, we have considered the astrophysical uses of a wide-field camera optimized for the detection of the first star signature and have suggested that studies related to Population III stars can be further improved with near-infrared observations with the upcoming Astro-F mission. The best opportunity to study Population III stars via IRB spatial fluctuations will come from a combined analysis of Astro-F and wide-field rocket-borne data in the same area on the sky. The higher resolution data from Astro-F can be used to clean low-resolution, but significantly wider field of view, data from the rocket experiment, such that clustering studies can be extended even below our pessimistic level of the Population III clustering considered.

This work is supported by the Sherman Fairchild Foundation and DOE DE-FG 03-92-ER40701 (A. C.), as well as by an NSF Astronomy and Astrophysics Postdoctoral Fellowship (B. K.). A. C. thanks the Aspen Center for Physics for hospitality while this work was initiated. We thank Mike Santos for providing us with Population III spectra and for useful discussions, Sasha Kashlinsky for an electronic table of the 2MASS power spectrum and details regarding this measurement, Aparna Venkatesan for useful comments and suggestions, and an anonymous referee for helpful suggestions.

REFERENCES

Abel, T., Bryan, G. L., & Norman, M. L. 2000, ApJ, 540, 39
Abel, T., Bryan, G. L., & Norman, M. L. 2002, Science, 295, 93
Bromm, V., Coppi, P. S., & Larson, R. B. 1999, ApJ, 527, L5
Bromm, V., Coppi, P. S., Larson, R. B. 2002, ApJ, 564, 23
Bullock, J. S., Weinberg, D. H., & Somerville, R. S. 2002, MNRAS, 329, 246
Cambrésy, L., Reach, W. T., Beichman, C. A., & Jarrett, T. H. 2001, ApJ, 555, 563
Cen, R. 2003, ApJ, 591, 12
Cooray, A., Hu, W., Huterer, D., & Joffre, M. 2001, ApJ, 557, L7
Cooray, A., & Sheth, R. P. 2002, Phys. Rep., 372, 1
Daddi, E., et al. 2003, ApJ, 588, 50
Eisenstein, D. J., & Hu, W. 1999, ApJ, 511, 5
Fan, X., et al. 2002, AJ, 123, 1247
Fukugita, M., & Kawasaki, M. 2003, MNRAS, 343, L25
Hauser, M., & Dwek, E. 2001, ARA&A, 39, 249
Jimenez, R., & Kashlinsky, A. 1999, ApJ, 511, 16
Kashlinsky, A., Mather, J. C., & Odenwald, S. 1996a, ApJ, 473, L9
Kashlinsky, A., Mather, J. C., Odenwald, S., & Hauser, M. 1996b, ApJ, 470, 681
Kashlinsky, A., & Odenwald, S. 2000, Science, 289, 246
Kashlinsky, A., Odenwald, S., Mather, J., Skrutskie, M., & Cutri, R. 2002, ApJ, 579, L53
Kendall, M. G., & Stuart, A. 1969, The Advanced Theory of Statistics, Vol. II (London: Griffin)
Knox, L. 1995, Phys. Rev. D, 52, 4307
Knox, L., Cooray, A., Eisenstein, D., & Haiman, Z. 2001, ApJ, 550, 7
Kogut, A., et al. 2003, ApJS, 148, 161
Lagache, G., Dole, H., & Puget, J.-L. 2003, MNRAS, 338, 555
Limber, D. B., 1954, ApJ, 119, 655
Madau, P., & Pozzetti, L. 2000, MNRAS, 312, L9
Magliocchetti, M., Salvaterra, R., & Ferrara, A. 2003, MNRAS, 342, L25
Matsumoto, M., et al. 2000, in ISO Surveys of a Dusty Universe, ed. D. Lemke, M. Stickel, & K. Wilke (Berlin: Springer), 96
Matsumoto, M., et al. 2001, in IAU Symp. 204, The Extragalactic Infrared Background and Its Cosmological Implications, ed. M. Harwit & M. G. Hauser (San Francisco: ASP), 101
ERRATUM: “FIRST STAR SIGNATURE IN INFRARED BACKGROUND ANISOTROPIES”
(ApJ, 606, 611 [2004])

Asantha Cooray, James J. Bock, Brian Keating, Andrew E. Lange, and T. Matsumoto

As a result of an error at the Press, Brian Keating’s name was incorrectly given as Brian Keatin on the title and contents pages for our article.
The Press sincerely regrets this error.